Characterization of Neuronal-Specific Tra2b Knock-Out Mice and Identification of Tra2b Splicing Targets

Inaugural-Dissertation

zur

Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Universität zu Köln

vorgelegt von

Markus Storbeck

aus Lünen

Köln 2014

This Doctoral Thesis “Characterization of Neuronal-Specific Tra2b Knock-Out Mice and Identification of Tra2b Splicing Targets” was performed at the Institute of Human Genetics, Institute of Genetics and Center for Molecular Medicine Cologne (CMMC) of the University of Cologne from November 2009 to 2013.

Berichterstatter: Prof. Dr. Brunhilde Wirth

Prof. Dr. Elena Rugarli

Tag der mündlichen Prüfung: 20.01.2014

Für meine Eltern

Acknowledgements

First, I would like to thank my supervisor Professor Dr. Brunhilde Wirth for giving me the opportunity to work on this highly interesting and challenging project, for her counseling, lively and motivating scientific discussions and for always having an open ear. Furthermore, I would like to thank her for her generous support to attend international meetings and further scientific training. Her positive way of thinking and her dedication to science always was a source of confidence for me.

I thank my examiner Professor Dr. Elena Rugarli and Professor Dr. Günter Schwarz for taking the chair in my thesis defense.

A big THANKS to all the present and past members of the 'SMA group', and the entire Institute of Human Genetics Cologne. Very big and special thanks go to Kristina Hupperich for her outstanding technical support, for giving up her usual breathing habits at the microtome and for always taking some time to show me where to find this Tra2b antibody ☺. Thanks to Markus Rießland for a lot and well appreciated scientific advice, and nonetheless to Lutz Garbes, Bastian Ackermann and Mohsen Hosseini for all these scientific and less-scientific discussions. Further thanks go to Miriam Jakubik, again to Kristina, and Ylva Mende for introducing me into the business all around the mice. Thanks to all the group members from the 'front lab' for giving me a *warm* welcome back then.

I thank Radu Wirth for writing and running his custom-made scripts to locate thousands of binding motifs. I would further like to thank Kesavan Meganathan, John Antonydas Gaspar and Agapios Sachinidis for a good collaboration and for the time they invested to perform the exon array analysis. Further, I thank David Elliott and Sushma Grellscheid for a great collaboration, good discussions and for lots of scientific input.

I want to thank my parents and my grandmother for their everlasting support and love. Your encouragement, motivation and confidence made my studies and this thesis possible.

At last I thank Lilian Martínez: Thanks for just being *you* and being always there for me when I need you most. Your patience and contagious calmness helped me through these days of rush.

Table of Contents

Directory of figures

Directory of tables

List of abbreviations

Abbreviation	Explanation
µ	micro
3’	3-prime (for nucleic acids)
3’ss	3-prime splice site
5’	5-prime (for nucleic acids)
5’ss	5-prime splice site
AD	Alzheimer’s disease
AP	alkaline phosphatase
approx.	approximately
ATP	Adenosine triphosphate
BP	branch point
bp	base pairs
BSA	bovine serum albumin
cAMP	cyclic adenosine monophosphate
cDNA	copy DNA
CDS	coding sequence
chr.	chromosome
DAPI	4',6-diamidino-2-phenylindole
dATP	deoxyadenosine triphosphate
dCTP	deoxycytidine triphosphate
dGTP	deoxyguanosine triphosphate
DNA	deoxyribonucleic acid
dNTP	deoxynucleoside triphosphate
dpc	dies post coitum (days after copulation)
ds	double strand
dTTP	deoxythymidine triphosphate
dUTP	Deoxyuridine triphosphate
e.g.	for example (*exempli gratia*)
EGFP	enhanced green fluorescent protein
ER	endoplasmic reticulum
et al.	and others (*et alii*)
etc.	*et cetera*

Abbreviation	Explanation
FCS	fetal calf (bovine) serum
fl	flox / floxed
FTDP-17	Frontotemporal dementia and parkinson linked to chr. 17
g	gramm
h	hour(s)
hnRNP	heterogenous ribonucleoprotein particle
HSP	heat shock protein
i.e.	that is (*id est*)
Ig	immunoglobulin
kb	kilobases
kDa	kilodalton
LB	Luria Bertani
M	molar
m	milli
max.	maximum
min	minute(s)
min.	minimum
miRNA	microRNA
mRNA	messenger RNA
n	nano
NBD	nucleotide binding domain
NMD	nonsense-mediated decay
p	pico
p	probability
p.a.	per analysis
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
poly Y	polypyrimidine tract
PND	Postnatal day
PTC	premature termination codon
RNA	ribonucleic acid
RNAi	RNA interference

Abbreviation	Explanation
RNP	ribonucleoprotein
rpm	revolutions per minute
RRM	RNA recognition motif
rRNA	ribosomal RNA
RT	reverse transcription
s	second(s)
s.e.m.	standard error of the mean
SDS	sodium dodecylsulfate
shRNA	small hairpin RNA
siRNA	small interfering RNA
snRNA	small nuclear RNA
snRNP	small nuclear ribonucleic particle
SR	Serine/arginine
SRPK	SR-protein kinase
ss	single strand
ST	sense strand
TBE	Tris buffered EDTA
TBS	Tris buffered saline
TBST	Tris buffered saline Tween-20
tg	transgene
T_m	melting temperature (oligonucleotides)
URL	Uniform Resource Locator
UTR	untranslated region
UV	ultraviolet
WB	Western Blot
WT	Whole transcriptome
wt	wild type

1. Introduction

The splicing factor TRA2B (SFRS10, **s**plicing **f**actor arginine(**R**)-serine(**S**)-rich **10**) belongs to the family of SR-splicing factors. It is involved in alternative splicing processes (1.1) and can modulate splice site selection by direct sequence-specific binding to RNA in a concentration-dependent fashion (Nayler et al. 1998; Clery et al. 2011). TRA2B is involved in splicing processes related to neuronal function (Stamm et al. 1999; Kondo et al. 2004; Glatz et al. 2006), development (Chen et al. 2004; Eswarakumar et al. 2005), vascularization (Shukla and Fisher 2008), obesity (Pihlajamaki et al. 2011) and probably spermatogenesis (Venables et al. 2000; Dreumont et al. 2010; Grellscheid et al. 2011a). It has achieved major attention in our laboratories as it is directly responsible for the splicing of *SMN* exon 7 (Hofmann et al. 2000; Hofmann and Wirth 2002). Functional loss of the *SMN1* gene causes spinal muscular atrophy (SMA), but the severity of the disease is primarily determined by the copy number of the *SMN2* gene that is present in every SMA patient (Wirth et al. 2006). Splicing correction of *SMN2* transcripts to promote exon 7 inclusion is the most promising approach for SMA therapy (1.6.1), as it has the potential to fully restore levels of functional SMN protein.

TRA2B has been demonstrated to promote inclusion of exon 7 into the *SMN2* transcript in a concentration-dependent manner (Hofmann et al. 2000). It was further shown that histone deacetylase inhibitors capacitate *SMN2* transcripts to achieve strongly increased exon 7 inclusion rates (Riessland et al. 2006; Garbes et al. 2009; Hauke et al. 2009; Riessland et al. 2010), which was facilitated by transcriptional upregulation of TRA2B (Brichta et al. 2003). However, higher TRA2B protein levels supposedly alter the splicing patterns of numerous transcripts. As only a limited number of transcripts are known to be processed by TRA2B, this study was aiming at the identification of TRA2B splicing targets *in vivo* using a mouse model. Previously generated *Tra2b* knock-out mice are early embryonically lethal (Mende et al. 2010). In this study a conditional *Tra2b*-knock-out mouse of the central nervous system was generated to allow identification and analysis of transcripts targeted by Tra2b in neurons, which the primary tissue of SMA disease pathology.

Splicing is a complex and highly regulated process and applies to the majority of multi-exon transcripts. Therefore, the following chapters focus on the control and different kinds of alternative splicing, as well as the basic principle of the splicing reaction and its regulation. Furthermore, known Tra2b-related splicing processes but especially its involvement in *SMN* exon 7 splicing will be explained in detail.

1.1. Alternative splicing

Alternative splicing is a mechanism regulating the selection or inclusion of exons into a mature mRNA. During the splicing process intronic sequences are removed from pre-mRNAs linking exons to form the mature mRNA. Thus, multiple mRNAs can emerge from a single pre-mRNA by combinatorial inclusion and exclusion of exons. This increases the informational content of the genome allowing a relatively small number of genes to encode a much larger number of mRNAs and proteins (Nilsen and Graveley 2010). About 95 % of all transcripts encoded by multiple exons undergo alternative splicing (Pan et al. 2008) and numerous splicing events are evolutionarily highly conserved (Kan et al. 2005). The most intriguing alternative splicing event is probably the generation of more than 38,000 possible distinct mRNAs from the *Dscam1* gene in *Drosophila melanogaster*. In *Dscam1* exons are selected out of 4 different exons clusters allowing this huge diversity of transcripts that is necessary for regular neuronal circuitry formation (Neves et al. 2004; Hattori et al. 2008).

As much as functional variability emerges from conserved alternative splicing events, erroneous splicing is a major contributor to disease. In the brain there is a higher occurrence of alternative splicing than in any other tissue (Yeo et al. 2004; Johnson et al. 2009; Zheng and Black 2013). Thus, 50 percent of known debilitating mutations in RNA-binding proteins cause neuronal-related disease (Licatalosi and Darnell 2006; Cooper et al. 2009).

Alternative splicing and the ways of combinatorial exon selection are very versatile (Figure 1) and can enable, alter, or abolish the function of the protein products in many ways (Faustino and Cooper 2003). Selection of different initial exons can lead to alternative translational start sites (different N-termini) or results in mRNAs with differently long 5’ UTRs (Figure 1 (1)). The most common type of alternative event is the cassette exon that is either included or excluded from a transcript (Figure 1(2)) (Stamm et al. 1999; Stoilov et al. 2004; Keren et al. 2010). Other splicing patterns for transcript-internal exons include alternative 3’ and 5’ splice sites (Figure 1(3,4)), retention of introns within the mature mRNA (Figure 1(5)) and mutually exclusive exons (Figure 1(6)). On protein level each of these events can cause inframe deletions or insertions, a shift of the reading frame, or the insertion of a stop codon. Generation of premature termination codons (PTCs) more than 50 nucleotides upstream of the terminal intron might target a transcript for nonsense-mediated decay (NMD), which is an important regulatory mechanism for the regulation of gene expression on a posttranscriptional level (Lareau et al. 2007; McGlincy and Smith 2008; Nilsen and Graveley 2010). Furthermore, the selection of alternative terminal exons can result in competition of different

polyadenylation or cleavage signals resulting in different C-termini or differently long 3' UTRs (Figure 1(7,8)).

Alternative exon selection is based on the interplay of different cooperating or antagonizing splice factors (1.3) and on the recognition of alternative splice sites. Mutations in splicing-determining sequences can therefore significantly alter splicing of exons and cause disease (Faustino and Cooper 2003; Cartegni et al. 2006; Liu and Gong 2008). Regulation of alternative splicing is in part tissue specific and also differentially regulated during development. This allows generation of protein products with extended or altered functionality from the same gene depending on the cell type or the developmental stage (Stamm et al. 1999; Kondo et al. 2004; Glatz et al. 2006). Moreover, alternative splicing has been shown to mediate cellular responses to different types of stress (Jacob et al. 2013; Wong et al. 2013).

Besides the sequence with splicing motifs and the trans-factors by which those are recognized, alternative splicing is also regulated epigenetically via chromatin structure and histone modifications (reviewed in (Luco et al. 2011)). Commitment of splicing factors and spliceosmal components (1.2) to exon splicing occurs co-transcriptionally, though completion of the splicing reaction mostly occurs after transcription (Tardiff et al. 2006). Binding of spliceosomal components and trans-factors in part depends on RNA topology (i.e. secondary structure) which can bury or expose sequence recognition motifs (Luco et al. 2011). In turn secondary structure formation underlies transcriptional speed and is biased by RNA polymerase performance and pausing (Kornblihtt 2005). Such effects are based on chromatin structure and especially histone modifications that impact transcription speed. Moreover, histone modifying proteins have been reported to physically interact with the U1 and U2 spliceosomal components (1.2) (Martinez et al. 2001; Gunderson and Johnson 2009). Further, nucleosomes themselves do not appear to be randomly positioned along genes but are enriched at exon-intron junctions, thus marking exons. A stretch of ~147 nucleotides is wrapped around a single histone octamer and it is speculated that the size of the average exon being similar to this is not a coincidence (Andersson et al. 2009; Schwartz et al. 2009; Chodavarapu et al. 2010).

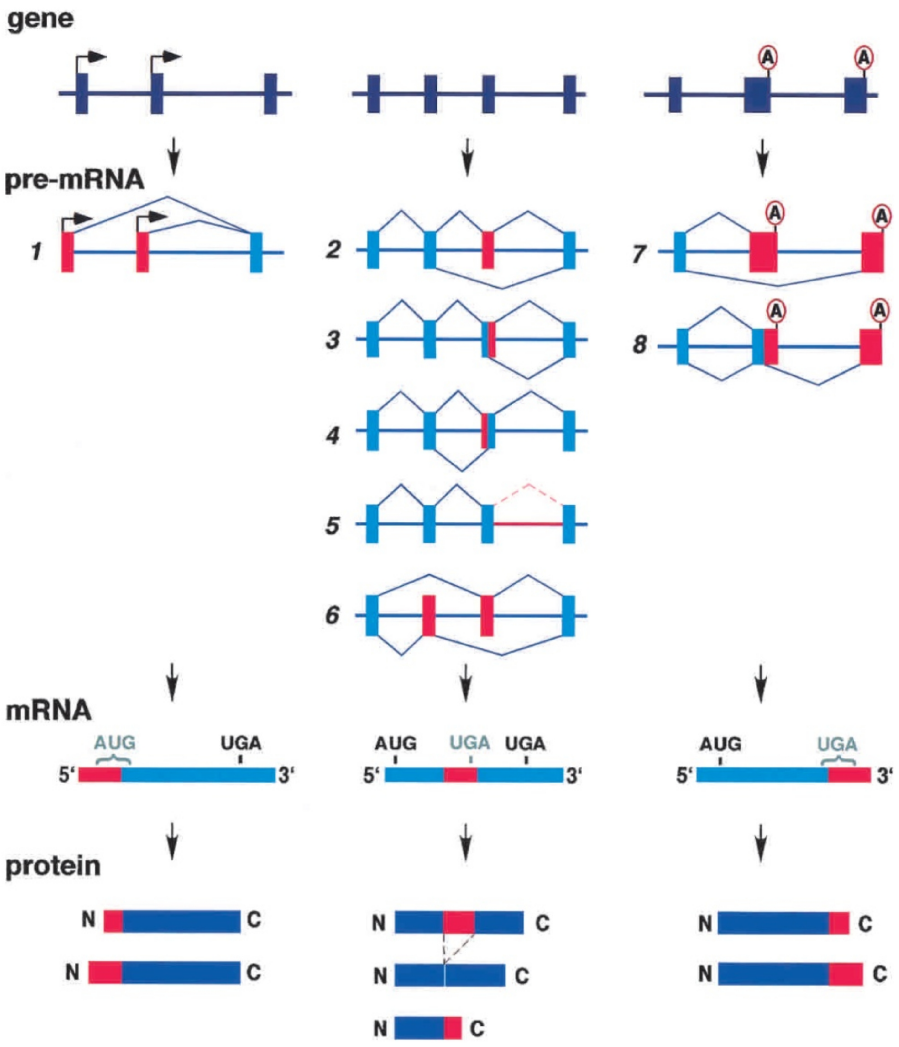

Figure 1 | Mechanisms of alternative splicing.
Alternative splicing modifies mRNA in different ways, thus significantly impacting the protein-coding information. Common mechanisms of alternative splicing are: (1) Selection of alternative initial exons (2) Cassette exons (3) alternative 5'ss (4) alternative 3'ss (5) intron retention (6) mutually exclusive exons (7) selection of alternative terminal exons and (8) competition of 5'ss with upstream poly-adenylation signals. (From (Faustino and Cooper 2003; Manley and Krainer 2010))

1.2. The splicing reaction and the spliceosome

Maturation of pre-mRNA to mRNA that can be translated into a peptide requires several processing steps. Maturation includes 5'-capping with methylated guanosine triphosphate, 3'-polyadenylation and splicing. During splicing long, non-coding sequence elements called introns are removed from the transcript, thus linking exons to each other forming the mature mRNA. This process is facilitated by the spliceosome, which is a multi-megadalton ribonucleoprotein (RNP) complex. In eukaryotes there are 2 types of spliceosomes as there are two types of introns that need to be recognized: The U2-dependent spliceosome (major spliceosome, reviewed by (Will and Luhrmann 2011)) and the less abundant U12-dependent spliceosome (minor spliceosome, reviewed by (Turunen et al. 2013)). Accordingly, these splicesomes recognize and process U2-type and U12-type introns, respectively. Both types of introns are different as to their recognition sequences. Sequence elements (cis-elements) defining an intron to be recognized by the spliceosome are limited to short conserved sequence stretches: The 5' and the 3' splice site (5'ss and 3'ss), the branch point (BP) and the polypyrimidine tract (poly Y), which is present in higher eukaryotes downstream of the branch site (Chen and Manley 2009) (Figure 2). The branch site is usually located 18-40 nucleotides upstream of the 3'ss.

Introns are removed from pre-mRNAs by two transesterification reactions. As shown in Figure 2A, the 2' hydroxyl group of the branch point adenosine performs a nucleophilic attack on the intronic 5'ss. Thereby, the 5'ss is cleaved and ligated to the branch point adenosine forming a lariat structure. The second transesterification is an attack by the 3' hydroxyl group of the exon to the 3'ss of the intron. During this reaction the 3' and 5' ends of the exons are linked forming the mRNA (Will and Luhrmann 2011).

Cis-elements found on the pre-mRNA cannot facilitate the splice reaction by themselves but require numerous trans-acting factors to assemble and bind the pre-mRNA. This renders proper positioning of the splice and branch sites within the spliceosomal complex to allow catalysis. The U2 spliceosome is assembled from the snRNPs (***s**mall **n**uclear **r**ibo**n**ucleic **p**articles*) U1, U2, U5 and U4/U6. Each snRNP mainly consists of a snRNA (***s**mall **n**uclear **RNA***), a group of 7 common Sm proteins as well as a group of unique (particle-specific) proteins. Assembly of the spliceosome is a highly ordered mechanism that is based on protein-protein, protein-RNA and RNA-RNA interactions (Collins and Guthrie 2000). Spliceosome assembly and catalysis of the splicing reaction occurs stepwise, as snRNA secondary structures, snRNP conformation and interactions connecting the spliceosome undego fundamental changes during the splice reaction (Will and Luhrmann 2011) (Figure 2B). The first step of spliceosomal

assembly is the binding of U1 snRNP to the 5’ splice site (Seraphin and Rosbash 1989). Other non-snRNP molecules like U2AF/65 and U2AF/35 (*U2 auxiliary factor*) bind to the branch site and the polypyrimidine tract. In the next step, U2 snRNA associates with the branch site via interaction with U2AF forming the pre-spliceosome (Ruskin et al. 1988). In contrast to U1 association, this step is ATP-dependent. U2 snRNP incorporation makes the branch site adenosine available for the nucleophilic attack of the 5’ splice site, which is the first out of 2 transesterification reactions during the splice process (Query et al. 1994; Furman and Glitz 1995). U4/U6 and U5 snRNP enter the assembling spliceosome (ATP-dependent) as a pre-assembled snRNP complex forming the catalytic - but still inactive - B complex. Massive rearrangements lead to the dissociation of U1 and U4 snRNP generating the catalytically active B complex (Will and Luhrmann 2011). Starting from this complex both splicing reactions are performed involving gross conformational rearrangements of the spliceosome (Umen and Guthrie 1995a; Umen and Guthrie 1995b). After completion the spliceosome dissociates and single snRNP components become available for a new cycle of spliceosome assembly.

The mode of spliceosome assembly largely depends on the intron length. For introns shorter than approx. 250 nucleotides spliceosome assembly occurs across the intron. For much longer introns – as it is normality in higher eukaryotes – the initiation of spliceosome assembly occurs across the exon, which in most cases is rather short. This principle is referred to as *exon definition* (Figure 2C). Here, U2AF and U2 snRNP associate to the intron upstream of the defined exon, while U1 snRNP binds to the 5’ss of the downstream intron (Berget 1995; Fox-Walsh et al. 2005). Both, the 3’ss upstream of an exon and the 5’ss downstream of an exon have been shown to be equally crucial for efficient exon recognition by the spliceosome and thus for efficient exon inclusion into a transcript (Shepard et al. 2011).

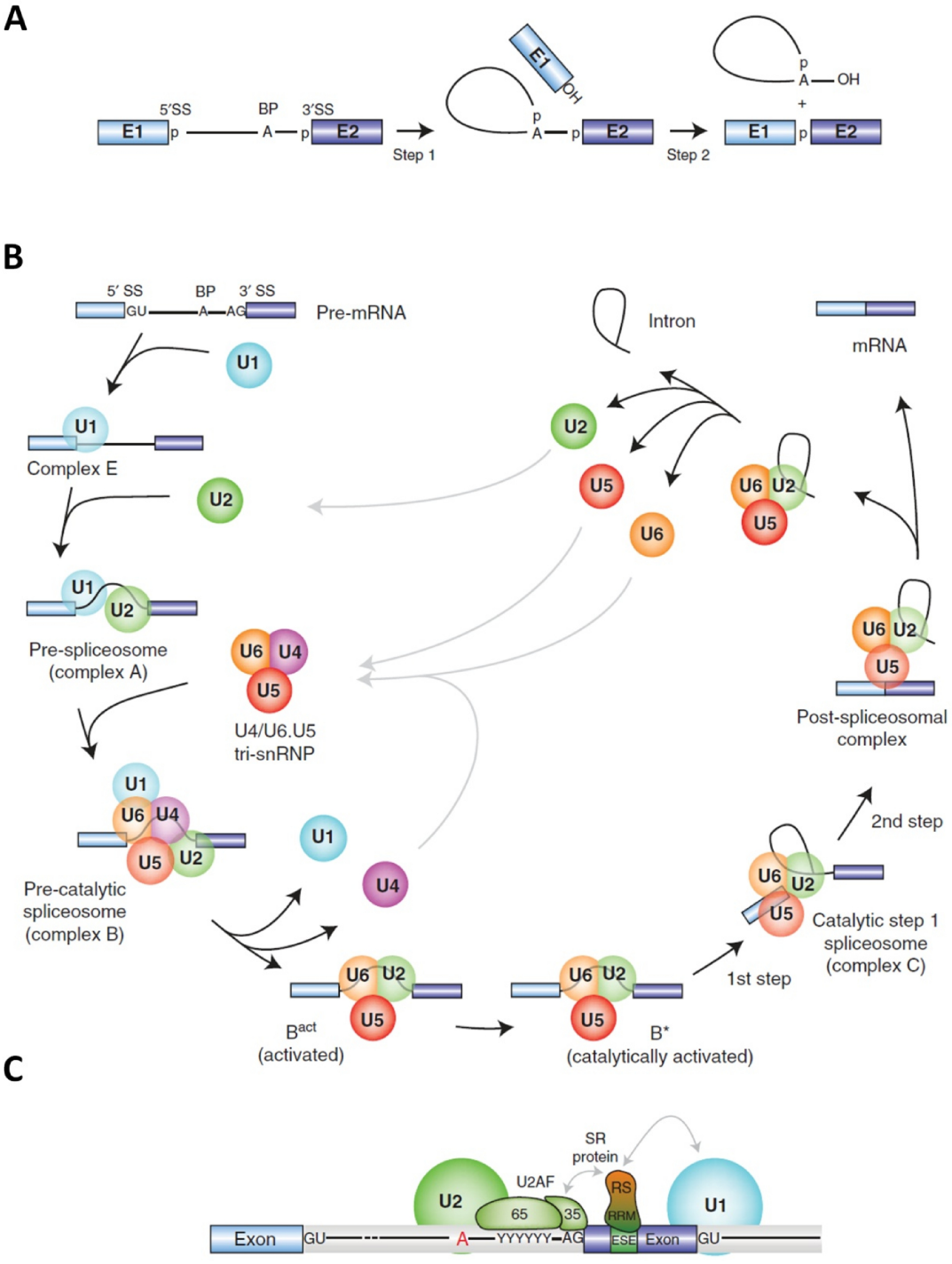

Figure 2 | The splicing reaction and the U2-dependent spliceosome.
(A) Introns contain conserved sequence motifs: the 5′ and 3′ splice sites, the branch point and the polypyrimidine tract that allow recognition of the intron by the spliceosome. Introns are removed from the pre-mRNA in 2 transesterification reactions. (B) The assembly of the spliceosome occurs stepwise and is based on multiple RNA-RNA, RNA-protein and protein-protein interactions that undergo massive rearrangements during the splicing process. Catalytic removal of the intron is facilitated by the active B complex incorporating snRNPs U2, U5 and U6. (C) In higher eukaryotes spliceosome assembly initiates due to exon recognition as the intermittent introns are usually very long. The spliceosome assembles spanning an exon by recognizing the 3'ss of the upstream intron and the 5'ss of the downstream intron. A, adenosine; p, phosphate; YYYYY, polypyrimidin tract; U2AF, U2 snRNP auxiliary factor; RRM, RNA recognition motif; ESE, exonic splicing enhancer. (Modified from (Mende 2008; Will and Luhrmann 2011))

1.3. SR proteins and SR-related proteins

Primary transcripts contain sequence motifs recognized by the spliceosome, which allows constitutive splicing events. Even though splicing regulatory cis-elements can be found in virtually every exon and intron, the sequence motifs are relatively degenerated and might be insufficient to promote effective splicing by the spliceosome. Besides the spliceosome (1.2) many other accessory splice factors (trans-factors) are involved in the orchestration of most splicing events (Faustino and Cooper 2003). The trans-factors use exonic or intronic sequence motifs that are distinct from those used by the spliceosome. Depending on their effect on the splicing reaction these sequences are referred to as being enhancing (inclusion) or silencing (skipping) and are therefore termed **e**xonic/**i**ntronic **s**plicing **e**nhancers/**s**ilencers (ESE/ESS, ISE/ISS). The number and position of such splicing regulatory elements as well as the modulating trans-factors that are recruited to these positions convey significant flexibility to exon definition (Hertel 2008; Shepard et al. 2011).

Typical trans-factors that modify the splicing reaction belong to the families of hnRNPs (*heterogeneous ribonucleoprotein particles*) or SR- and **SR**-**r**elated (SRr) (**s**erine-a**r**ginine-rich) splice factors. Members of both families associate to the primary transcript co-transcriptionally and help regulating the recruitment of spliceosomal components (Blencowe et al. 1999). Both groups may either exert a positive or a negative effect on exon recognition – they may cooperate or antagonize each other - depending on their binding positions (Busch and Hertel 2012). Members of the SR and SRr family do not only participate in alternative splice processes but rather make a general contribution to most constitutive splicing events as well (Graveley 2000). SR proteins are evolutionarily conserved and contain 1 or 2 N-terminal RNA recognition motifs (RRM) and a serine/arginine-rich (SR) region that is variable in length but comprises at least 50 amino acids with more than 40 % SR content (Busch and Hertel 2012; Risso et al. 2012). There are 12 SR-proteins described in humans (Figure 3) which mostly retain a similar domain composition. Only TRA2B has a unique composition with a single RRM flanked by two RS-rich domains (also in 1.4). The number of SR-proteins varies between organisms and it has been proposed that the number of SR-proteins increases with the complexity of alternative splicing (Busch and Hertel 2012). In humans every gene contains on average 8 introns. Only 4 % of genes in yeast contain introns and classical SR-rich proteins are fully absent in yeast. This demonstrates that SR-proteins are functionally highly related to complex alternative splicing (Plass et al. 2008; Busch and Hertel 2012).

RRM domains of SR-rich proteins can directly bind to the primary transcript in a sequence specific manner but have also been reported to mediate interactions to proteins

(Michlewski et al. 2008; Pelisch et al. 2010). The SR domain is a mediator of protein-protein interactions and its serine residues undergo extensive phosphorylation which regulates protein interactions as well as subcellular and subnuclear localization (Xiao and Manley 1998; Risso et al. 2012; Li et al. 2013). Based on protein-protein interactions with U1 snRNP and U2AF, SR proteins can initiate spliceosome assembly, thus strengthening potentially weak splice sites that are not well defined (Ruskin et al. 1988; Fu 1993; Staknis and Reed 1994; Blencowe et al. 1999). Phosphorylation and thus activation of SR proteins enabling protein interactions is facilitated by SR protein kinases (SRPKs) and by Clk/Sty kinases (Ngo et al. 2005). Given the versatile functions and possible binding positions of SR proteins, phosphorylation can promote as well as inhibit exon inclusion. The whole process of exon definition and its balance depends on the interplay of SR(r) protein abundance and their activity regulated by kinases and phosphatases, their binding sites on the primary transcript and the presence or recruitment of cooperating or antagonizing factors (Blencowe et al. 1999; Cowper et al. 2001; Erkelenz et al. 2013). All together orchestrate fine tuning of alternative splicing which makes the prediction of alternative splicing events a big challenge.

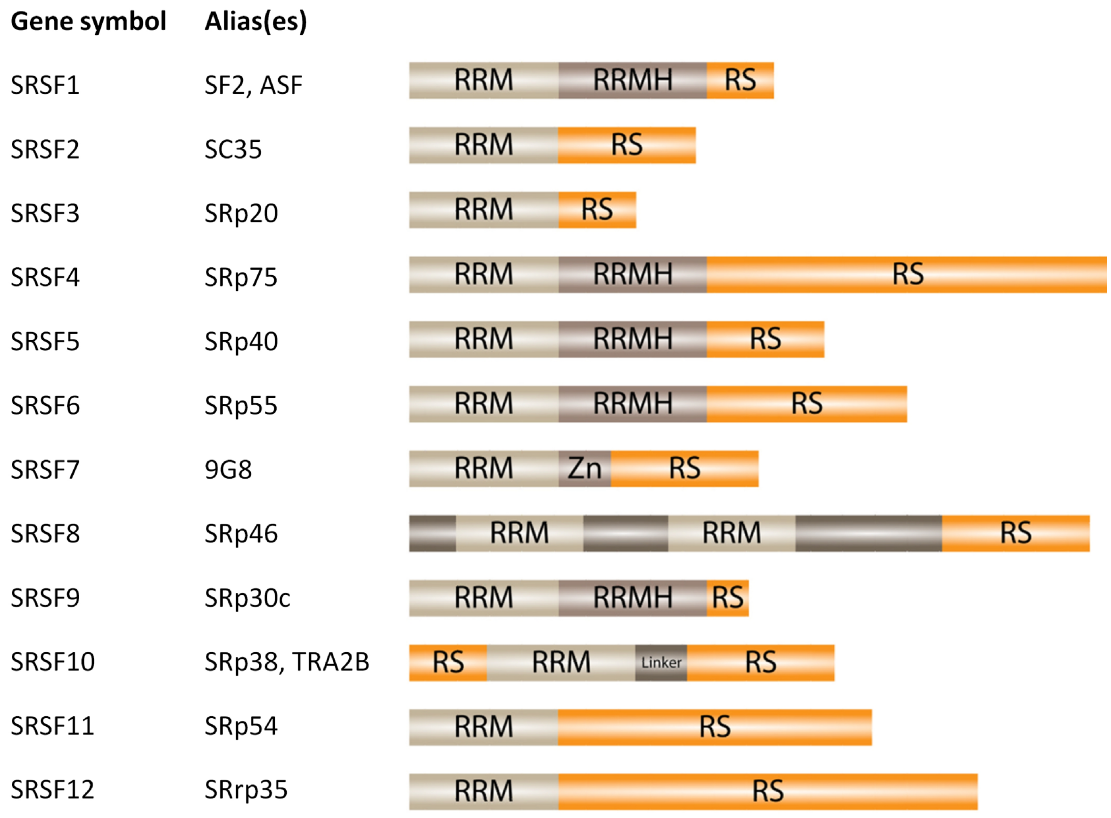

Figure 3 | Domain structure of human canonical SR-proteins.
Typically SR-family members are defined by N-terminal RRMs followed by a downstream RS domain. TRA2B holds a special role as its domain composition with a single RRM domain flanked by serine-arginine-rich domains differs from all other SR-proteins. RRM(H), RNA recognition motif (homolog); RS, arginine-serine-rich; Zn, zinc-finger. (Modified from (Manley and Krainer 2010; Busch and Hertel 2012))

Next to their important role in alternative splicing, SR-rich proteins have been implicated in miRNA-mediated translational repression, miRNA processing and post-translational modification. Moreover, there is evidence of general translational regulation by SR-rich proteins via the regulation of mRNA stability and mRNA export from the nucleus (reviewed in (Sanford et al. 2008; Risso et al. 2012)).

1.4. The SR-related splice factor TRA2B (SFRS10)

The splicing factor TRA2B (SFRS10*,* **s**plicing **f**actor arginine(**R**)/serine(**S**)-rich **10**) belongs to the family of arginine-/serine-rich splicing factors and contributes to alternative splicing of specific exons. Its ortholog *transformer-2* was first described in *Drosophila melanogaster* where it regulates sex differentiation, that is controlled by alternative splicing (Nagoshi et al. 1988; Baker 1989). While sex in the fly is initially defined by the ratio of X-chromosomes and autosomes, specific splicing of the *doublesex* (*dsx*) gene is a major determinant for maintaining female sex commitment in the somatic cell lineage. *Transformer-2* is strongly expressed in females and promotes in cooperation with *transformer* the expression of the female-specific *doublesex* isoform by including exon 4 into the *dsx* transcript. In male flies expression levels of *transformer-2* are low and exon 4 is skipped from the *dsx* transcript, giving rise to the male *doublesex* isoform. Chromosomal females that carry loss of function mutations in *transformer-2*, transform physiologically and behaviorally into viable males which are sterile, however (Belote and Baker 1982; Butler et al. 1986). Mice do not tolerate complete ubiquitous ablation of their equivalent ortholog *Tra2b*. Embryos deficient for *Tra2b* die at early gestational stages around 7.5 dpc (Mende et al. 2010). This implicates the splicing factor Tra2b in having a pivotal role during development in mammals.

The human and murine genes of *TRA2B* (*Tra2b*) are located on chromosome 3q27.2 in humans and on chromosome 16qB1 in mice. *TRA2B* encodes a 288 amino acid (33.7 kDa) protein that is 100 % identical to the murine version and still preserves 58 % sequence similarity (45 % identity) compared to *Drosophila melanogaster (Segade et al. 1996)*. The sequence identity between humans and mice suggests a highly conserved function of TRA2B in mammals probably with a high number of common target transcripts. In the cell TRA2B localized to nuclear speckles, which is the typical subnuclear localization expected for splice factors of the SR-family (Nayler et al. 1998; Chen et al. 2003).

The domain structure of TRA2B is different from other human SR-proteins (Figure 3) as TRA2B exhibits a unique arrangement of its 2 SR-domains. All other SR-proteins described in humans possess a single SR-domain that is always located C-terminally of the RRM. The

canonical isoform *TRA2B-1* possesses a single central RRM flanked by two SR-domains at the N- and C-terminal end (Busch and Hertel 2012) (Figure 4). Splicing of human *TRA2B* gives rise to 5 different mRNAs of which 2 produce functionally different protein products. If included, exon 2 introduces several inframe premature termination codons. Therefore, isoforms including exon 2 do not yield a protein product. The canonical *TRA2B-1* and the *TRA2B-3* (also termed *TRA2B-ΔRS1*) isoform share the typical initiation codon in exon 1. These isoforms skip either exon 2 or both exons 2 and 3, respectively. The latter leads to production of the shorter *TRA2B-ΔRS1* isoform that lacks the N-terminal SR-rich domain (Nayler et al. 1998) (Figure 4).

Even though *TRA2B* is ubiquitously expressed, strongest expression can be found especially in neuronal tissue and testis but also in the heart and intestine (Chen et al. 2003; Grellscheid et al. 2011a). Generation of the *ΔRS1* isoform is conserved in flies and mammals and its expression has been reported to be spatiotemporally highly regulated by Clk-kinases (1.3). Still, no distinct function in splicing has been related to the *ΔRS1* isoform (Mattox et al. 1990; Nayler et al. 1998; Daoud et al. 1999; Stoilov et al. 2004). Overexpression of *TRA2B-ΔRS1* has failed to promote inclusion of TRA2B-dependent exons of the *Krba1*, *Pank2*, *Tra2a* and *HIPK3* transcripts, suggesting that the *RS1* domain is likely crucial for protein-protein interactions to allow exon inclusion (Grellscheid et al. 2011a; Grellscheid et al. 2011b).

Depending on the targeted transcript both SR-domain and RRM are necessary to facilitate exon inclusion. However, TRA2B-ΔRRM constructs (removal of the RRM) have been demonstrated to exert a negative effect on splicing of target exons, which are effectively spliced by TRA2B-1. This indicates that cooperative and independent function of SR-domains and RNA recognition motifs in TRA2B is likely (Dauwalder and Mattox 1998; Grellscheid et al. 2011a). This mechanism bears the potential of tissue-specific alternative exon selection depending on *TRA2B* isoform expression in different tissues.

TRA2B facilitates splicing of its own transcript to regulate its expression levels in a negative feedback-loop. Upon inclusion of the cassette exon 2, premature termination codons are inserted rendering the transcript non-functional (Stoilov et al. 2004). TRA2B binds to multiple ESEs within exon 2 to activate its inclusion. In the mouse, exon 2 is the only *Tra2b* exon that is alternatively regulated. All other detectable isoforms in the mouse are only distinguishable by differently long 3'UTRs, based on alternate usage of polyadenylation signals. These give rise to 4 functional full-length transcripts with 2.1, 2.0, 1.5 and 1.4 kb in size (Segade et al. 1996). Hyperphosphorylation of TRA2B has firstly been shown to alter subnuclear and subcellular localization (Li et al. 2013), and secondly, to increase protein-protein interactions via the SR-domains, which could sequester HTRA2B from RNA (Stoilov et al. 2004). Both would promote skipping of *TRA2B* exon 2 and might equally impact other target

transcripts as well. Other examples of how differential phosphorylation of TRA2B can promote or inhibit exon inclusion are the exon splicing processes of *Mapt* (Microtubule associated protein tau) exon 10 and *SMN2* (Survival motor neuron 2) exon 7 (Novoyatleva et al. 2008). Modulation of TRA2B phosphorylation to enhance *SMN2* exon 7 inclusion was even considered as a therapeutic approach for spinal muscular atrophy (Zhang et al. 2011) (also see chapter 1.6).

TRA2B binds and recognizes exons via short GAA-rich sequence motifs (Tacke et al. 1998; Stoilov et al. 2004; Grellscheid et al. 2011a). Unraveling the complex structure of RRM-RNA interactions revealed embedding of a 6 nucleotide stretch into the RRM. Inside this, only the central AGAA motif is specifically recognized by the TRA2B RRM (Clery et al. 2011). Moreover, a direct but weaker interaction has been demonstrated between the RRM and a CAA sequence motif, for which a stem-loop secondary structure was a prerequisite, however (Tsuda et al. 2011).

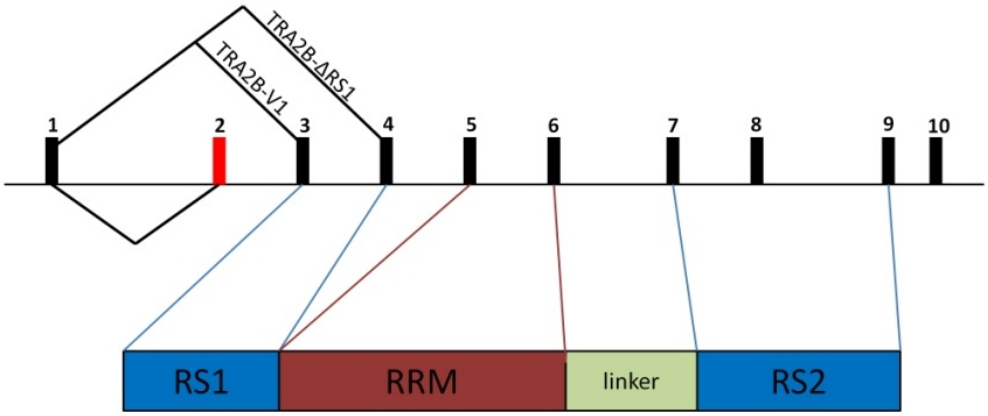

Figure 4 | Domain structure and splicing of human *TRA2B*.
The *TRA2B* gene consists of 10 exons and undergoes alternative splicing forming a total of 5 different mRNAs. Isoforms including exon 2 do not produce a protein product due to insertion of premature termination codons. Two functionally different isoforms are produced: the canonical *TRA2B-V1* (exon 2 skipped) and *TRA2B-ΔRS1* (exons 2 and 3 skipped). The TRA2B protein comprises a central RNA recognition motif (RRM) flanked by 2 serine/arginine-rich (RS) domains. The domains are encoded by the exons they are linked to. Thus, alternative splicing of exon 3 can either include or exclude the N-terminal RS1-domain. (Figure inspired by (Stoilov et al. 2004; Mende 2008))

1.5. TRA2B-mediated alternative splicing and disease

Testis and spermatogenesis

TRA2B is classified to be a non-essential splicing factor as it alone cannot restore splicing activity to S100 cytoplasmic extracts (Tarn and Steitz 1994; Tacke et al. 1998; Zahler 1999). This and the fact that *TRA2B* deficiency itself cannot be compensated and is not tolerated in mammals (Mende et al. 2010) might point towards unique functions and unique splicing targets of TRA2B. Thus, target exons might dedicatedly rely on TRA2B.

Specific TRA2B-related splicing processes have been implicated in human disease and in distinct physiological functions in specific tissues. *TRA2B* is strongly expressed in testis (Grellscheid et al. 2011a) and its impact on exon inclusion has been shown to be inverted by direct interaction with the splicing-regulator RBMY on testis-specific transcripts (Venables et al. 2000; Dreumont et al. 2010). TRA2B has been shown to evoke inhibitory effects on specific exons of the testis-enriched transcript *ACIN1* and *CREB*, which was turned into a promoting effect upon *RBMY* overexpression. Gross alterations of TRA2B function occur upon interaction with modulating factors in a tissue-specific manner. Moreover, the testis-specific T-exon of the homeodomain interacting protein kinase 3 (*HIPK3*-T) has been shown to be tightly regulated by TRA2B in a highly concentration-dependent manner (Venables et al. 2005). For effective splicing of *HIPK3-T*, *TRA2B* needed to be expressed at elevated levels and required multiple clustered exonic AGAA binding motifs. Thereby, the effect of switching exon inclusion on or off in a TRA2B-concentration-dependent fashion has been shown to rely on the intrinsic splice site strength (Grellscheid et al. 2011b). TRA2B can exert greatest impact and concentration-dependent modulation on exon splicing on a weak splice site background (Hertel 2008; Grellscheid et al. 2011b). Its elevated expression and direct involvement in the splicing of specific transcripts implicate TRA2B having important physiological functions in testis and probably spermatogenesis.

Tauopathies

TRA2B involvement has also been demonstrated in numerous neuronal-specific processes. The microtubule associated protein tau (encoded by the *MAPT* gene) binds to microtubules and helps to regulate their stability and assembly. *Mapt* is spliced with a total of 6 different isoforms. Splicing of *Mapt* exon 10 is determining the number of microtubule-bidning domains in the protein. Transcripts containing exon 10 will possess 4 binding domains (isoform 4R), while lack of exon 10 will produce the 3R isoform (3 binding domains) (Goedert and Jakes 1990). The number of microtubule-binding domains determines the phosphorylation

of Tau and thus the microtubule binding properties. Disturbance of this regulated system is associated with tauopathies like Alzheimer's disease (AD) and Frontotemporal Dementia and Parkinson linked to chromosome 17 (FTDP-17) (Wang and Liu 2008; Zhou et al. 2008; Denk and Wade-Martins 2009). Indeed, mutations associated with FTDP-17 aggregate in *MAPT* exon 10 and can either disturb protein function by amino acid exchange or disrupt splicing regulatory elements (Liu and Gong 2008; Wang and Liu 2008). Splicing regulation of *MAPT* exon 10 is directly regulated by TRA2B. Inclusion of the fourth microtubule binding domain is strictly regulated in neurons and both 3R and 4R isoforms coexist in a defined ratio. Splicing defects of exon 10 have equally been demonstrated to cause isoform imbalances causing neurodegeneration present in tauopathies *(Jiang et al. 2003; Kondo et al. 2004; Wang et al. 2005)*. Supporting this idea, coordinate upregulation of *TRA2B1* and *MAPT-4R* transcripts have been found in postmortem temporal cortices of AD patients (Glatz 2005; Glatz et al. 2006).

Development and neuronal function

Further transcripts that are known targets of TRA2B include fibroblast growth factor receptor 2 (*Fgfr2*) and clathrin light chain b (*Cltb*). *Fgfr2* is implicated in neuronal and non-neuronal development (Ford-Perriss et al. 2001) while *Cltb* along with other clathrins encodes a vesicle coating protein and is involved in the regulation of endocytosis (reviewed in (Brodsky 2012)). Splicing of *Cltb* exon 5 (also exon EN) has been shown to be indirectly and negatively regulated by TRA2B *in vitro* (Stamm et al. 1999). The neuronal *Cltb* isoform II (exon 5 included) comprises regulatory features for the C-terminal calmodulin binding domain. This isoform has been demonstrated to be sensitive to calcium and it has been proposed that skipping of *Cltb* exon 5 (isoform III) might keep vesicles in a calcium-insensitive state in non-neuronal tissue (Stamm et al. 1992; Pley et al. 1995). Here TRA2B might have a tissue-specific function in mediating vesicular control in response to calcium.

Fibroblast growth factor receptors (*FGFRs*) mediate multiple biological activities including development, tissue homeostasis and metabolic processes. FGFRs are transmembrane proteins that convey signals to the cell depending on the binding of specific extracellular ligands (i.e. FGFs, fibroblast growth factors). Splicing of *FGFR* isoforms is a common mechanism that makes FGFRs very flexible by allowing modulation of their kinase- or ligand binding-domains (Kalinina et al. 2012; Belov and Mohammadi 2013). TRA2B regulates the splicing of *FGFR2* exons 8 and 9 in a mutually exclusive fashion (Figure 1 (6)) (Chen et al. 2004; Eswarakumar et al. 2005). These two exons both encode the C-terminal half of the D3 extracellular domain. Alternative - but mutually exclusive – usage of these exons modifies the

extracellular domain, thus changing ligand binding specificity (Chen et al. 2004). This might enable distinct functionality of FGFR2 in different tissues depending on exon selection.

Smooth musculature can be divided into slow, intermediate and fast phenotype, based on their rates of contraction. These phenotypes are specified during development depending on neural input and mechanical load (Somlyo and Somlyo 1994; Owens 1995). Alternative splicing is a known phenomenon to regulate numerous muscle-related transcripts to allow diversification of smooth musculature into fast and slow contractile types (Sobue et al. 1999). TRA2B has been reported to positively control inclusion of exon 23 of the Myosin phosphatase targeting subunit I (*MYPT1*). Expression levels of *TRA2B* are strongly elevated in fast smooth muscle (as compared to slow smooth muscle) and highly correlate with exon 23 inclusion into the *MYPT1* transcript. This implicates TRA2B in the specification of the fast smooth muscle phenotype (Shukla and Fisher 2008).

1.6. Spinal muscular atrophy: The role of TRA2B in SMN exon 7 splicing

Spinal muscular atrophy (SMA) is an autosomal recessive disease and the leading genetic cause of infant death (Crawford and Pardo 1996). SMA is characterized by degeneration and loss of α-motor neurons in the anterior horns of the spinal cord leading to increasing weakness and atrophy of the voluntary muscles (Zerres and Rudnik-Schoneborn 1995; Markowitz et al. 2004). For the very most, SMA is caused by homozygous disruption of the survival motor neuron 1 gene (*SMN1*) (Lefebvre et al. 1995; Wirth 2000). In humans there are two copies of the *SMN* gene termed *SMN1* and *SMN2*. Both encode identical proteins and genomically only differ by 5 nucleotides that are either intronic, synonymous or located in untranslated regions (Burglen et al. 1996). Even though encoding the same protein, the *SMN2* transcript is spliced differently from the *SMN1* transcript which is caused by a C to T transition at position 6 of *SMN2* exon 7 (Figure 5). This results in a dramatically lower inclusion rate of exon 7 into the *SMN2* transcript (Lorson et al. 1999; Monani et al. 1999). Only 10-20 % of *SMN2* transcripts have exon 7 included resulting in the production of functional full-length SMN protein. The protein produced from *SMN2* Δ7 transcripts is truncated and has shown impaired oligomerization properties as well as reduced protein stability (Lorson et al. 1998; Burnett et al. 2009).

The intrinsic splice site strengths of *SMN* exon 7 are rather weak causing generally poor exon definition (Lim and Hertel 2001; Singh et al. 2004). Caused by this, exon 7 splicing strongly depends on the function of accessory trans-factors working in concert with spliceosomal components. The C to T transition in *SMN2* exon 7 causes disruption of an exonic

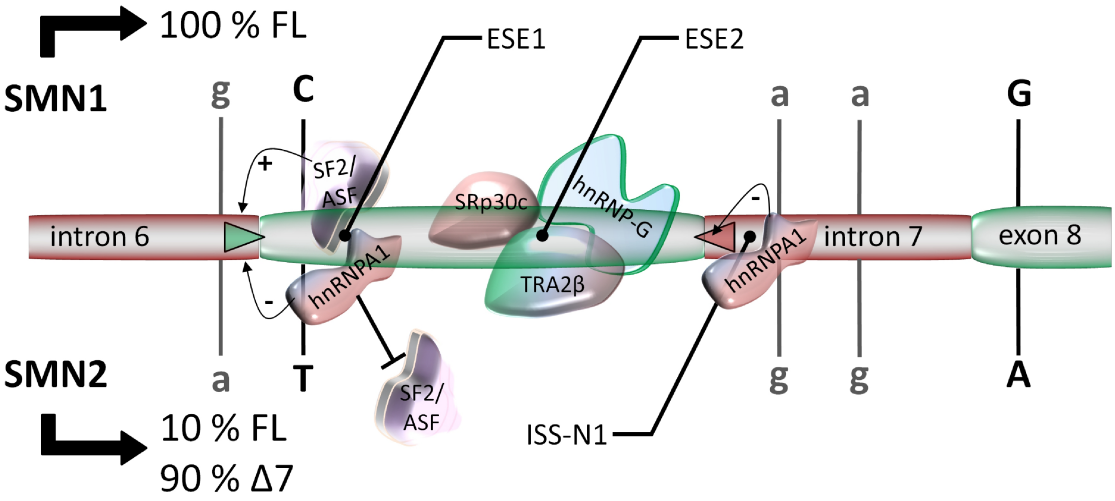

Figure 5 | Model of *SMN* exon 7 splicing.
SMN1 and *SMN2* differ in their splicing pattern due to a C to T transition at position 6 of exon 7. SF2/ASF binds to ESE1 and interacts with U2AF/U2 at the 3'ss (green arrow) to promote splicing of intron 6. In *SMN2* ESE1 is disrupted by the C to T transition and SF2 can no longer antagonize hnRNP-A1, that exerts a negative effect on the 3'ss promoting exon 7 skipping. Thus, *SMN2* produces transcripts that in 90 % of the cases lack exon 7 (Δ7). TRA2B specifically binds GA-rich sequence motifs at the ESE2 and acts in concert with SRp30c and hnRNP-G to promote exon 7 inclusion. An intronic splicing silencer (ISS-N1) at the beginning of intron 7 exerts a negative effect to the 5'ss via hnRNP-A1.

splicing enhancer (Figure 5, ESE1) and generates an exonic splicing silencer (Lefebvre et al. 1995; Cartegni and Krainer 2002; Kashima and Manley 2003). Binding of the SR-splicing factor SF2/ASF to ESE1 enhances exon recognition at the 3'ss via interaction with U2AF/U2, but its binding depends on the cytosine nucleotide at position 6. SF2 cannot bind to the *SMN2* pre-mRNA due to the C to T transition and does no longer antagonize binding of the inhibitory factor hnRNP-A1. This exerts a strong negative effect on the 3'ss which promotes exon 7 skipping from *SMN2* transcripts (Kashima and Manley 2003; Cartegni et al. 2006). By using an *SMN2* minigene it has been demonstrated that TRA2B can revert the *SMN2* splicing pattern by strongly promoting exon 7 inclusion in a concentration-dependent fashion, thereby restoring the *SMN2* splicing pattern up to 80 % to produce functional full-length transcripts (Hofmann et al. 2000). TRA2B specifically and directly binds a GA-rich sequence motif at the ESE2 in the central region of exon 7 and acts in concert with SRp30c and hnRNP-G (Hofmann and Wirth 2002). hnRNP-G has been shown to contribute in great extent to exon 7 inclusion but does not directly associate with *SMN2* exon 7 in a sequence-specific manner (Hofmann and Wirth 2002). Instead hnRNP-G has been demonstrated to interact with TRA2B, which suggested hnRNP-G being recruited to *SMN2* exon 7 via TRA2B. This makes TRA2B a crucial determining factor for *SMN* exon 7 inclusion that can act in parallel to SF2/ASF. Indeed, overexpression of TRA2B has been demonstrated to increase SMN protein levels produced from endogenous transcripts, which makes TRA2B a potent candidate that could be exploited to correct the splicing pattern of *SMN2* transcripts (Hofmann and Wirth 2002).

1.6.1. Splicing correction as a therapeutic approach for SMA

The number of *SMN2* gene copies in humans is variable between individuals. As full absence of both *SMN1* and *SMN2* is incompatible with life, all SMA patients carry at least one or two *SMN2* gene copies. The more *SMN2* copies are present, the more functional SMN protein is produced. For that reason the *SMN2* copy number correlates inversely with the SMA disease severity, which makes *SMN2* the major disease determining modifier in SMA (Wirth 2000; Wirth et al. 2006). Splicing correction of *SMN2* that is present in every patient could allow compensating the loss of *SMN1*. Indeed, numerous approaches aiming to establish a possible therapy for SMA have been focusing on increasing the amount of functional SMN protein produced from the *SMN2* gene by enhancement of *SMN2* expression and splicing correction.

Histone deacetylase inhibitors (HDACi) are substances inhibiting HDAC function, thus attenuating chromatin condensation and generally enhancing gene expression. Numerous substances have been demonstrated to enhance *SMN2* expression, among them valproic acid (Brichta et al. 2003), M344 (Riessland et al. 2006), SAHA (Riessland et al. 2010) and LBH589 (Garbes et al. 2009). A general increase of total *SMN2* expression also increases the amount of functional full-length transcripts accordingly. Interestingly, valproic acid and M344 did further cause correction of the *SMN2* splicing pattern *in vitro* resulting in increased amounts of full-length *SMN2* transcripts. This has been demonstrated to be attributable to a more than 3-fold reinforced expression of TRA2B (Brichta et al. 2003; Riessland et al. 2006). Due to their dual mode of action these results underlined the value of HDACi for possible treatment strategies for SMA. The combinatorial effect of enhanced transcription from a probably low copy number of *SMN2* genes and the correction of the splicing pattern to yield more full-length *SMN2* transcripts constitutes a promising approach. As to its potency in splicing correction of exon 7, TRA2B holds a central role in such treatment strategy. However, it is not predictable which other transcripts or exons might be affected by increased TRA2B levels. Knowing that TRA2B exerts splicing effects depending on its protein concentration (Hofmann et al. 2000; Hofmann and Wirth 2002) it appears likely that the regulation of TRA2B target exons will change upon its upregulation. Few TRA2B splicing targets have been described and alteration of their splicing patterns has been associated with disturbed protein function and probably disease (1.5). HDAC inhibitors (and especially VPA) are not solely used as a therapeutic approach for SMA, but have been considered for treatment of other conditions like epilepsy (Davis et al. 1994), bipolar disorder (Cipriani et al. 2013) and for the prophylaxis of episodic migraine (Linde et al. 2013). Furthermore, VPA has been found to exert positive impact in fighting cancer (Cha et al. 2013;

Han et al. 2013). This has mainly been attributed to its ability to transcriptionally upregulate tumor suppressor genes (Qiu et al. 2013).

In frame of HDACi-related treatment strategies for various conditions, the identification of TRA2B splicing targets and the characterization of their physiological role as well as the effect evoked by altered splicing are of major interest and are the central objective of this study. Previous studies have generally used *in vitro* approaches involving minigene constructs. Via this method the impact of a splicing factor on an exon can be safely determined, but it does not fully reflect splicing under physiological conditions, as it uses artificial gene constructs and non-physiological splicing factor concentrations (i.e. overexpression). Therefore, this study is aiming at the identification of novel TRA2B splicing targets *in vivo* using conditional knock-out mice. The identity of human TRA2B and murine Tra2b on protein level implies that splicing targets identified in the mouse are likely also TRA2B splicing targets in humans. The use of conditional Tra2b knock-out mice will be further explained in chapter 1.7

Besides the use of HDACi other methods have been researched to correct the splicing pattern of *SMN2*. Different strategies using antisense oligonucleotides (ASO) to modulate exon 7 splicing inclusion have been developed. ASO bind sequence-specifically to their target sequence and block the interaction of trans-factors with mRNA. Successful restoration of the *SMN2* splicing pattern has been shown for an ASO targeting the intronic splicing silencer ISS-N1 (Figure 5) at the beginning of intron 7 thus blocking the binding of the inhibitory factor hnRNP-A1. This could restore *SMN2* splicing in a humanized SMA mouse model (Hua et al. 2008). Other ASO designs have been aiming to modulate competition of exon joining to promote exon7 inclusion. An ASO targeting the 3'ss of *SMN2* exon 8 has been shown to incapacitate exon 8 to compete for joining with the exon 6 5'ss, thus promoting exon 7 joining (Lim and Hertel 2001). As to the limitations of ASO to only block binding sites, bifunctional RNAs have been developed to not only block inhibiting factors but also actively promote exon 7 inclusion. Like ASO, bifunctional RNAs comprise an antisense sequence for specific sequence targeting but further contain enhancing sequence elements (ESE) that constitute a basis for the recruitment of enhancing factors like HTRA2B or SF2 (Skordis et al. 2003; Baughan et al. 2006; Baughan et al. 2009). In general, ASO-based approaches appear highly specific as modulation of splicing is restricted to transcripts that are specifically targeted by the complementary ASO sequence. As a treatment strategy these methods harbour the potential of having an unmatched low number of off-targets as opposed other drugs that affect splicing of many transcripts.

As an ASO-independent approach the use of modified cantharidins has been proposed to modulate TRA2B phosphorylation (Zhang et al. 2011). Cantharidins are protein phosphatase (PP) inhibitors and inhibition of PP1 has been demonstrated to reduce TRA2B phosphorylation at Thr-33. Thus, the use of modified cantharidins has been demonstrated to enhance exon 7 inclusion in *SMN2* minigenes and SMA patient fibroblasts by increased TRA2B activity (Zhang et al. 2011). However, cantharidins and their derivatives are not very specific and both protein phosphatase 1 and 2 are essential for the maintenance of general splicing processes.

1.7. Mouse genetics

During the past decades the mouse has become the most frequently used laboratory animal model organism. Its high genetic similarity to humans and broad opportunities of genetic manipulation of the mouse has allowed the study of versatile biological processes and diseases. The mouse, with 99 % of its genes having direct counterparts in humans, hold the key for the study of human biology and disease (Gunter and Dhand 2002).

As introduced previously (1.6.1), this study is aiming at the identification of novel TRA2B splicing targets *in vivo* using a conditional *Tra2b* knock-out mouse model. The generation of ubiquitous *Tra2b* knock-out mice has been previously performed and mice fully depleted of Tra2b were found to be early embryonically lethal (Mende 2008; Mende et al. 2010). In anticipation of that the applied targeting strategy for the generation of these mice included the introduction of loxP sites at the *Tra2b* locus. This allows for conditional knock-out that is mediated by tissue-specific expression of Cre-recombinase (1.7.1). In an attempt to analyze Tra2b-related splicing processes especially in the context of SMA (1.6), motoneuron-specific knock-out mice have been generated using Cre-recombinase expressed under control of the Hb9 promoter. These mice did not develop an SMA-like phenotype. Changes in FL-*Smn* transcripts have not beed detected upon *Tra2b* depletion, even though the Δ7 isoform has been found to be upregulated by up to 4-fold (Mende et al. 2010). In this study a conditional knock-out mouse depleted of *Tra2b* in the central nervous system should be generated. By expression of Cre-recombinase under control of the rat Nestin promoter and enhancer (Tronche et al. 1999) the *Tra2b* locus gets disabled in neuronal and glial precursor cells of the developing central nervous system (1.7.1.). This allows to study Tra2b-related splicing processes in a broader cell variety of neuronal origin.

More detailed information about Cre-mediated recombination and the generation of neuronal-specific knock-out mice is given in the chapters 1.7.1 and 4.2.2. As severe developmental brain defects were encountered during the analysis of neuronal-specific *Tra2b*

knock-out mice, chapter 1.7.2 gives a short overview about the neuronal and cortical development in the mouse.

In 2011 the group of Elizabeth Patti has independently generated a *Tra2b* knock-out mouse using the gene trap technology (Pihlajamaki et al. 2011). In concordance with the results obtained during mouse breedings by our workgroup (Mende et al. 2010), homozygous *Tra2b* knock-out pups could not be detected indicating early embryonic lethality. Expression of TRA2B has been demonstrated to be reduced in human obesity and has been suggested to be involved via splicing of the *LPIN1* gene (Pihlajamaki et al. 2011).

1.7.1. The Cre/loxP system of site-directed recombination

The mouse is the most frequently used mammalian model organism. Genetic manipulation is relatively easy and a great number of genes and their functions are being studied in knock-out mouse models. In a knock-out mouse the gene of interest is disabled to reversely study its function. Knock-out is usually achieved by gene targeting (homologous recombination) of gene loci and their disruption (Capecchi 1989; Mansour et al. 1990). Numerous studies yield limited success, as mice might not tolerate absence of some genes during development resulting in early lethality (Lewandoski 2001).

Conditional knock-out strategies emerged to overcome that problem allowing gene depletion that is limited to a specific tissue or are set to occur at a defined time point. A commonly used system to achieve tissue specific knock-out is the Cre-loxP system, which allows targeted recombination upon expression of *Cre-recombinase*. *Cre-recombinase* originates from *bacteriophage P1* and belongs to the integrase family of specific recombinases (reviewed by (Nagy 2000)). The 38 kDa integrase catalyzes the recombination between two of its specific recombination sites, so called loxP sites (Hamilton and Abremski 1984; Sauer and Henderson 1989). LoxP sites are 34 bp consensus sequences that contain an 8 bp core spacer sequence flanked by 2 palindromic 13 bp sequences. The orientation of the loxP site is determined by the asymmetric core sequence. Cre-recombinase does not require additional co-factors to operate. That makes it well-suited for the use in a foreign cellular environment as it can perform recombination independently (Hamilton and Abremski 1984; Lewandoski 2001).

During the reaction of recombination one Cre-recombinase molecule binds to each palindromic half of a loxP site. Cre-recombinase forms a tetramer combining two loxP sites located at different genomic locations. Recombination occurs within the central spacer region of the loxP site causing the post-recombination site to be assembled by each two halves of the complementary pre-recombination sites (Voziyanov et al. 1999). As a result of recombination the sequence in between of two loxP sites is excised. The first use of Cre-recombiase-mediated

recombination to generate conditional knock-out mice was reported in 1994 by the group of Prof. Klaus Rajewski (Gu et al. 1994).

In the present study the effect of *Tra2b* depletion in the central nervous system was analyzed. Neuronal-specific knock-out mice were generated by cross-breeding *Tra2b*$^{fl/fl}$ animals with a *Cre-recombinase* transgenic mouse line expressing *Cre-recombinase* under control of the rat *Nestin* promoter and enhancer, which is active in neuronal and glial precursor cells (Tronche et al. 1999). It is first activated at 10.5 dpc in the neural tube, the somites and in migrating cells of the neural crest. At 15.5 dpc it has been reported to be strongly expressed in the ventricular and subventricular zone of the cortex (Dahlstrand et al. 1995). Generation of the *Tra2b* mouse line by targeting the murine *Tra2b* locus via homologous recombination facilitating insertion of loxP sites was performed as described previously (Mende 2008; Mende et al. 2010).

1.7.2. Cortical development in the mouse

Neuronal development in the mouse starts with the induction of ectodermal cells to form the neuroectoderm that develops to the neural plate. These steps occur during gastrulation between 6.5 – 9.5 dpc. During neurulation the neural plate undergoes multiple folding steps to finally form the neural tube. The neural tube will develop to form the central nervous system, while its lumen will form the ventricular system. First brain structures start to develop at around 9 dpc at the anterior end of the neural tube and can be subdivided into prosencephalon (forebrain), mesencephalon (midbrain) and rhombencephalon (hindbrain). These basic structures will later form distinct brain areas (Figure 6). The prosencephalon subdivides into diencephalon and telencephalon. The diencephalon will form all thalamic structures, while the most anterior part of the neural tube, the telencephalon, will give rise to the cerebral cortex. The ventricular system consisting of 4 cavities that are interconnected by the aqueduct allows the transport of cerebrospinal fluid. Both hemispheres of the telencephalon comprise one ventricle termed lateral ventricle. A third and fourth ventricle is located within the diencephalon and rhombencephalon, respectively (Kaufman and Bard 1999).

The mammalian cortex is divided into functionally distinct areas, which strongly differ between species and are considered to be determining for mental and perceptive performance. The formation of these areas is referred to as regionalization or arealization. To date, it is not fully understood how aeralization is regulated. However, traits like individual gene expression patterns that are autonomous to a specific region appear equally important as extrinsic input from thalamocortical afferents (O'Leary et al. 1994; O'Leary and Nakagawa

2002). The latter are neuronal connections that link a thalamic nucleus to a specific cortical region. Thalamic input shows a high level of specificity with regard to the cortical area and cortical layer that are targeted by the afferents. Most thalamic input terminates in layer IV of the neocortex. Layers V and VI send reciprocal connections back to the thalamic nuclei (Lopez-Bendito and Molnar 2003).

The six-layered neocortex is the largest and the most specialized region of the cerebral cortex (Krubitzer and Huffman 2000). Most cortical neurons emerge from the ventricular and subventricular zones (VZ/SVZ) of the lateral ventricles. The first fully differentiated, post-mitotic neurons form the pre-plate. Subsequently generated neurons coming from the VZ/SVZ migrate to the outside along radial glia to later form the different layers of the cortex (Krubitzer and Huffman 2000; Lopez-Bendito and Molnar 2003). Thalamic and cortical development occur synchronously and the time frames of neurogenesis in cortex and thalamus strongly overlap. In mice the most significant changes in cortical patterning occur during the second and third week of gestation. Between 13 and 18 dpc the neocortex and nuclei of the dorsal thalamus build up reciprocal connections. In mammals thalamic afferents reach their cortical destination area before their actual target neurons are generated. It has been proposed that thalamic projections are homing in on their target area sensing specific signals. Continuous interactions and signaling between the thalamus and the cortex is necessary to organize cortical patterning, for arealization and for the development of an intact cortical circuitry (Lopez-Bendito and Molnar 2003).

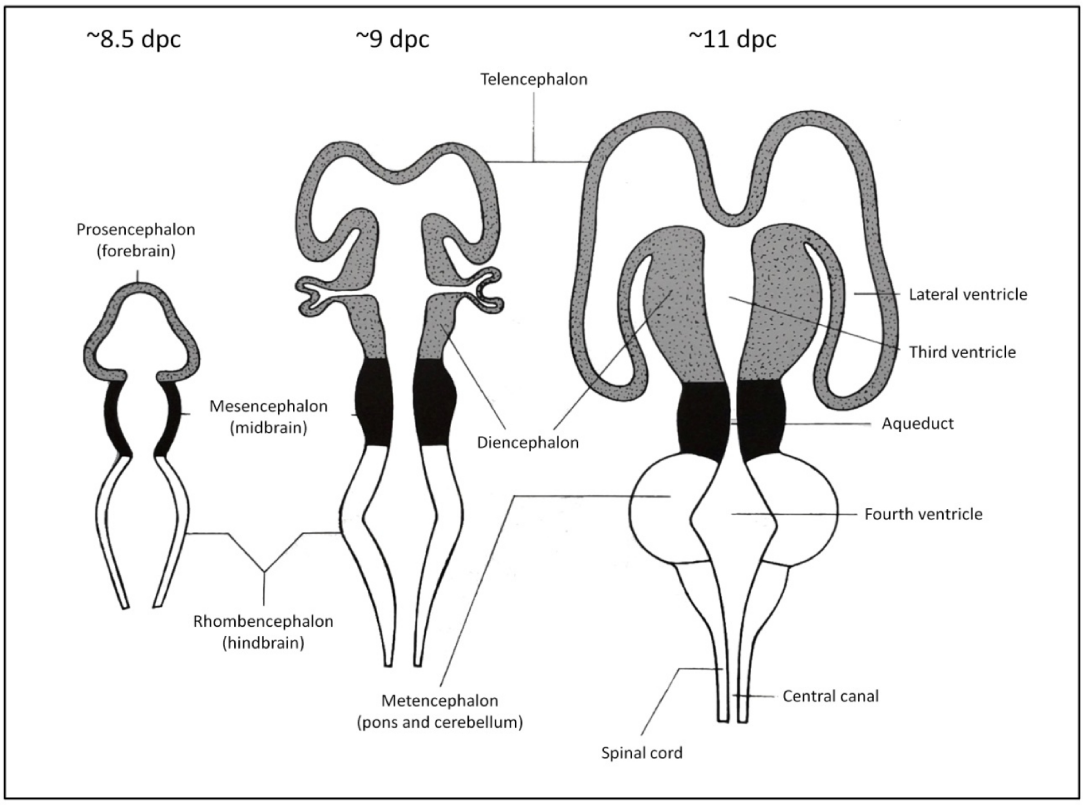

Figure 6 | The early murine brain development.
Subdivision of specializing areas of the neural tube can first be observed at around 8.5 dpc. The neural tube splits into Prosencephalon, Mesencephalon and Rhombencephalon. Telencephalon and Diencephalon emerge from the Prosencephalon. Formation of the ventricular system is observable at around 11 dpc, as the lateral ventricles form at the Telencephalon, the third ventricle at the Diencephalon and the aqueduct at the Mesencephalon. The cerebral cortex will emerge from the Telecephalon at subsequent developmental stages. (Modified from (Kaufman and Bard 1999))

2. Aims

SMA is the leading genetic cause of infant lethality and is caused by functional absence of the *SMN1* gene. Every patient carries at least one copy of the almost identical *SMN2* gene. Caused by alternative splicing *SMN2* produces only 10 % of functional transcripts, which depending on the *SMN2* copy number can ameliorate the SMA phenotype but is insufficient to prevent disease onset. Based on this, modulation of the *SMN2* expression and correction of *SMN2* splicing have been major subjects of research in developing a therapy for SMA.

The use of HDAC inhibitors has proven itself to be a promising approach as a treatment strategy for SMA. HDACi treatment has been demonstrated to not only enhance general expression of *SMN2* but also to correct the *SMN2* splicing pattern to primarily produce functional full-length transcripts. This has been shown to occur via massive upregulation of *TRA2B* expression, which in turn results in effective inclusion of exon 7 into *SMN2* transcripts.

Despite the beneficial effect of TRA2B upregulation on *SMN2* splicing, it can be assumed that the splicing of other transcripts targeted by TRA2B will be affected by its upregulation. Indeed, few TRA2B targets have been described and their missplicing has been associated with aberrant function and disease. For that reason it is of utmost importance to identify target transcripts of TRA2B, which constitute off-targets in SMA-related and general HDAC inhibitor treatment strategies.

This study aims at the *in vivo* identification of Tra2b splicing targets using a neuronal-specific knock-out mouse model. Therefore, *Tra2b* was depleted in neurons of the central nervous system by using Nestin-Cre transgenic mice. Neuronal-specific knock-out mice were viable until birth, but showed severe developmental defects of the brain. This allowed full developmental analysis as opposed to ubiquitously depleted *Tra2b* knock-out mice, which died very early during development.

Firstly, brains of conditional knock-out mice should be morphologically and histologically analyzed to identify the underlying reasons for the observed developmental defects. Secondly, differentially spliced Tra2b targets should be identified using whole transcriptome sequencing and exon arrays. In preparation to that and as a proof of concept, bona fide Tra2b target transcripts should be validated *in vivo* using whole brain RNA of neuronal-specific knock-out mice. Newly identified candidate transcript should be confirmed by quantitative PCRs. Thirdly, the splicing mechanism of respective exons and the effect of increased or decreased Tra2b levels should further be investigated in minigene splicing assays. Lastly, the identified splicing targets and their missplicing should be evaluated as to their physiological role and possible impact on neuronal development.

3. Materials

3.1. Devices and Utilities

Type of device	Designation	Manufacturer
Balance	ER	Kern
Balance (fine scale)	ARJ 120-4M	Kern
Centrifuge	Allegra X22-R	Beckmann Coulter
Centrifuge	Avanti J-20XPI	Beckmann Coulter
Centrifuge	5415R	Eppendorf
Centrifuge	5415D	Eppendorf
Centrifuge	5804	Eppendorf
Centrifuge	Galaxy Mini	VWR
Electrophoresis chamber	MGV-620T	C.B.S & Scientific
Electrophoresis chamber	SGE-020-02	C.B.S & Scientific
Electrophoresis chamber	E-H6	Febicon
Electrophoresis chamber	Mini-Protean 3 cell	BioRAD
Freezer (-80°C)	HERA freeze	Heraeus
Gel documentation	ChemiDoc XRS	BioRAD
Heating block	HTMR132	HLC
Heating cabinet	Function Line	Heraeus
Heating magnetic stirrer	MR3001	Heidolph
Homogenizer	T10 basic, Ultra Turrax	IKA
Incubator	Innova 44	New Brunswick Scientific
Incubator	Innova 4230	New Brunswick Scientific
Incubator, tissue culture	HERA cell 150	Heraeus
Laminar airflow cabinet	HERA safe	Heraeus
Light source (Fibre optic)	KL 1500 LCD	Leica
Microplate reader	Safire²	Tecan
Microscope (Fluorescence)	Imager M2	Zeiss
Microscope (stereo)	S8AP0	Leica
Microscope camera	AxioCam MRm	Zeiss
Microscope camera	AxioCam ICc 1	Zeiss
Microscope camera	AxioCam ERc 5s	Zeiss
Microtome	RM2255	Leica

Type of device	Designation	Manufacturer
Microwave	R-898(AL)-A	Sharp
Orbital Shaker	3015	GFL
pH meter	inoLab pH level 1	WTW
Photometer	BioPhotometer	Eppendorf
Photometer	NanoDrop 1000	Peqlab
Pipette	Research (2.5/10/20/100/ 200/ 1000/5000 µl)	Eppendorf
Pipette (automatic)	Research Pro (10/100 µl)	Eppendorf
Pipette (repetitive)	Multipette Plus	Eppendorf
Pipettor	Easypet	Eppendorf
Power supply	PowerPac Basic	BioRAD
Power supply	PowerPac HC	BioRAD
Power supply	PowerPac 1000	BioRAD
real-time Thermocycler	7500 Real Time PCR System	Applied Biosystems
real-time Thermocycler	Light Cycler 1.5	Roche
Roll incubator	SRT9	Stuart
Safety gas-burner	Fireboy Plus	IBS
Shaker	Type 3015	GFL
SpeedVac	Concentrator 5301	Eppendorf
Thermocycler	DNAengine Tetrad2	MJ Research
Thermocycler	DNAengine Dyad	MJ Research
Thermocycler	S1000	BioRAD
Thermocycler	C1000 Touch	BioRAD
Tissue processor	TP1028	Leica
Vacuum pump	PM12640-026.3	Biometra
Vortex	444-1372	VWR
Water bath	1083	GFL
Water bath	FBC620	FisherBrand

3.2. Chemicals

Chemical	Manufacturer
2-Mercaptoethanol (99 %, p.a.)	AppliChem
2-Propanol (≥ 99.5 %)	AppliChem
Acetic acid	AppliChem
Acetone	AppliChem
Agarose (low melting)	Biozym
Agarose for routine use	Sigma-Aldrich
Ammonium persulfate	AppliChem
Ampicillin	AppliChem
Bacto Agar	AppliChem
Bovine serum albumin (BSA)	Sigma-Aldrich
Bromophenol blue	AppliChem
Citrate	AppliChem
Coomassie Brilliant Blue R-250	AppliChem
D-glucose	AppliChem
Diethylpyrocarbonate (DEPC)	AppliChem
Dimethyl sulfoxide (DMSO)	Sigma-Aldrich
Disodium hydrogen phosphate (p.a.)	AppliChem
Dithiothreitol (DTT)	AppliChem
Emetine dihydrochloride hydrate	Sigma-Aldrich
Ethanol (≥ 99.5 %, p.a.)	AppliChem
Ethidium bromide (1 % in H_2O)	AppliChem
Ethylenediaminetetraacetic acid (EDTA)	AppliChem
Formaldehyde (37 %, p.a.)	Sigma-Aldrich
Glycerol (86 %, p.a.)	AppliChem
Glycine	AppliChem
Hydrochloric acid (37 %, fuming)	AppliChem
Hydroxymethylaminoethane (Tris)	AppliChem
Kanamycin	AppliChem
Magnesium chloride hexahydrate	Merck
Magnesium sulfate	Merck
Methanol (≥ 99.9 %, p.a.)	AppliChem
Milk powder (low fat)	AppliChem

Chemical	Manufacturer
Mowiol	Roth
Paraformaldehyde	Fluka
Potassium chloride (p.a.)	AppliChem
Potassium dihydrogenphosphate	AppliChem
Sodium chloride (p.a.)	AppliChem
Sodium dihydrogen phosphate dihydrate (p.a.)	AppliChem
Sodium dodecyl sulfate (SDS, 10 % in H_2O)	AppliChem
Sodium hydroxide	AppliChem
Tertramethylethylenediamine (TEMED)	AppliChem
Triton X-100	AppliChem
Trizma base	Sigma-Aldrich
Tryptone (microbiology grade)	AppliChem
Tween-20 (Polysorbate 20)	AppliChem
Water (HPLC grade)	Sigma-Aldrich
Xylene	AppliChem
Xylene cyanol	AppliChem
Yeast extract (microbiology grade)	AppliChem

3.3. Kits

Unless noted otherwise, all kit-related procedures have been performed according to the manufacturer's instructions.

Name	Manufacturer
BigDye Terminator V1.1 Sequencing Kit	Applied Biosystems
DNeasy Blood and Tissue Kit	Qiagen
EndoFree Plasmid Maxi Kit	Qiagen
gDNA Blood & Tissue Kit	Qiagen
GeneChip WT Terminal Labelling and Hybridization kit	Affymetrix
Hybridization Wash and Stain Kit	Affymetrix
Light Cycler Fast Start DNA Master SYBR green I	Roche
Mouse Exon 1.0 ST Array	Affymetrix
pcDNA 3.1 GFP-TOPO TA Expression Kit	Life Technologies
pcDNA 3.1/V5-HisTOPO TA Expression Kit	Life Technologies

Name	Manufacturer
Phase Lock Gel (Heavy 2 ml)	5 Prime
Pure Yield Plasmid Midiprep System	Promega
QIAprep Spin Miniprep Kit	Qiagen
QIAquick Gel Extraction Kit	Qiagen
QIAquick PCR Purification Kit	Qiagen
QIAshredder	Qiagen
Quant-iT RiboGreen RNA assay Kit	Life Technologies
QuantiTect Reverse Transcription Kit	Qiagen
RNase-free DNase Kit	Qiagen
RNeasy Mini Kit	Qiagen
Superscript II Reverse Transcription Kit	Life Technologies
Superscript VILO Reverse Transcription Kit	Life Technologies
Vectastain ABC	Vector Laboratories
WT Expression Kit	Ambion

3.4. Reagents and equipment

3.4.1. Reagents

Reagent	Manufacturer
1kb DNA ladder	Life Technologies
AquaPlus Mix 30% (29:1 Acrylamide/Bisacrylamide)	AppliChem
Bradford reagen	AppliChem
Chloroform/Isoamylalcohol (24:1)	AppliChem
Complete Mini Protease Inhibitors	Roche
dATP	Peqlab
dCTP	Peqlab
dGTP	Peqlab
dTTP	Peqlab
Eosin-Y solution	Sigma-Aldrich
Eukitt mounting medium	Fluka
GeneRuler 100bp DNA Ladder Plus	Thermo
Mayer’s Hematoxylin	Sigma-Aldrich
Pager Ruler Prestained Protein Ladder	Thermo

Reagent	Manufacturer
PBS (10x)	VWR
Ponceau S	Sigma Aldrich
QIAzol Lysis Reagent	Qiagen
Restore Western Blot Stripping Buffer	Thermo
Ripa buffer	Sigma-Aldrich
Super Signal West Pico ECL Substrate	Thermo
TBE (10x)	AppliChem

3.4.2. Enzymes

Enzyme	Manufacturer
Accu Prime DNA Polymerase	Life Technologies
BamHI (Endonuclease)	New England Biolabs
EcoRI (Endonuclease)	New England Biolabs
EcoRV (Endonuclease)	New England Biolabs
NotI (Endonuclease)	New England Biolabs
Platinum Taq DNA Polymerase	Life Technologies
Proteinase K	AppliChem
PstI (Endonuclease)	New England Biolabs
Recombinant Taq DNA Polymerase	Life Technologies
RNase A	Sigma-Aldrich
T4 DNA Ligase	Promega
XhoI (Endonuclease)	New England Biolabs

3.4.3. Cell culture reagents and media

Type	Name	Manufacturer
Antibiotic	Penicillin/Streptomycin	Life Technologies
Antibiotic	G418 (Geniticin)	Sigma-Aldrich
Antibiotic	Zeocin	Life Technologies
Fungicide	Amphotericin B	Promocell
Medium	PBS	Life Technologies
Medium	DMEM	Life Technologies
Medium	Opti MEM	Life Technologies

Type	Name	Manufacturer
Medium	Trysin / EDTA	Life Technologies
Serum	Fetal Bovine Serum	Biochrom
Transfection	DharmaFECT I	Thermo
Transfection	Lipofectamine 2000	Life Technologies
Transfection	Mirus TransIT LT	Mirus

3.4.4. Equipment for laboratory mouse work

Type	Name	Manufacturer
Ear tag applicator	1005-s1	National Band & Tag Co.
Ear tags	1005-1	National Band & Tag Co.
Forceps	FM002R	Aesculap
Forceps	DB047R	Aesculap
Scissors	BC321R	Aesculap
Scissors	BC8641R	Aesculap
Scissors	BC100R	Aesculap

3.5. Solutions and media

Name	Composition
Ampicillin stock solution	50 mg/ml in 50 % EtOH
Citrate buffer	10 mM citrate; pH 6
Coomassie destaining solution	10 % acetic acid; 30 % Methanol in water
Coomassie staining solution for acrylamide gels	0.25 % Coomassie Brilliant Blue R-2500; 10 % acetic acid; 50 % Methanol in water; filter
DEPC H_2O	0.1 % (v/v) DEPC in H_2O; stir over night; autoclave
DNA sample buffer, 10x	100 mM EDTA (pH 8); 1 % (w/v) SDS; 50 % (v/v) Glycerin; 0.1 % (w/v) Bromophenol blue; 0.25 % (w/v) Xylene cyanol
dNTP Mix (1.25 mM per dNTP)	5 µl dATP; 5 µl dCTP; 5 µl dGTP; 5 µl dTTP; 480 µl H_2O
Kanamycin stock solution	50 mg/ml in H_2O
LB agar	3.75 g Bacto Agar in 250 ml LB-medium

Name	**Composition**
LB medium	5 g Bacto-Trypton; 2.5 g Bacto-yeast-extract; 2.5 g NaCl; ad 500 ml ddH_2O
Lysis buffer (tissue)	200 mM NaCl; 100 mM Tris-HCl; 5 mM EDTA; 0.2 % (w/v) SDS; pH 8.5; 200 µg/ml proteinase K (fresh)
Paraformaldehyde (4 %)	4 % paraformaldehyde, dilute in 30 ml H_2O at 60°C; add 1 M NaOH to clear the solution; cool down and filter; add 50 ml 2x PBS, pH 7.4
Resolving gel (SDS-PAGE), 12 %	3.3 ml H_2O; 4 ml 30 % acrylamide mix; 2.5 ml 1.5 mM Tris pH 8.8; 100 µl 10 % SDS; 100 µl 10 % APS; 4 µl TEMED
SDS Electrophoresis buffer, 10x	0.25 M Tris; 1.92 M Glycine; 1 % (w/v) SDS; pH 8.3
SDS sample buffer (2x)	0.757 g Tris-base; 20 ml glycerol; 10 mg Bromophenol blue; 6 g SDS; ad 90 ml H_2O; add 1/10 Vol. β-ME before use
SOC medium	2 % (w/v) Tryptone, 0.5 % (w/v) yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 20 mM D-Glucose, pH 7
Stacking gel (SDS-PAGE), 4 %	1.4 ml H_2O; 330 µl 30 % acrylamide mix; 250 ml 1.5 mM Tris pH 6.8; 20 µl 10 % SDS; 20 µl 10 % APS; 2 µl TEMED
TBST	0.5 % (v/v) Tween-20 in TBS
TE buffer	10 mM Tris; 1 mM EDTA; pH 8
TE^{-4} buffer	10 mM Tris-HCl; 0.1 mM EDTA; pH 8
Transfer buffer (Western Blot)	15 mM Tris-base; 150 mM Glycine; 20 % (v/v) Methanol

3.6. Cell culture

3.6.1. Cell lines

The following cell lines were employed in this study:

- HEK293T cells: Cells are derived from human embryonic kidney and were originally transformed by using sheared adenovirus 5 DNA (Graham et al. 1977). HEK 293**T** cells carry the SV40 large-T antigen.
- NSC34 cells: Cells were produced by fusion of **n**euroblastoma cells with embryonic **s**pinal **c**ord cells. NSC34 cells display a neuron-like phenotype are considered to resemble many aspects of motoneuron development (Cashman et al. 1992).

3.7. Mouse inbread strains

The *Tra2b* knock-out mouse model has previously been generated by Dr. Ylva Mende (Mende et al. 2010). The genomic *Tra2b* locus was altered by homologous recombination to insert loxP recombination sites flanking the exon 4 of *Tra2b*. Cre-recombinase-mediated recombination results in deletion of exon 4 causing a shift in the codon frame. This results in multiple premature termination codons in the downstream sequence. Here the *Tra2b* mouse line was crossbread with a *Nestin-Cre* transgenic mouse line (Tronche et al. 1999) to achieve neuronal-specific knock-out of *Tra2b*. All animals were maintained on a >96 % C57BL/6N background.

3.8. Antibodies

3.8.1. Primary antibodies

Antibody	Host species	Blocking reagent	Dilution	Incubation time	Supplier	Product number
α-Caspase3	rabbit, polyclonal	milk	IHC: 1:1,000	o.n.	R&D Systems	AF835
α-Tra2b	rabbit, polyclonal	IHC: milk WB: milk	IHC: 1:8,000 WB: 1:2,000	o.n. o.n.	Epigentech	Produced
α-Ki67	rabbit, polyclonal	milk	IHC: 1:1,000	o.n.	Abcam	Ab15580
α-p21	rabbit, polyclonal	milk	WB: 1:1,000	o.n.	ProteinTech	10355-1-AP
α-GFP	mouse, monoclonal	milk	WB: 1:250	o.n.	N/A	N/A
α-β-actin	mouse, monoclonal	milk	WB: 1:30,000	1 hours	Sigma-Aldrich	A5316

3.8.2. Secondary antibodies

Antibody	Host species	Blocking reagent	Dilution	Incubation time	Supplier	Product number
α-rabbit IgG (biotinylated)	goat	milk	1:2,000	30 min	Vector Laboratories	PK-6101
α-mouse IgG (HRP conjugate)	goat	milk	1:5,000	1 h	Jackson Immuno Research	N/A
α-rabbit IgG (HRP conjugate)	goat	milk	1:5,000	1 h	Cell Signaling	7074S

3.9. Oligonucleotides and PCR conditions

All oligonucleotides used and optimized in this study are listed. *Amplicon* defines the target of the PCR including the species or the exon if applicable. Oligonucleotide sequences are listed in 5' to 3' direction, while forward primers are designated as *fw*, reverse primers as *rev*. The given identification number corresponds to the inhouse primer database. *Product* gives the length of the PCR product(s) in base pair. *Platform* states the kit, polymerase or PCR system that has been used for the respective primer pair. *AccuPrime* and *Platinum Taq* refer to the respective proof reading polymerase kits from *Life Technologies*. *PCR* refers to *Recombinant Taq DNA Polymerase* (*Life Technologies*). *RT-PCR* refers to semi-quantitative PCR on a cDNA template using *Recombinant Taq DNA Polymerase* (*Life Technologies*). *Roche LC* indicates primer usage for quantitative real-time PCR on the *Roche Light Cycler 1.5*. The column *°C* indicates the optimal annealing temperature. *$MgCl_2$* or *$MgSO_4$* gives the optimal concentration of $MgCl_2$ or $MgSO_4$ in the respective PCR. *Cycles* indicate the typical number of

PCR cycles used. In quantitative real-time PCRs the cycle number is indicated as RTQ, as the cycle number is irrelevant.

3.9.1. Oligonucleotides for cloning

Amplicon	Oligonucleotide 5'>3'			product [bp]	Platform	°C	$MgSO_4$ [mM]	Cycles
mmSgol2 ex4 minigene	4537	fw	AAAAAACTCGAGGTTTTGGACTCTCCATTTAAATGG	2138	Accu Prime	55	0.4 mM $MgSO_4$	35
	4538	rev	AAAAAAGGATCCCTTAGCAGTTAGGGAAGGAAACC					
hsSGOL2 ex4 minigene	4533	fw	AAAAAACTCGAGCCATGTAAGTCTATGTCAAAGTAC	1977	Accu Prime	55	0.4 mM $MgSO_4$	35
	4534	rev	AAAAAAGGATCCATCCTGATGGGAATAATCCAGAAG					
mmTubd1 ex4 minigene	4539	fw	AAAAAACTCGAGAGCTCTTGCCCAGTAGTTCC	1238	Accu Prime	55	0.4 mM $MgSO_4$	38
	4540	rev	AAAAAAGGATCCCTCTAGTGAGTATTTACCAAGGAG					
hsTUBD1 ex 4 minigene	4535	fw	AAAAAACTCGAGCACATATCCAAATATCACATGTACC	1123	Accu Prime	55	0.4 mM $MgSO_4$	38
	4536	rev	AAAAAAGGATCCCATTCTGGAAGCCATAACTTAAC					
mm tNasp	4594	fw	AAAAAACTCGAGCTATGGCCACAGAGTCTACAGC	sNasp: 1347	Platinum Taq	60	2 mM $MgSO_4$	38
	4595	rev	AAAAAAGGATCCTTAACATGCAGTGCTTTTAATTGTAG	tNasp: 2322				
pSPL3 Exon trap vector sequencing primers	4508	fw	TCTGAGTCACCTGGACAACC	seq	-	-	-	-
	4541	fw	GTGAACTGCACTGTGACAAGCTGC	seq	-	-	-	-
	4542	fw	CAACCTCAAAGGCACCTTTGC	seq	-	-	-	-
	4544	fw	AAATTGTGGGTCACAGTCTATTATGG	seq	-	-	-	-
	4509	rev	ATCTCAGTGGTATTTGTGAGC	seq	-	-	-	-
	4543	rev	CACCTTCTTCTTCTATTCCTTCG	seq	-	-	-	-
	4545	rev	ATCTGATGCACAAAATAGAGTGGTG	seq	-	-	-	-
NLS-Cre-recombinase	4083	fw	ATGAAGAAGAAGAGGAAGGTGTCC	1060	Taq	56	1.5	34
	4084	rev	GGTGCTCGAGCTGATCGCCATCTTCC					
NLS-Cre sequencing	4085	fw	ACTGGTTATGCGGCGGATCC	seq	-	-	-	-
	4028	rev	GGACAGGGCCATCGCCAATTGG	seq	-	-	-	-

3.9.2. Oligonucleotides for genotyping

Genotyping of mice was performed as previously described using the following primers and conditions (Mende et al. 2010).

Amplicon	Oligonucleotide 5'>3'			product [bp]	Platform	°C	$MgCl_2$ [mM]	Cycles
Nestin-Cre transgene	3478	fw	CGCTTCCGCTGGGTCACTGTCG	tg: 500	PCR	58	1.5	35
	3479	rev	TCGTTGCATCGACCGGTAATGCAGGC					
mmTra2b (floxed/wt)	3482	fw	AAGGCGTTCTAGATCAAAGTCCAG	wt: 821 flox: 950	PCR	60	1.5	35
	2340	rev	CGAGAGGGCACGAGAGGACAATC					
mmTra2b (wt/-)	2507	fw	GTAGTGAAGGTCTGTTACCC	wt: 827 KO: 256	PCR	58	1.5	38
	3263	rev	CACCAGAAAATTTAACAAATTAAAC					

3.9.3. Oligonucleotides for semi-quantitative RT-PCR and quantitative PCR

Amplicon	Oligonucleotide 5'>3'			product [bp]	Platform	°C	$MgCl_2$ [mM]	Cycles
mmAppbp2 E4	4277	fw	GGCTGGTACAGTGATGCTGAAAAGG	118	Roche LC	61	3	RTQ
	4278	rev	GCACATGAAGCAACCTCACACAACACTCC					
mmAppbp2 Δ4	4279	fw	CATTTGCTCCACCACTGTTTTCAGG	166	Roche LC	61	3	RTQ
	4280	rev	CGCACATGAAGCAACCTAAGACAAAGC					
mmCasp1	4565	fw	TTATTTCTCAATCTGTATTCACGCC	FL: 541	RT-PCR	60	2.5	32
	4566	rev	CTCTGAAGGATTTTCTTTCCATAAC	Δ2: 274				
mmCasp1 Δ2	4567	fw	ACGCCATGGCTGCTCCATC	186	RT-PCR	58	1.5	32
	4566	rev	CTCTGAAGGATTTTCTTTCCATAAC					
mmCdkn1a	4491	fw	GAGCAAAGTGTGCCGTTGTCTCTTCG	414	RT-PCR	60	1.5	30
	4492	rev	CTGCGCTTGGAGTGATAGAAATCTGTCAG					
mmCdkn1a	4493	fw	AGGCCCAGTACTTCCTCTGCCCTG	195	Roche LC	62	1.5	RTQ
	4492	rev	CTGCGCTTGGAGTGATAGAAATCTGTCAG					
mmCltb E5 (CltbII)	3939	fw	ATGGAGATGTGTTTCAGGAGGCTAAC	276	Roche LC	63	3	RTQ
	3941	rev	GCTGCTGGTAGAACGCTTTGTC					
mmCltb Δ5 (CltbIII)	3939	fw	ATGGAGATGTGTTTCAGGAGGCTAAC	258	Roche LC	64	3	RTQ
	3940	rev	CTTCCGATGCCCTGTTGTTGATC					
mmEomes E4	4494	fw	ATAAGATGTACGTTCACCCAGAATC	161	RT-PCR	60	1.5	32
	4495	rev	AGCCTCGGTTGGTATTTGTGC					
mmEomes Δ4	4496	fw	ACAACAACAACACACAGATCACC	118	RT-PCR	60	1.5	32
	4497	rev	GATGGAGTTAACCTGTCATTTTCTG					
mmEomes	4502	fw	ATCAGATTGTCCCTGGAGGTC	322	RT-PCR	60	1.5	34
	4503	rev	AAGGTCGGGTCAGGGTAATAC					
mmEzh2 Δ17	4596	fw	AGAATACTGTGGGGAGATTTTGTG	281	RT-PCR	60	1.5	30
	4597	rev	GCAGCTGTTTCAGAGAGAAGG					
mmFgfr2 IIIb	3992	fw	CCTGAAGCACTCGGGGATAAATAGC	155	Roche LC	65	3	RTQ
	3993	rev	CTTGCTGTTTGGGCAGGACAGTG					
mmFgfr2 IIIc	3611	fw	CAGCATCGGAGCCTTATTATGG	365	Roche LC	65	3	RTQ
	3612	rev	GTTACATTCCGAATATAGAGAACC					
mmGapdh	3999	fw	GGCTGCCCAGAACATCATCC	169	Roche LC	63	3	RTQ
	4000	rev	GTCATCATACTTGGCAGGTTTCTC					
mmGca	4569	fw	GTTATTCTGACAGCTACTCCCCTG	FL: 257	RT-PCR	60	1.5	32
	4570	rev	GTTCTGCTTCCAGGCATTAAGAG	Δ4: 213				
hsHPRT	4262	fw	GCTATTGTAATGACCAGTCAACAGGGGAC	168	Roche LC	62	3	RTQ
	4263	rev	CCTTGACCATCTTTGGATTATACTGCC					
mmHprt	2811	fw	AAGGAGATGGGAGGCCA	209	RT-PCR	60	1.5	>26
	2812	rev	GTTGAGAGATCATCTCCACCAAT					
mmHtt E60	4273	fw	CCCAGTTTGAAATGATGTATCTGACGC	136	Roche LC	63	3	RTQ
	4274	rev	GCCACAGTTTTGTCCATTCCAAGGACAGC					
mmHtt Δ60	4275	fw	GAAGTGGTTCGATCTGACAAAACTGTGGC	186	Roche LC	63	3	RTQ
	4276	rev	GGACAGCAGATAGTCACTAACAACTGG					
mmKif11 E12	4498	fw	AAGTTTTAGAGCCATGAATGGAAAG	367	RT-PCR	60	1.5	32
	4499	rev	CTCATCGATTGCTCTCTTGCG					
mmKif11 Δ12	4500	fw	AGGAGCTCAGTAAGTTGCTTAAC	179	RT-PCR	58	2.5	32
	4501	rev	ACTGCCATCCTTAATCAATTCTTC					
mmMapt E10 (4R)	3990	fw	TATCAAACACGTCCCGGGTGGAGGC	257	Roche LC	67	3	RTQ
	3991	rev	CATTCTCCCTGAAGGTCAGCTTGTGG					

Amplicon	Oligonucleotide 5'>3'			product [bp]	Platform	°C	$MgCl_2$ [mM]	Cycles
mmMapt Δ10 (3R)	3931	fw	GAGGTGGCAAGGTGCAAATAGTCTAC	291	Roche LC	61	3	RTQ
	3934	rev	CCACGGGTGACTTATACACAATTTCTGC					
mmNasp E7	4049	fw	GGAGTGCATGTAGAAGAGG	115	Roche LC	65	1	RTQ
	4042	rev	CGTCATAAACCTGTTCTCTC					
mmNasp Δ7	4037	fw	AATGGAGTGTTGGGAAATGC	118	Roche LC	62	3	RTQ
	4043	rev	CTGAGCCTTCAGTTTCATCTAC					
mmNrg1 E5	4504	fw	GAAGTCAGAGCTTCGAATCAAC	238	RT-PCR	60	2.5	32
	4505	rev	GGATGTAGATGTGGATGAAGAAG					
mmNrg1 Δ5	4506	fw	CTATGTGTCCTCAGCCACATC	176	RT-PCR	60	1.5	32
	4507	rev	GGCAACGATCACCAGTAAACTC					
mmPrkd3 E18	4577	fw	TGCGTTTATGTACCCACCAAATC	172	RT-PCR	60	2.5	30
	4578	rev	AGTTTCAAATTCTCTGAGGTCAAGC					
mmPrkd3 Δ18	4580	fw	GAGAAATTTCCAGTGAAGGATTATC	156	RT-PCR	62	2.5	32
	4579	rev	CATAATGAAATGCTTTGGGTATTCC					
mmRcor2 E4	4581	fw	GCTATAGCAACAAGGAGCTGAAG	168	RT-PCR	60	2.5	32
	4582	rev	GTGATTTCTCCACATCGTGTTTATG					
mmRcor2 Δ4	4584	fw	GTGTATCAGATGCCAAACGCAC	198	RT-PCR	60	2.5	32
	4583	rev	AATCACCTTGTCAGGGAGCATC					
pSPL3 minigene	4508	fw	TCTGAGTCACCTGGACAACC	262 + exon size	RT-PCR	62	1.5	28
	4509	rev	ATCTCAGTGGTATTTGTGAGC					
mmSgol2 E4	4510	fw	TCGAAGAATTACTACCGAAAAGATG	182	RT-PCR	60	1.5	32
	4511	rev	CTTGGTGGAACTCAGAAAGACTG					
mmSgol2 Δ4	4510	fw	TCGAAGAATTACTACCGAAAAGATG	113	RT-PCR	60	3	34
	4521	rev	GAAAAGAACCTTGGTGGAACAAG					
mmSlu7 E15	4512	fw	TGGTGCGACTGCAGAAGAG	145	RT-PCR	60	2.5	32
	4513	rev	CTTCCGTTCCTCATCGTCG					
mmSlu7 Δ15	4522	fw	CCTCCTGGAGGCACTGAATG	152	RT-PCR	60	2.5	32
	4523	rev	TCTGCCGTTTCATCCTGTAGG					
mmSox6 E5	4514	fw	AATGAACAGTGTTACTTTTGGAACC	259	RT-PCR	60	2.5	30
	4515	rev	CTGGGTAATCATGGTGGAGAG					
mmSox6 Δ5	4524	fw	CGTCTACCTCACCACATAAGC	176	RT-PCR	60	2.5	32
	4525	rev	CTCTCAGGTGTACCCTCTTG					
mmSt18	4587	fw	ATGGTCTGGGTCATGTAAACAAC	FL: 313	RT-PCR	60	2.5	32
	4588	rev	AGCAGACTCTCATTGCTCTGTTC	Δ22: 209				
mmTcf4 E18	4516	fw	GAACTAGTCTTCTCTCAGCCAACAG	443	RT-PCR	60	2.5	30
	4517	rev	GTTGATATCTCGGACCCTCAG					
mmTcf4 Δ18	4526	fw	AAGACCCTTACAGAGCAATAACG	129	RT-PCR	60	2.5	32
	4518	rev	AAAGCCTCGTTGATATCTCGGAC					
mmTop2a E11	4519	fw	GCTTTGGATCAACATGTCAATTAAG	231	RT-PCR	60	2.5	30
	4520	rev	AGTCTCCCTCAGTCAGGATAAGTG					
mmTop2a Δ11	4527	fw	GAAAAGTTCATCAAAGCTGGAGC	154	RT-PCR	62	2.5	32
	4528	rev	ACATTGAGTATCTTTCCTCTAAGAG					
hsTRA2B E2	4563	fw	GGAAGATGAAAGAAGTCAGAATTTG	250	Roche LC	64	3	RTQ
	4564	rev	AGACCTAGATTCAGATCGGGAC					
hsTRA2B Δ2	4562	fw	CTACGGCGAGCGGGAATC	151	Roche LC	64	3	RTQ
	4564	rev	AGACCTAGATTCAGATCGGGAC					
mmTra2a E2	4040	fw	GTTGTAGCCGTCGCCTTCT	206	Roche LC	64	3	RTQ
	4044	rev	CTTGATTTATCTTCCACATTCTTGG					

Amplicon	Oligonucleotide 5'>3'			product [bp]	Platform	°C	$MgCl_2$ [mM]	Cycles
mmTra2a Δ2	4040	fw	GTTGTAGCCGTCGCCTTCT	155	Roche LC	66	3	RTQ
	4046	rev	GAGACTCTCTGCCCTCGAAG					
mmTra2b E2	4269	fw	GAACTACGGCGAGCGGGTTAATG	161	Roche LC	64	3	RTQ
	4270	rev	CAAGTGGGACTTCTGGTCTGATAATTAGC					
mmTra2b Δ2	4271	fw	AGAACTACGGCGAGCGGGAATC	191	Roche LC	64	3	RTQ
	4272	rev	CCTTGTATAATGCCTTCTAGAACTTCTTC					
mmTra2b	3218	fw	TAGAAGGCATTATACAAG	186	Roche LC	63	3	RTQ
	3219	rev	CTCAACCCAAACACGC					
mmTubd1 E4	4529	fw	CAGTACTCCAGTTCCTTGAAAATG	157	RT-PCR	60	1.5	34
	4531	rev	AGATCTTATGGACAGCGTCATTCTC					
mmTubd1 Δ4	4532	fw	GGACAACGAGGCATATCAAGC	225	RT-PCR	60	1.5	36
	4530	rev	GTTCTGCACAATGACCCGTATG					
mmHoxa3	4546	fw	GTCTCTCCCCCTCAAAGTGC	340	RT-PCR	60	1.5	32
	4547	rev	AGATCTTGATCTGGCGCTCG					
mmHoxa4	4548	fw	GCGCCGTCAACTCCAGTTAT	255	RT-PCR	60	1.5	32
	4549	rev	CCGAGGCAGTGTTGGAAGAT					
mmHoxa5	4550	fw	AATGGCATGGATCTCAGCGT	473	RT-PCR	60	1.5	32
	4551	rev	GTCTGGTAGCGAGTGTAGGC					
mmHoxa6	4552	fw	CACCGACCGGAAGTACACAA	249	RT-PCR	60	1.5	32
	4553	rev	CACTTCATGCGCCGATTCTG					
mmHoxb5	4554	fw	CTCTGAGCGGCTCTTACAGG	496	RT-PCR	60	1.5	32
	4555	rev	CCGTCTGGCCCAGTCATATC					
mmHoxb6	4556	fw	ACAAGAGCGTGTTCGGAGAG	262	RT-PCR	60	1.5	32
	4557	rev	CGCCGGTTCTGAAACCAAAT					
mmHoxb7	4558	fw	GCTCGAACCGAGTTCCTTCA	303	RT-PCR	60	1.5	32
	4559	rev	TGTCTTTCGGTGAGGCAGAG					
mmHoxb8	4560	fw	ACTTCTACGGCTACGACCCT	294	RT-PCR	60	1.5	32
	4561	rev	CGTGCGATACCTCGATCCTC					
mmHoxc5	4562	fw	AGCTAAGAGCAGTGGGGAGA	298	RT-PCR	60	1.5	32
	4563	rev	CCACTTCATCCTGCGGTTCT					
mmHoxc6	4564	fw	ACCGGATCTACTCGACTCCC	337	RT-PCR	60	1.5	32
	4565	rev	TCCTTCTCCAGTTCCAGGGT					
mmHoxd3	4566	fw	TATGAGACCTGGCACTGGGA	385	RT-PCR	60	1.5	32
	4567	rev	GCGGTTGAAGTGGAACTCCT					
SV40 T-antigen	4052	fw	CCAGGATTTAAGGAAGAAGC	142	RT-PCR	58	1.5	30
	4053	rev	GGAGTTTCATCCTGATAAAGG					

3.10. Small interfering RNAs

Target transcript	Gene accession no.	Sense strand	Supplier	Cat. No.
hTRA2B	NM_001243879	5'-GGACUACUAUAGCAGAUCATT-3'	Qiagen	SI02653252
hTRA2B	NM_001243879	5'-CGAUCUGAAUCUAGGUCUATT-3'	Qiagen	SI02653504

3.11. Software, web applications and databases

3.11.1. Routinely employed software

Name	Purpose	Manufacturer
AutoStitch 2.2	Image processing (panoramic assembly)	University of Britsish Columbia
EndNote X3	Reference management	Thomson Reuters
Excel 2003/2007	Plots, charts, data processing, statistics	Microsoft
Image J 1.40g	Densitometric analysis	NIH
Lasergene Suite	Sequence analysis	DNAstar Inc.
Light Cycler Software	Quantitative real-time PCR analysis	Roche
Paint Shop Pro Photo XII	Image processing	Corel
Power Point 2003/2007	Image illustration	Microsoft
Quantity One 4.5.1	Densitometric analysis	BioRAD
Word 2003/2007	Word processing	Microsoft
ZEN Blue Edition	Fluorescence imaging and analysis	Zeiss
Mausoleum	Laboratory animal management	Dr. Hanns-Eugen Stöffler
PyRAT	Laboratory animal management	Scionics

3.11.2. Routinely consulted databases and web applications

Name	URL
AnalyzerSpliceTool	http://ibis.tau.ac.il/ssat/SpliceSiteFrame.htm
Ensembl	http://www.ensembl.org/index.html
ESE Finder	http://rulai.cshl.edu/cgi-bin/tools/ESE3/esefinder.cgi?process=home
ExonPrimer	http://ihg.gsf.de/ihg/ExonPrimer.html
Genecards	http://www.genecards.org/
Ingenuity Systems	http://www.ingenuity.com/
MGI	http://www.informatics.jax.org/
NCBI	http://www.ncbi.nlm.nih.gov/
OligoCalc	http://basic.northwestern.edu/biotools/OligoCalc.html
PrimerBLAST	http://www.ncbi.nlm.nih.gov/tools/primer-blast/
Pubmed	http://www.ncbi.nlm.nih.gov/pubmed
UCSC	http://genome-euro.ucsc.edu/
UniProt	http://www.uniprot.org/

3.12. Plasmids

Vectors that were produced in frame of this study are labeled with an asterisk and a reference to the section describing the generation of the respective vector. For amplification chemicompetent E. coli (TOP10 or DH5α) were transformed with the respective vector using heat shock transformation (4.3.1). Vectors were conserved as bacterial glycerol stocks (4.3.3).

Vector	**Insert**	**Resistance (in bacteria)**
pBABE-zeo-largeT (Hahn et al. 2002b)	SV40 large-T antigen	Ampicillin
pcDNA3.1-CT-GFP-NLS-Cre *(5.2.2)	NLS-Cre-recombinase	Ampicillin
pEGFP-C2 (Clontech Laboratories)	none	Kanamycin
pEGFP-C2-HTRA2B (Hofmann et al. 2000)	HTRA2B	Kanamycin
pEGFP-C2-tNasp *(5.8.2)	tNasp	Kanamycin
pMAX-GFP (Amaxa Biosystems)	GFP	Kanamycin
pSPL3 (Life Technologies)	none	Ampicillin
pSPL3_hsSGOL2_770nt-ex4-1129nt *(5.7.1)	hsSGOL2 exon4 minigene	Ampicillin
pSPL3_hsTUBD1_468nt-ex4-438nt *(5.7.1)	hsTUBD1 exon4 minigene	Ampicillin
pSPL3_mmSgol2_1062nt-ex4-998nt *(5.7.1)	mmSgol2 exon4 minigene	Ampicillin
pSPL3_mmTubd1_468nt-ex4-553nt *(5.7.1)	mmTubd1 exon4 minigene	Ampicillin
pTriEx-1-HTNC (Peitz et al. 2002)	His-TAT-NLS-Cre-recombinase	Ampicillin

4. Methods

4.1. Eukaryotic cell culture

All cell culture procedures were performed under sterile conditions using a laminar flow workbench. To prevent contamination with bacteria or fungi the growth medium was routinely complemented with penicillin, streptomycin and amphotericin B unless noted otherwise. Eukaryotic cells were grown in sterile incubators at 37°C, 5 % CO_2 and 95 % relative humidity.

4.1.1. Cultivation of eukaryotic cells

HEK293T, NSC34 and murine fibroblasts were cultivated in D-MEM (*Dulbeccos's Modified Eagle Medium*) complemented with 10 % fetal bovine serum. Routinely, cells were cultivated in 75 cm² or 25 cm² tissue culture flasks and split into new flasks when reaching confluence. Cells were washed with 1x PBS without Mg^{2+} and Ca^{2+} and trypsinized using Trypsin-EDTA until cells detached from the surface. Trypsinization was stopped with at least 1 volume of culture medium. The cell suspension was homogenized and evenly distributed to new tissue culture flasks. Depending on the cell type and proliferation rate cells were distributed at different density. Routinely, HEK293T cells were split 1:40, NSC34 cells 1:5, primary fibroblast 1:5 and immortalized murine fibroblasts cell lines 1:20.

4.1.2. Counting cells using a Neubauer cytometer

Some applications require defined cell numbers and growth densities to work optimally. To determine the number of cells in a suspension, a Neubauer counting chamber was used. After trypsinization of cells (4.1.1) cells were well homogenized in excess of culture medium. Roughly 40 µl of the homogenized suspension were applied to the counting chamber. Under a standard microscope the cell number was determined in 4 quadrants of the counting chamber and the mean was calculated. The cell concentration in the cell suspension is mean x 10^4 cells per ml.

4.1.3. Generation of murine embryonic fibroblasts (MEF)

Murine embryonic fibroblasts were isolated from 13.5 dpc mouse embryos. Mice of the desired genotype were crossbred in a timed breeding (4.2.4) to determine the progression of pregnancy. Pregnant female mice were sacrificed by cervical dislocation and embryos were prepared from the uterus under sterile conditions and washed in PBS. Visible organs like heart and liver were removed from the embryos. The heads were used for the isolation of DNA for genotyping purposes. The embryo was ground in a 70 µm nylon cell filter using a plunger of a

20 ml syringe. The ground material was flushed through the cell filter with 25 ml of cell culture medium freshly complemented with antibiotics. Cells were pelleted by spinning for 5 minutes at 200 x g. Cells were resuspended in 6 ml of culture medium and plated into a 6 cm tissue culture plate. After 24 h cells were 3 times carefully washed in PBS to remove cellular debris. Cells were expanded while providing fresh culture medium daily.

4.1.4. Transfection of HEK293T and NSC34 cells

For transfection of HEK293T and NSC34 cells, best efficiency and lowest cytotoxicity was achieved using the *DharmaFECT I transfection reagent* (Thermo). Cells were either transfected with plasmid DNA (3.12), siRNA (3.10) or both simultaneously. Experiments were carried out in 6-well tissue culture plates with a final cultivation volume of 2 ml. To transfect 2 µg of plasmid DNA and/or 50 pmol of siRNA, 4 µl of *DharmaFECT I transfection reagent* were routinely used. 2.5 x 10^5 HEK293T or NSC cells were plated per 6-well 1 day prior to transfection in D-MEM complemented with fetal bovine serum without antibiotics. For transfection experiments assayed after 72 h or 96 h the cell number was reduced down to 1.8 x 10^5 cells to prevent overgrowth during the experiment. Transfection mixes were prepared according to the manufacturer's instructions using *OptiMEM* to pre-dilute nucleic acids and transfection reagent. The completed transfection mix was complemented to 2 ml with culture medium without antibiotics. For RNA analysis cells were harvested after 24 h, 36 h or 48 h. Proteins were isolated after 48 h, 72 h or 96 h.

To survey transfection efficiency cells were transfected with the *pMAX-GFP* vector allowing for easy assessment of the transfection rate via GFP fluorescence. As a negative control for RNAi experiments the *AllStars negative siRNA* (*Qiagen*) was used.

4.1.5. Stable transfection of primary murine embryonic fibroblasts

In this study MEFs were stably transfected with *pBABE-zeo-largeT* and *pcDNA3.1-CT-GFP-NLS-Cre*. This allowed stable expression of the *SV40 large T-antigen* (immortalization) (Hahn et al. 2002a) and stable expression of *Cre-recombinase*. Best results for the transfection of primary MEFs were obtained using the *Mirus TransiT LT transfection reagent* (Mirus). 1 x 10^5 cells were plated to a 6-well tissue culture plate 18 h prior to transfection. Transfection mixes were prepared as recommended by the manufacturer. Best result were achieved using 2.8 µl transfection reagent per µg of plasmid DNA and a total of 2.5 µg of plasmid DNA per 6-well. Transfection efficiency was surveyed by transfecting *pMAX-GFP* to assess the rate of transfection by GFP fluorescence. Selection for stably transfected cells was started one to two days after transfection. Routinely, selection was continued for 2 months using an appropriate

antibiotic. Cells transfected with *pBABE-zeo* were selected with 200 µg/ml *Zeocin* and cells transfected with *pcDNA3.1* were selected with 2,000 µg/ml *G418*.

4.1.6. Treatment of cells with emetine

To allow detection of transcripts undergoing nonsense-mediated decay (NMD) cells were treated with the protein synthesis inhibitor emetine. Immortalized MEFs were grown in culture medium until reaching ~90 % confluence. Emetine was provided to the cells at a concentration of 100 µg/ml for 10 h. Cell were harvested directly and RNA was isolated for quantitative analysis.

4.1.7. Cryoconservation of cells

Confluent cell layers were detached from the surface by trypsinization (4.1.1). Cells were pelleted by centrifugation at 200 x g for 5 minutes and resuspended in cryo-medium containing DMSO as a cryoprotective agent. For all cell types used in this study a cryomedium containing 10 % DMSO in fetal bovine serum was applicable. Cells from a growth area of 25 cm² were resuspended in 1 ml of cryo-medium and transferred to a cryo-tube. These were frozen to -80°C using an isopropanol-based freezing aid, which allowed controlled cool down rates of approx. 1°C per minute. For long term storage cells were transferred to liquid nitrogen storage.

To take frozen cells back to culture the frozen cryo-tube was quickly thawed in a 37°C water bath. The thawed cell suspension was transferred to pre-warmed culture medium and immediately plated to a 25 cm² tissue culture flask.

4.2. Working with laboratory mice

4.2.1. Mouse experiments

All procedures involving laboratory animals were performed according to the German laws of animal welfare and were approved by the '*Landesamt fuer Natur, Umwelt und Verbraucherschutz NRW*' under the reference numbers 50.203.2-K 38, 19/05 and 8.87-51.05.20.10.106. Animals used in this study were kept in the animal facility of the Institute for Genetics at the University of Cologne and in the central animal facility if the Center of Molecular Medicine Cologne (CMMC).

4.2.2. Generation of conditional knock-out mice

Mice breedings performed in this study are based on the *Tra2b*$^{flneo/+}$ and *Tra2b*$^{fl/+}$ lines generated by Ylva Mende (Mende et al. 2010). These lines have previously been backcrossed to achieve >96 % C57BL/6N background. To maintain this background all *Cre*-lines used in this study were also on C57BL/6N background.

Homozygously floxed *Tra2b* animals were generated by crossbreeding *Tra2b*$^{fl/+}$ animals with each other. To generate neuronal-specific *Tra2b* knock-out mice Tra2b$^{fl/fl}$ mice were first crossbred with *Nestin-Cre*$^{tg/0}$ animals to obtain *Tra2b*$^{fl/+}$*; Nestin-Cre*$^{tg/0}$ animals. *Nestin-Cre* transgenic mice express *Cre-recombinase* under control of the rat *nestin* promoter and enhancer and enable expression of *Cre-recombinase* in neuronal and glial precursors (Dahlstrand et al. 1995; Tronche et al. 1999; Mignone et al. 2004). Nestin expression is first activated at 10.5 dpc in the neural tube, the somites and in migrating cells of the neural crest. At 15.5 dpc it has been reported to be strongly expressed in the ventricular and subventricular zone of the cortex (Dahlstrand et al. 1995). To obtain neuronal-specific *Tra2b* knock-out animals *Tra2*$^{fl/fl}$ animals were crossbred with *Tra2b*$^{fl/+}$*; Nestin-Cre*$^{tg/0}$ animals. This crossbreeding allowed generation of control mice (*Tra2b*$^{fl/+}$ *and Tra2b*$^{fl/fl}$), heterozygous conditional knock-out mice (*Tra2b*$^{fl/+}$*; Nestin-Cre*$^{tg/0}$) and homozygous conditional knock-out mice (*Tra2b*$^{fl/fl}$*; Nestin-Cre*$^{tg/0}$) in a single litter (Figure 8 A).

4.2.3. Tagging and genotyping

Offspring of a breeding was weaned 3 weeks after birth and animals were separated by sex. To allow accurate animal management each animal was tagged with a numbered ear tag. For genotyping purposes a biopsy of the tail tip was taken (approx. 3 mm) according to the German laws for animal welfare (TierSchG §6). Oligonucleotides and genotyping conditions are listed in 3.9.2.

4.2.4. Timed breedings

Many applications like the generation of MEFs (4.1.3) or the detailed analysis of murine development require precise knowledge about the developmental stage of the analyzed embryo. For this purpose controlled timed matings were performed. Two non-pregnant females were routinely mated with one male over night. The following morning females were surveyed for the presence of a vaginal plug indicative of copulation during the night. Vaginal plugs are remnants of the male ejaculate that can be present up to 24 h after mating. Positively tested females were isolated and were considered to carry embryos at the stage of 0.5 dpc assuming fecundation in the middle of the night.

4.2.5. Isolation of prenatal mouse embryos

Mouse embryos were prepared following a time breeding (4.2.4). The pregnant mother was sacrificed at the desired developmental stage by cervical dislocation. The uterus harboring embryos was fully transferred to PBS. Embryos were prepared from the uterus and a tail biopsy was performed for genotyping purposes. Isolated embryos were either fixated in PFA (4.7.1) or embryonic brains were isolated and frozen in liquid nitrogen for subsequent RNA (4.4.2.3) analysis.

4.3. Microbiology Methods

All procedures involving cultivation of bacteria were performed under sterile conditions. Unless indicated otherwise bacteria were always cultivated under selective pressure depending on the plasmid propagated. In this study the *endA* deficient *E. coli* strains *TOP10* and *DH5α* were used to amplify the plasmid of interest. Both were obtained as chemically competent stocks from *Life Technologies*.

4.3.1. Heat shock transformation of chemically competent bacteria

Chemically competent cells are sensitive to temperature as to their ability to take up plasmid DNA. Cells were stored at -80°C and were thawed on ice directly before usage. Routinely, 2 µl of ligation product or approx. 100 ng of plasmid DNA were added to the bacterial vial and mixed gently by flicking the tube. Bacteria were incubated on ice for 20 minutes and then heat shocked for 30 seconds at 42°C in a water bath. Bacteria were directly chilled on ice for 2 minutes and 250 µl of pre-warmed S.O.C medium was added. The mix was incubated for 1 h at 37°C in a shaking incubator to allow expression of resistance genes. Various amounts (e.g. 20 µl and 200 µl) were plated on selective agar and incubated at 37°C over night.

4.3.2. Selection of positive clones

Colonies from a selective agar plate were picked using a sterile pipette tip. Routinely, the bacterial material obtained from one colony was resuspended in 20 µl of sterile H_2O. 10 µl of this water was used as a template in a colony PCR. The colony PCR is a standard PCR (4.4.3.1) that is supposed to detect the presence of the desired insert cloned into a vector using specific primers that are optimally positioned in the vector and the insert sequence. Screening of PCR products via agarose gel-electrophoresis allowed selection of clones with the supposedly correct insert in correct orientation. The remaining bacteria-containing water of correctly tested clones was used to inoculate 5 ml selective LB-cultures to generate bigger amounts of plasmid for further analysis.

4.3.3. Preparation of bacterial glycerol stocks

Generation of glycerol stocks from bacterial cultures allows long term storage of plasmids within bacteria without the need of re-transforming and screening of clones. Glycerol stocks were prepared by growing bacteria in a 5 ml culture of selective media. Cultivation was stopped in the exponential phase of growth (i.e. the culture medium becomes slightly opaque). 750 µl of bacterial culture were mixed with 250 µl of sterile glycerol as a

cryoprotective agent. Cryo vials were stored at -80°C. Bacterial cultures were inoculated from a frozen glycerol stock by transferring a minor piece of ice into selective culture media.

4.4. Molecular Biology Methods

4.4.1. Isolation of DNA

In the following chapter different methods of DNA purification will be explained. All methods vary in the obtained purity of DNA and the method of DNA isolation was chosen according to the downstream application. All isolated DNAs were routinely stored at -20°C.

4.4.1.1. Isolation of genomic DNA from murine tail biopsies

A single tail biopsy sample was lysed in 500 µl tissue lysis buffer (3.5) complemented with 200 µg/ml *Proteinase K* at 55°C in a shaking heating block at approx. 500 rpm. Hair and insoluble debris were removed by 15 minute centrifugation at 16,100 x g. The supernatant was mixed with 1 volume isopropanol in a new reaction tube to precipitate the DNA. Precipitation was continued for 30 minutes at room temperature. DNA was pelleted for 30 minutes at 16,100 x g and washed with 1 volume of EtOH. After repeated centrifugation the EtOH was discarded and the DNA was dried in a 37°C heated cabinet or in the *Concentrator* device. At last the pellet was resuspended in 80 µl TE^{-4} (3.5) containing 50 µg/ml *RNase A* to degrade possible RNA contamination.

4.4.1.2. Isolation of genomic DNA from tissues and cells

To obtain DNA of good purity from cells and animal tissues the *gDNA Blood & Tissue kit* (*Qiagen*) was used according to the manufacturer's instructions. For routine use, genomic DNA from pelleted cells and small amounts of tissue was isolated as previously described for tail biopsies (4.4.1.1). Depending on the tissue size the lysis step was increased in length. To improve DNA purity the lysate was optionally subjected to phenol-chloroform extraction (4.4.1.3).

4.4.1.3. Isolation of genomic DNA from tissues (Phenol-Chloroform)

Following tissue lysis in lysis buffer with *proteinase K* cellular debris was pelleted for 20 minutes at 16,100 x g. To minimize material loss during phenol-chloroform extraction a gel-based phase separating system was employed. *Phase Lock Gel Heavy* (5 Prime) tubes were prepared by spinning 1 minute at 16,100 x g. The cleared lysate was transferred to the *Phase Lock* tube. 1 volume phenol-chloroform was added (i.e. ½ vol. phenol + ½ vol.

chloroform/isoamylalcohol (24:1) and mixed by vortexing. The *Phase Lock* tube was rotated for 5 minutes and then spun for 5 minutes at 16,100 x g. The upper aqueous phase was transferred to a fresh *Phase Lock* tube and previous steps were repeated with 1 volume of phenol-chloroform. The procedure was finally repeated with 1 volume of pure chloroform/isoamylalcohol. After centrifugation the aqueous phase was mixed with 3 volumes of ice cold 100 % EtOH and 0.1 volumes of 3M NaAc (pH 5.2). DNA was precipitated over night at -20°C and afterwards pelleted for 20 minutes at 16,100 x g. DNA was washed in 70 % EtOH and repeatedly centrifuged. After drying in a heating cabinet the pellet was resuspended in TE^{-4} with 50 µg/ml RNase A.

4.4.1.4. Preparation of plasmid DNA from bacteria

Depending on the sensitivity and the needs of downstream applications a high purity and large amounts of plasmid DNA were required. For routine analysis of cloning products a small yield of plasmid DNA was sufficient. For this purpose the *Spin Miniprep Kit* (*Qiagen*) was used with a culture volume of max. 10 ml. Transfection experiments in eukaryotic cell culture require high amounts of plasmid DNA and a solid endotoxin removal. To obtain transfection grade plasmid DNA the *PureYield Plasmid Midiprep System* (*Promega*) was routinely employed with a culture volume of 200 ml. For highest amounts of transfection grade plasmid DNA the *Endo-free Plasmid Maxi Kit* (*Qiagen*) was used with culture volumes of up to 500 ml. Purified plasmid DNA was stored at -20°C.

4.4.1.5. Determination of DNA concentration via spectrophotometry

Concentrations of DNA solutions were determined using a *NanoDrop ND-1000 spectrophotometer* (*Peqlab*). From the measured absorption at the wavelengths of 260 nm and 280 nm the software calculated the DNA concentration and determined the $^{260nm}/_{280nm}$ ratio, which is an indicator for DNA purity. A value of 1.8 is indicative of high purity. Lower values point to possible remnants of ethanol, while higher values indicate the presence of RNA.

4.4.2. Isolation of RNA

RNA was handled under RNase-free conditions including appropriately treated reaction tubes and stuffed pipette tips. RNA in aqueous solution was always kept on ice and was stored at -80°C.

4.4.2.1. Isolation of RNA from eukaryotic cells

RNA was isolated using the *RNeasy Mini Kit* (*Qiagen*) according to the manufacturer's instructions. Depending on the amount of RNA needed, RNA was isolated from 6-well or 6 cm

tissue culture plates. The culture medium was removed from cells and cells were directly lysed in *RLT buffer* (*Qiagen*) and collected using a cell scraper. Cell lysates were homogenized using *QiaShredder* (*Qiagen*). The on-column *RNase-free DNAse* digestion was routinely performed in all RNA isolations to prevent contamination with genomic DNA interfering with downstream applications. RNA was eluted in RNase-free ddH_2O.

4.4.2.2. Isolation of RNA from tissues

Isolation of RNA from small tissue fragments or from tissues with low fat content was equally performed using the *RNeasy Mini Kit* (*Qiagen*) and *QiaShredder* (*Qiagen*). To ensure optimal breakup of organ tissue, specimens were homogenized using a rotor-stator homogenizer (*Ultra Turrax, IKA*) before proceeding with the RNA extraction as described in 4.4.2.1.

4.4.2.3. Isolation of RNA from fatty tissues

Isolation of RNA from tissues with high fat content (e.g. brain) requires additional phenol-chloroform extraction to reduce fat and protein content and to maintain high RNA purity. 100 mg of fatty tissue were broken up and homogenized in at least 1 ml of *QIAzol Lysis Reagent* (*Qiagen*) using a rotor stator homogenizer (*Ultra Turrax, IKA*). The suspension was repeatedly homogenized using *QiaShredder* (*Qiagen*) and incubated for 5 minutes at room temperature to allow dissociation of nucleoprotein complexes. To minimize material loss during phenol-chloroform extraction a gel-based phase separating system was employed. *Phase Lock Gel Heavy* (5 Prime) tubes were prepared by spinning 1 minute at 16,100 x g. Up to 750 µl of the homogenized lysate were transferred into the *Phase Lock* tube and mixed with 200 µl of chloroform/isomalyalcohol (24:1). Tubes were rotated at room temperature for 5 minutes and centrifuged for 10 minutes at 16,100 x g. If the lysate volume exceeded 750 µl the process was repeated in a fresh *Phase Lock* tube. The upper aqueous phase containing RNA was transferred into a new tube and mixed with 1 volume of RNA-grade 70 % EtOH. The mixture was then applied to *RNeasy* (Qiagen) columns and RNA extraction was continued according to the manufacturer's instructions and as described in 4.4.2.1.

4.4.2.4. Determination of RNA concentration via spectrophotometry

Concentrations of RNA solutions were initially determined using a *NanoDrop ND-1000 spectrophotometer* (*Peqlab*). From the measured absorption at the wavelengths of 260 nm and 280 nm the software calculated the RNA concentration and determined the $^{260nm}/_{280nm}$ ratio, which is an indicator for RNA purity. A value of 2.0 is indicative of high purity. The

determination of RNA concentrations with a standard spectrophotometer is insufficiently precise to allow reliable quantification of transcripts via real-time PCR. For quantitative downstream applications RNA concentrations need to be determined more accurately (4.4.2.5).

4.4.2.5. Determination of RNA concentration (Ribo-Green method)

Accurate determination of RNA concentrations was performed using the *Quant-IT RiboGreen RNA assay* (*Life Technologies*) as described in the manufacturer's instructions. This technique is based on the *RiboGreen* dye, which increases its fluorescence when bound to RNA. This allows RNA quantification that is unbiased by compounds absorbing at similar wavelengths. The assay was carried out in dark 96-well plates (*#655076, Greiner*) on a *Safire²* microplate reader (*Tecan*). *RiboGreen* allows absolute RNA quantification using an included RNA standard series ranging from 0 to 100 ng/µl. Therefore, the measured samples were pre-diluted to match the standard range. By fluorescence measurement of the standard series a standard curve was established by linear regression. The linear equation was used to determine absolute concentrations from fluorescence values obtained from samples.

4.4.2.6. Evaluation of RNA integrity (Experion gel-electrophoresis system)

For highly sensitive applications (e.g. exon array analysis) the purity and integrity of RNA samples was further analyzed on the *Experion* gel-electrophoresis system (*BioRAD*) using the *Experion RNA StdSens Kit* (*BioRAD*) according to the manufacturer's instructions. This system analyzes RNA integrity by assessing band sizes of the 18S and 28S ribosomal RNAs. These molecules should appear at distinct sizes of 1874 nt and 4718 nt, respectively. Gross deviation from these values or smear is indicative of impaired RNA integrity. Degraded samples or samples with impaired integrity have been excluded from exon array analysis or have been replaced by intact RNA samples.

4.4.3. Polymerase chain reaction

The polymerase chain reaction (PCR) was developed by Kary Banks Mullis in 1983 (Mullis et al. 1992). PCR utilizes thermostable DNA polymerases to extend 3'-OH ends of DNA oligonucleotides specifically binding to a DNA matrix molecule. Primer extension leads to cyclic amplification of the template, as each extension product serves as a new template in the subsequent cycle.

In this study PCR was a fundamental tool that many applications were based on. PCR was used for the genotyping of mice, for the generation of protein encoding sequences during cloning and for the quantification of transcripts from single stranded cDNA.

4.4.3.1. The standard PCR

For routine PCR, as applied for genotyping purposes but also for semi-quantitative PCR from a single stranded cDNA, the *Recombinant Taq DNA Polymerase kit* (*Life Technologies*) was used. A typical PCR setup on a 25 µl scale is shown in Table 1.

Table 1 | Typical PCR setup.
PCRs using the *recombinant Taq DNA polymerase* (*Life Technologies*) were routinely performed in a 25 µl scale. 10x PCR buffer and a 50 mM $MgCl_2$ solution were provided with the kit.

component	amount
H_2O (ad 25 µl)	14.35 µl
PCR buffer (10x)	2.5 µl
$MgCl_2$ (50 mM)	0.75 µl (1.5 mM)
dNTP Mix	4 µl
fw oligonucleotide (10 µM)	1 µl
rev oligonucleotide (10 µM)	1 µl
Taq DNA polymerase	0.4 µl
Template (20-200 ng)	1 µl

A starting point to assess performance of a PCR included $MgCl_2$ concentrations of 1.5 mM and primer concentrations of 0.4 µM. Amplification of targets that were difficult to amplify or that yielded unspecific products was enhanced by adaption of PCR conditions. Increasing the $MgCl_2$ concentration boosts polymerase performance on cost of primer binding specificity. Decreasing the $MgCl_2$ concentration can make primer binding very specific, but the amplification reaction might become very ineffective. Similar effects can be achieved by adjustment of the annealing temperature. Low temperatures cause primers to effectively bind the template but allow unspecific primer binding. High temperatures increase binding specificity but might also reduce the number of primers binding the template, thus decreasing PCR efficiency. The annealing temperature of a PCR program should be adjusted according to the oligonucleotide properties. Melting temperatures of oligonucleotides were calculated using the *OligoCalc* web application as listed in 3.11.2. A typical PCR program is shown in Table 2.

Normal Taq DNA polymerase is prone to insertion of wrong nucleotides and usually causes insertion of mutations into the amplified sequence. For that reason, proof reading polymerases were used for cloning procedures to preserve the original sequence. Proof reading polymerases used were *AccuPrime Polymerase* and *Platinum Taq Polymerase* (*Life Technologies*) as stated in 3.4.2. These polymerases vary from *recombinant Taq DNA polymerase* as they use different buffer systems and work best at 68°C. Proof reading PCRs were performed as suggested by the manufacturer.

Table 2 | Standard PCR program.
The usual PCR begins with an initial denaturation step melting the double stranded DNA template followed by 38 cycles of denaturation, annealing and extension. Finally, incomplete product are complemented for 10 minutes at 72°C. The annealing temperature varies between PCRs and depends on the oligonucleotide properties. The extension temperature is chosen according to the enzymes optimum temperature (i.e. 72°C for Taq polymerase) which may vary between enzymes). Extension time depends on the amplicon length and processing speed of the polymerase.

initial denaturation	5 min	95°C	
denaturation	30 sec	95°C	38 cycles
anneling	30 sec	50-68°C	
extension	~1min/kb	72°C	
final extension	10min	72°C	

4.4.3.2. Colony PCR

Colony PCR was used to identify bacterial clones that have successfully taken up a vector during transformation. Furthermore, the orientation of an insert within the vector was analyzed. Colony PCR is based on a standard PCR setup as described in Table 1 and Table 2. Bacteria of a colony were resolved in 20 µl of sterile water as described in 4.3.2. 10 µl of this solution served as a template for colony PCR. As template DNA was enclosed within intact bacteria the initial denaturation step of the PCR was increased to 10 minutes to allow effective lysis of the bacteria. Primers used for colony PCR were located on the vector sequence and within the insert. This allowed not only detecting the presence but also the orientation of the insert.

4.4.3.3. Clean-up of PCR products

Applications like ligations, digests by endonucleases and sequencing reactions require removal of PCR buffer, enzymes and free nucleotides from PCR products. For PCR clean-up the *QIAquick PCR Purification Kit* (*Qiagen*) was used according to the manufacturer's instructions. PCR products were eluted in ddH_2O.

4.4.3.4. Design of oligonucleotides

The design of oligonucleotides can significantly determine PCR specificity and efficiency. There are some guidelines for oligonucleotide design that were roughly followed to design primers for this study. Oligonucleotides used in PCRs are usually 18-30 nucleotides long and should have a GC content between 30 % and 70 %. Primers should end on at least 1 but at maximum 3 consecutive G or C nucleotides at the 3' end. Forward and reverse primers should have similar melting temperatures to be compatible in a PCR. Both primers should not be complementary with themselves or with each other to prevent formation of hairpins or primer dimers. Many of these molecular features were analyzed during primer design using the *OligoCalc* web application (3.11.2). To minimize the number of unspecific but as to their sequence similar targets *Primer BLAST* was performed (3.11.2). All oligonucleotides used in this study, their sequences and optimal PCR conditions are listed in 3.9. Oligonucleotides used in this study were produced by *Metabion* and *Integrated DNA Technologies* (*IDT*).

4.4.3.5. Reverse Transcription

cDNA synthesis was routinely performed using *SuperScript VILO* (*Life Technologies*) or the *QuantiTect Reverse Transcription kit* (*Qiagen*) according to the manufacturers instructions. For quantitative real-time PCR the *QuantiTect kit* (*Qiagen*) was used exclusively. Prior to cDNA synthesis the RNA concentration was determined by the *RiboGreen* method (4.4.2.5). For quantitative real-time PCR 150 ng of RNA was reversely transcribed in a 10 µl setup. For generation of a standard series 600 ng of RNA was used in 20 µl.

4.4.3.6. Semi-quantitative Reverse Transcription PCR

Semi-quantitative Reverse-Transcription PCR uses single-stranded cDNA as a template. After amplification of a specific cDNA copy of a transcript, original transcript amounts can be estimated by measuring band intensities on an agarose gel. This method is an end-point assay and evaluates amounts of PCR products at the end of a PCR. Reaction kinetics in every PCR include an exponential amplification phase followed by a saturation phase in which product amounts are no longer increasing. Therefore, in semi-quantitative RT-PCRs the cycle number has to be reduced as far as possible to allow valid densitometric band analysis, while the reaction was still performing exponentially. Running semi-quantitative PCRs for too many cycles will mostly result in equal band intensities, as all PCRs reached saturation, even though original cDNA levels might have been different. To overcome inaccuracies in quantification based on unequal loading of PCR template or unequal loading of the gel, all semi-quantitative PCRs were performed as multiplex PCR with a housekeeping gene.

Semi-quantitative RT-PCR was performed using *Recombinant Taq DNA Polymerase* (*Life Technologies*). PCR conditions of all applied PCRs were thoroughly optimized and the cycle numbers were reduced until bands became clearly weak but were still measurable. Densitometric analysis was performed using the *ChemiDoc XRS* system (*BioRAD*) in combination with the *QuantityOne* Software (*BioRAD*) or *ImageJ* (3.11.1).

4.4.3.7. Quantitative real-time PCR

Quantitative real-time PCR (qPCR) overcomes the limitations of semi-quantitative RT-PCRs. Even though it underlies the same reaction kinetics as any other PCR, template amounts can be robustly determined. In contrast to the semi-quantitative PCR, qPCR measures product amplification as it happens during exponential phase. In this study qPCRs were performed on the *Roche Light Cycler 1.5* using the *Light Cycler Fast Start DNA Master SYBR green I kit* (*Roche*) according to the manufacturer's instructions. The *SYBRgreen* I dye preferentially binds to DNA but also to RNA. Its absorption maximum is at 494 nm (blue). When excited *SYBRgreen* emits green light (521 nm). The fluorescence is strongest when bound to double stranded DNA but is much weaker when bound to single stranded DNA or RNA. After each primer extension step of a qPCR, *SYBRgreen* fluorescence is measured at 80°C. At this temperature double stranded DNA caused by primer dimmers or cDNA hairpins is denatured but longer PCR products are stable allowing a precise determination of PCR product amounts in every PCR cycle. As *SYBRgreen* binds double stranded DNA independently of the sequence, all PCR products including unspecific side products are measured. Therefore, specific product amplification is vital for *SYBRgreen*-based systems. Product specificity was surveyed by dissociation curve analysis and subsequent agarose gel electrophoresis. All qPCR primers used in this study amplify their target specifically at the indicated conditions (3.9.3).
For use as a template for qPCR, cDNA was produced as described in 4.4.3.5. The template solution was generated by diluting cDNA 1:5 in TE^{-4} buffer (e.g 8 µl of cDNA (from 150 ng RNA in 10 µl) + 32 µl of TE^{-4}). A standard series was generated from a 600 ng RNA reverse transcription reaction as shown in Table 3. The cycle threshold of each standard concentration was plotted against the concentraton (log) and the standard curve was extrapolated by linear regression. This curve was then reversely used to calculate the concentrations of samples from their respective cycle threshold value. Samples were routinely measured as duplicates and averaged using the arithmetic mean. For a 10 µl qPCR setup 3 µl of template solution were routinely used. The composition of a typical qPCR is shown in Table 4.

Table 3 | Dilution of a cDNA standard series for qPCR.
cDNA was diluted 1:2.5 in TE^{-4} buffer to generate templates with different concentrations ranging from 64 (i.e. 64/64) to 1 (i.e. 1/64 of the starting cDNA concentration)

Standard sample	cDNA	diluent
standard 64 (1:2.5)	16 µl cDNA (from 600 ng RNA in 20 µl)	24 µl TE^{-4}
standard 32 (1:5)	20 µl standard 64	20 µl TE^{-4}
standard 16 (1:10)	20 µl standard 32	20 µl TE^{-4}
standard 8 (1:20)	20 µl standard 16	20 µl TE^{-4}
standard 4 (1:40)	20 µl standard 8	20 µl TE^{-4}
standard 2 (1:80)	20 µl standard 4	20 µl TE^{-4}
standard 1 (1:160)	20 µl standard 2	20 µl TE^{-4}

Table 4 | Typical qPCR setup for the *Roche LightCycler 1.5*.
qPCRs in the *Roche Light Cycler 1.5* using the *SYBRgreen I Master Mix* were routinely performed in a 10 µl scale. H_2O, $MgCl_2$ solution, *SYBRgreen Master Mix* and *Fast Start* enzyme were provided in the kit. A $MgCl_2$ concentration of 3 mM was suitable for most applications but was decreased in some reactions to improve primer specificity.

component	amount
H_2O (ad 10 µl)	2.8 µl
$MgCl_2$ (25 mM)	1.2 µl (3 mM)
fw oligonucleotide (10 µM)	1 µl
rev oligonucleotide (10 µM)	1 µl
SYBRgreen Master Mix	0.833 µl
LC Fast Start Enzyme	0.166 µl
diluted cDNA template	3 µl

4.4.4. Agarose gel electrophoresis

DNA fragments of different sizes were separated by agarose gel-electrophoresis for analytical and preparative purposes. Different concentrations of agarose (e.g. 0.5 % - 3 %) were used according to the size of DNA fragments to analyze. For agarose concentrations higher than 1.5 % an agarose mix consisting of $^1/_3$ standard agarose and $^2/_3$ low melting agarose was used. For preparative applications including in-gel ligation of DNA 0.8 % of pure low melting agarose was used. Before loading, the gel was chilled to 4°C to become fully solid. Low melting gels were run at low voltage (e.g. 60 Volts) to prevent melting. For preparation agarose was melted in TBE buffer and supplemented with ethidium bromide (1:10,000). This allowed visualization of DNA when exposed to UV-light. DNA samples were mixed with 6x loading buffer, transferred to the gel and separated at 60 Volts to 180 Volts depending on the

gel size and agarose concentration. Results were documented using the *ChemiDoc XRS* documentation system (*BioRAD*).

4.4.5. Cloning Procedures

4.4.5.1. TOPO TA cloning

Cloning of DNA fragments into the eukaryotic expression vectors *pcDNA3.1* was performed using the *pcDNA3.1 TOPO TA Expression Kits* (*Life Technologies*). Different versions of the *pcDNA3.1* vector allow expression of fusion proteins with either GFP or V5-His tag. Vectors used for cloning are listed in 3.12. The *TOPO TA* ligation method uses 3'-adenine overhangs produced by *Taq polymerase* during PCR or previous 3'-adenylation for the ligation of a PCR product into the *pcDNA* vector. Topoisomerase coupled to the linearized vector allows effective but non-directional ligation of insert and vector. Ligation and transformation procedures were performed according to the manufacturer's instructions. Also the ligation of in-gel DNA fragments obtained from low melting agarose gels (4.4.4 and 4.4.5.4) was successfully performed using *TOPO TA* vectors.

4.4.5.2. Restriction Digest of DNA

Digestion of DNA with endonucleases was performed using restrictions enzymes and buffer systems from *New England Biolabs* (3.4.2). Digests were performed as a preparative step for directional ligation of DNA fragments into a vector or for analytical purposes. Nucleic acids used for digestion were either purified plasmids in ddH_2O (4.4.1.4) or purified PCR products in ddH_2O (4.4.3.3). Double digests were performed depending on buffer compatibility.

4.4.5.3. Extraction of DNA from agarose gels

DNA was routinely extracted from agarose gels using the *QIAquick Gel Extraction kit* (*Qiagen*) according to the manufacturer's instructions. DNA was eluted in ddH_2O.

4.4.5.4. DNA ligation and in-gel ligation

Ligation procedures were performed using *T4 DNA ligase* (*Promega*) as suggested by the manufacturer. Ligation was performed in a 20 µl reaction scale when using purified vector and insert sources. For ligation of in-gel vector and insert fragments the volume was increased to 50 µl to compensate for changes in the buffer system caused by the presence of TBE from the gel buffer. Digested PCR products and plasmids were resolved on a 0.8 % low melting agarose gel as described in 4.4.4. Bands were excised using a sharp scalpel while minimizing

the amount of total agarose cut. Agarose containing DNA was transferred to reaction tubes and heated to 70°C for 5 minutes to allow homogenization of the DNA agarose mixture. Depending on band intensities observed on the gel a total of up to 30 µl of agarose solution (including insert and vector) was used in a 50 µl ligation setup. Ligation was either performed at 14°C over night or for 2 hours at room temperature. 2 µl of a completed ligation reaction were used for heat-shock transformation of chemically competent bacteria as described in 4.3.1.

4.4.6. DNA sequencing

Sequencing of purified plasmid DNA and purified PCR products was performed at *GATC Biotech*. PCR products were routinely sequenced using custom made primers. Inserts of vectors were sequenced using custom primers and vector-primers provided by *GATC*. Sequence alignment and analysis was performed using the SeqMan application from the Lasergene Suite (DNAstar).

4.5. High throughput screening for alternatively spliced exons

Aiming on the discovery of alternatively spliced exons regulated by *Tra2b* in the murine developing central nervous system, two high throughput screening approaches have been performed. In both approaches relative abundances of single exons were compared to the abundance of the respective whole transcript to allow detection of differences in exon inclusion between control (*Tra2b*$^{fl/+}$) and conditional knock-out aimals (*Tra2b*$^{fl/fl}$*; Nestin-Cre*$^{tg/0}$).

4.5.1. Whole transcriptome sequencing

Whole transcriptome sequencing was performed by the Cologne Center for Genomics (CCG) using a RNA-sequencing approach from reversely transcribed cDNA. mRNA was isolated by poly-A selection. Sequencing was performed as a paired-end sequencing run with 2 x 95 bp of read length. Sequence data was aligned to the murine transcript database at *Ensembl*. All bioinformatic analyses and data processing was performed by Dr. Peter Frommolt (CCG). Exon and transcript abundances were estimated using a RPKM (***r**eads **p**er **k**ilobase per **m**illion mapped reads*) based model. Statistically, the number of reads mapping to an exon reflects the relative abundance of that exon. To exclude effects of differential gene expression, exon-specific RPKM-values were normalized to the RPKM-value of the respective whole transcript. In total 3 control animals have been compared to 3 conditional knock-out animals using whole brain RNA. Normalized RPKM-values for each exon were grouped according to the genotype and statistically compared. Candidate exons were sorted according to differences in RPKM, which suggested a differential usage of that exon in control and knock-out animals. Candidate

exons obtained in this approach were analyzed for validation of alternative splicing events by qPCR.

Many methodological details of the transcriptome data analysis and their variations had direct effect on data and outcome. Therefore, numerous items are described in detail in chapter 5.5 of the result section.

4.5.2. Mouse exon arrays

Mouse exon arrays were performed in the laboratories of Prof. A. Sachinidis at the Institute of Neurophysiology at the University of Cologne. Array analysis and preparative work was carried out by Dr. Kesavan Meganathan. The main workflow of mouse exon array preparation and labeling procedures is shown in Figure 7. All procedures have been carried out according to the manufacturer's instructions as described in the respective manuals of the *WT Expression Kit* (*Ambion*) and the *WT Terminal Labeling and Hybridization Kit* (*Affymetrix*). Whole brain RNA samples of *Tra2b*$^{fl/+}$ and *Tra2b*$^{fl/fl}$*; Nestin-Cre*$^{tg/o}$ mice were prepared as follows:

RNA samples were used as starting material and their quality and concentration was analyzed on the *Experion* gel-electrophoresis system (*BioRAD*) as described in 4.4.2.6. From mRNA first strand cDNA synthesis was performed using T7 promoter-coupled oligo(dT) primers and 100 ng of total RNA. The mRNA template was digested by RNAse H digest and the complementary DNA strand was synthesized using DNA-polymerase to produce ds cDNA. In an overnight *in vitro* transcription reaction antisense cRNA was produced using T7 RNA-polymerase. cRNA was purified using a magnetic bead system provided with the kit to remove enzymes, salt and free nucleotides. The total cRNA yield was determined using the Nanodrop-1000 spectrophotometer as described in 4.4.1.5. More than 20 µg of cRNA should have been produced from 100 ng RNA. 10 µg cRNA was used to generate single sense strand cDNA using random primers and a nucleotide mix containing dUTP. The cRNA template was digested by RNAse H and sense strand cDNA was purified using the provided magnetic bead system. The cDNA yield was determined using the Nanodrop-1000 spectrophotometer (4.4.1.5). More than 6 µg of cDNA should have been produced from 10 µg cRNA. Single stranded sense cDNA was fragmented using uracil-DNA-glycosylase (UDG) and apurinic/apyrimidinic endonuclease (APE1) cleaving DNA at positions of incorporated dUTP. cDNA ends were labeled using terminal deoxynucleotidyl transferase (TdT) with the DNA labeling reagent provided with the WT labeling kit. Labeled cDNA was hybridized to the array for 16 h and subsequently processed using an *Affymetrix* fluidics station for washing, staining and repeated washing, while using

components provided with the Hybridization Wash and Stain kit (*Affymetrix*). Arrays were scanned using an Affynetrix GeneChip scanner.

Bioinformatic analyes were carried out by Dr. John Antonydas Gaspar. In short: Both gene level and exon level analysis were performed using *R*, *Bioconductor package* and *Partek Genomic Suite* tools. Preprocessing of exon array data included background correction using RMA (Bolstad et al. 2003), adjusting for GC-content, Quantile normalization and probeset summarization using mean and setting log probes using base 2. The probeset class, Gene extended was used. Preprocessing was performed by *Partek Genomic Suite*. Gene level analysis was performed in core extended mode. Transcripts showing very low expression signal values of less than 25 across arrays were eliminated and only 23,331 transcripts (annotated and non-annotated) were taken for statistical analysis. Moderated t statistics of linear model analysis (Smyth 2004) was employed to filter the differentially expressed transcripts. Threshold values of t score-p value of ≤0.05 and fold-change value 1.5 were used to generate a significant gene list. *Partek Genomic Suite* was used for statistical filtering of the alternatively spliced variants. Prior to alternative splicing calculation the probe sets with signals <3 (at log 2 scale) were excluded. The threshold values used for statistical filtering were set to p-value ≤0.05 and fold-change values of 1.5 for high confidence (personal communication with Dr. Gaspar).

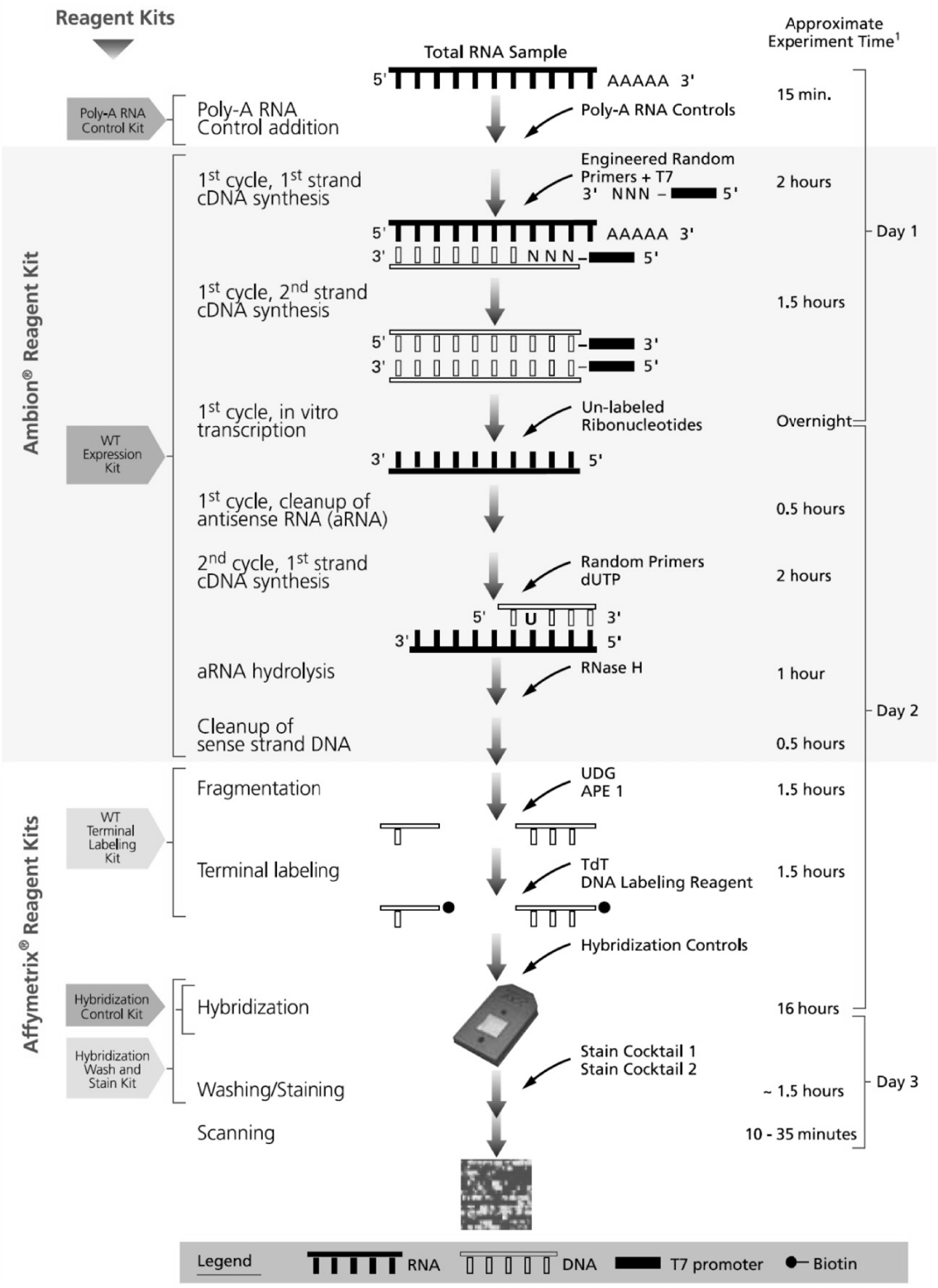

Figure 7 | Whole transcript sense target labeling workflow.
The *Ambion® WT Expression Kit* (*Ambion*) was used to amplify and generate sense strand cDNA with unbiased coverage of the transcriptome. cDNA was labeled and prepared using the *GeneChip® WT Terminal Labeling and Hybridization Kit* (*Affymetrix*). The array used was *GeneChip® Mouse Exon 1.0 ST Array* (*Affymetrix*). The figure is based on the Affymetrix product information.

4.6. Proteinbiochemistry

4.6.1. Extraction of proteins

Proteins were generally isolated in RIPA buffer (*Sigma-Aldrich*) containing a mix of protease inhibitors (*Complete Mini Protease Inhibitors, Roche*). 1 tablet of protease inhibitors was resolved in 10 ml of RIPA buffer. All isolation steps were carried out at 4°C and proteins were routinely stored at -80°C.

4.6.1.1. Extraction of proteins from eukaryotic cells

Cells were detached from growth surface by trypsinization and pelleted as described in 4.1.1. Cells were washed in PBS without Ca^{2+} and Mg^{2+} and transferred to new reaction tube. Depending on cell type and cell number the pellet was resuspended in an appropriate amount of ice cold RIPA buffer (e.g. 40 µl RIPA for a confluent 6-well of HEK293T or NSC34 cells). For efficient lysis cells were incubated on ice for 30 minutes. Lysate was centrifuged for 30 minutes at 16,100 x g and 4°C to sediment cellular debris. The supernatant containing soluble proteins was transferred to a new pre-chilled reaction tube. Isolated proteins were snap-frozen in liquid nitrogen and stored at -80°C.

4.6.1.2. Extraction of proteins from tissues

Lysis and protein extraction from animal tissues requires additional steps for efficient tissue break up and lysis. To obtain protein from organs (e.g. brain, liver, lung, spleen or testis) the tissue was homogenized in RIPA buffer containing protease inhibitors using a rotor stator homogenizer (*Ultra Turrax, IKA*). The homogenizer was cleaned after each homogenization step in 1 % SDS, 70 % EtOH and ddH_2O. Depending on tissue size 200 µl to 400 µ of RIPA buffer was used. Tissues with high protein content (e.g. brain) generally required a larger lysis volume. For effective lysis the homogenate was incubated on ice for 30 minutes. Subsequent steps were performed as described in 4.6.1.1.

4.6.1.3. Determination of protein concentration via Bradford protein assay

Protein concentrations were determined using the protein assay according to Bradford (Compton and Jones 1985). The Bradford reagent contains the dye Coomassie-Brilliant-Blue G-250, which forms complexes with cationic and nonpolar amino acids in acidic environment. This stabilizes the dye in its non protonated form which causes a shift of light absorbance maximum from 470 nm to 595 nm. 1 µl of protein lysate was mixed with 499 µl of Bradford reagent (*Applichem*) and incubated at room temperature for 10 minutes. The concentration

was determined by measuring the absorption at 595 nm on a standard spectrophotometer (*Eppendorf*) using a standard of 0.5 µg to 5.0 µg of BSA for calibration.

4.6.2. Discontinuous SDS-PAGE

During SDS-PAGE (*sodium dodecyl sulfate polyacrylamide gel-electrophoresis*) proteins are resolved according to their molecular weight within an electric field. Proteins of different size or molecular weight proceed through the polyacrylamide gel with different velocities. Furthermore, the anionic detergent SDS is applied, which adapts the mass to charge ratio of proteins. Proteins become negatively charged and their speed in the gel is almost exclusively determined by their mass. SDS also denatures secondary and tertiary structures. Thus proteins of similar mass proceed towards the anode at same velocities without being impaired by their innate conformation. Hydrogen- and disulfide-bonds are severed due to boiling in presence of reducing compounds like dithiothreitol or β-Mercaptoethanol (Laemmli 1970).

In this study discontinuous SDS-PAGEs were carried out. Routinely, a 4 % stacking gel and a 12 % resolving gel was used to separate most target proteins. Discontinuous SDS-PAGE allows proteins to focus in a sharp band within the stacking gel and to simultaneously start resolving in the resolving gel. This effect is based on a discontinuous buffer system with low pH of 6.8 in the stacking gel and higher pH of 8.8 in the resolving gel. Proteins traveling through the stacking gel have the same charge and their speed is hardly biased by their size, as the gel matrix is insufficiently dense (4 %). In the stacking gel proteins are focused between two ion fronts: The quickly moving Cl^- and the slowly moving glycine, which has its isoelectric point *pI* at 5.97 and therefore close to pH 6.8. As the ion front approaches the resolving gel (pH 8.8) glycin overtakes the proteins, which are then no longer focused and resolve according to their size caused by interaction effects with the dense gel matrix.

SDS gels were prepared between two vertical glass plates. First, the resolving gel (3.5) was prepared and isopropanol was pipetted on top to produce a sharp edge. After polymerization was complete, isopropanol was removed and the stacking gel (3.5) was poured on top.

The total protein concentration of lysates was determined according to Bradford (4.6.1.3). 7.5 µg to 20 µg of total protein were supplemented with 1 volume of 2x SDS sample buffer and boiled for 5 minutes before being loaded to the gel. As a protein ladder 10 µl of the *PageRuler™ Prestained Protein Ladder* (*Thermo*) were loaded. Electrophoresis was performed at low initial voltage of 80 Volts while proteins were moving through the stacking gel. Voltage was increased up to 180 Volts while resolving.

4.6.3. Transfer of proteins to PVDF blotting membranes

Proteins that have been resolved by SDS-PAGE (4.6.2) were electrophoretically transferred to PVDF (*polyvinylidene fluoride*) blotting-membranes. All transfers have been performed via wet-blot (tank–blot) according to Towbin (Towbin et al. 1979). Protein transfer was carried out at 4°C for 2 hours at 100 Volts or over night at 30 Volts. Successful transfer of proteins could be recognized due to pre-stained protein ladders that were visible on the blotting membrane.

4.6.4. Immunostaining of proteins on PVDF membranes

The immunoblot or Western Blot utilizes antibodies to specifically detect proteins. Primary antibodies specifically bind epitopes of the target protein. The primary antibody is then detected by secondary antibodies, which are conjugated to HRP (*horseradish peroxidase*). The desired protein can finally be detected by chemiluminescence using *Super Signal West Pico ECL Substrate* (*Thermo*). This substrate contains H_2O_2 and luminol. As HRP cleaves H_2O_2, luminol will luminesce.

After blotting (4.6.3) the membrane was shortly washed in TBST and blocked with 6 % milk powder in TBST at 4°C for 2 h or overnight. Primary antibodies were applied in TBST with 2 % milk powder and incubated at 4°C. Concentration and the duration of incubation time for each antibody are stated in 3.8.1. The membrane was washed 5 times for 5 minutes in TBST to remove non-bound antibodies. The secondary antibody was applied in TBST with 2 % milk powder at concentrations indicated in 3.8.2. All secondary antibodies used for immunoblots were HRP-conjugates. To remove non-bound secondary antibody the membrane was washed 5 times for 5 minutes in TBST. To visualize the protein of interest the membrane was incubated for 5 minutes in chemiluminiscence substrate. Results were recorded using the *ChemiDoc XRS* system (*BioRAD*). Densitometric analysis was performed using the *QuantityOne* Software (*BioRAD*) or *ImageJ* (3.11.1).

4.6.5. Restoring of PVDF membranes for re-blotting

For repeated immunological detection of proteins on previously used PVDF membranes bound antibodies were removed by shaking the membrane for 7 to 15 minutes in *Restore Western Blot Stripping Buffer* (*Thermo*). Afterwards the membrane was washed twice in TBS before proceeding with a new immunostaining.

4.7. Histological Methods

In this study brains of conditional knock-out mice were histologically analyzed at various developmental stages. To obtain mice of the desired genotype at the desired stage timed breedings were performed as described in 4.2.4.

4.7.1. Histological preparation of animal tissues

Pregnant females were sacrificed by cervical dislocation and embryos were isolated as described in 4.2.5. Isolated embryos were washed in PBS and fixated in PBS with 4 % PFA at 4°C for 1 day or longer to prevent tissue decay. Embryos at developmental stages later than 15.5 dpc were skinned prior to fixation to allow efficient uptake of PFA. Preparation of the specimen for micro-dissection included fully automated dehydration and paraffin embedding using a tissue processor (*TP1028, Leica*). After micro-dissection, paraffin sections were flattened in a water bath and dried over night at 37°C.

4.7.2. Haematoxylin/Eosin staining of paraffin-embedded sections

HE staining (haematoxylin and eosin staining) was used as an initial standard staining procedure to allow gross histological analysis. Upon HE staining nuclei are stained dark blue while most parts of the cytoplasm (i.e. proteins) are stained red to pink.

Paraffin-embedded sections were first deparafinized for 2 x 15 minutes in xylene and rehydrated in a series of EtOH solutions of decreasing concentrations (each 2 min in 100 % EtOH, 96 % EtOH, and 70 % EtOH). Sections were shortly washed in PBS and H_2O. Staining with haematoxylin was performed for 6 minutes. Specimens were shortly rinsed in water and then washed for 15 minutes in fresh water. Finally, sections were shortly rinsed in ddH_2O. Eosin staining was performed for 1 minute before shortly washing in H_2O. At the end specimens were dehydrated with an EtOH series of increasing concentration (each 1min in 70 % EtOH, 96 % EtOH and 100 % EtOH). For sample preparation sections were incubated for 1 minute in xylene and mounted with *Eukitt* mounting medium.

4.7.3. Immunostaining of paraffin-embedded sections

Paraffin-embedded sections were first deparafinized in xylene and rehydrated in a series of EtOH solutions of decreasing concentrations (each 2 min in 100 % EtOH, 96 % EtOH, and 70 % EtOH) followed by 10 minutes incubation in ddH_2O. To improve epitope accessibility antigen retrieval was performed by boiling specimens 3 times for 5 minutes in citrate buffer in a microwave at 600 W. Specimens were left 45 minutes for cool down before washing for 5 minutes in TBS. Sections were blocked for 45 minutes at room temperature in a moist chamber (20 % goat serum, 1 % BSA in TBS) and afterwards washed 3 times for 5 minutes in TBS. The

primary antibody was applied in a moist chamber over night at 4°C in TBS with 3 % milk powder. Concentrations of primary antibodies for immunohistochemistry are stated in 3.8.1. To remove non-bound antibodies sections were washed 3 times in TBS. Biotinylated secondary antibodies were applied in a moist chamber for 30 minutes in TBS with 3 % milk powder and samples were repeatedly washed 3 times for 5 minutes in TBS. In a 30 minute reaction avidin-conjugated peroxidase was coupled to biotinylated antibodies using the *Vectastain ABC kit* (*Vector Laboratories*) according to the manufacturer's instructions. Specimens were washed 3 times for 5 minutes with TBS were then incubated with DAB (diaminobezidine) staining solution. Enzymatic turnover of H_2O_2 causes DAB to produce a brownish stain. After sufficiently strong staining was achieved, sections were rehydrated and mounted as previously described (4.7.2).

4.7.4. Microscopy

Parffin-embedded brain sections were microscopically analyzed. Images were recorded on a Zeiss AXIO microscope and the Zeiss M2 imaging system using the ZEN software (Zeiss, Jena, Germany). Routinely, whole mount overview images were taken using 1.25x magnification. For high resolution analysis multiple images have been taken to finally cover the whole sample using 10x magnification. Images were merged and blended using the Demo version of the Autostitch software. Overlap of image recording at the microscope allowed perfect panoramic assembly of single images. Images used for brain analysis that are shown in this work are high resolution panoramic assemblies consisting of multiple 10x magnification shots.

4.8. Statistical analyses

Data obtained in this study was generally undertaken to statistical evaluation to analyze the technical error of a measurement or to determine the significant difference between groups of biological replicates. All statistical analyses were carried out using *Excel 2007* (Microsoft). Some measurements were recorded as duplicates (e.g. qRT-PCR, Light Cycler 1.5, n=2) or triplicates (RiboGreen method of determining the RNA concentration, n=3) to allow calculation of the mean to minimize the technical error caused by device inaccuracies or inaccurate pipetting. To assess biologically significant differences, groups of biological replicates (e.g. mice of a specific genotype or equally transfected but individual cell cultures) were compared to each other. N gives the number of biological replicates; i.e. the number of individuals per group observed. The average of each group was calculated by the arithmetic mean, while the dispersion of readout values within a group was represented by the standard

error of the mean (s.e.m.). Error bars within bar charts represent the s.e.m. unless indicated otherwise. *Student's t-Test* was employed to determine the statistical significant difference between groups. The probabilities for data points belonging to indifferent groups are are 0.05 (*), 0.01 (**) or 0.001 (***).

The statistical analyses of raw data obtained in high throughput screening approaches are described in the sections 4.5.1 and 5.5.1 for whole transcriptome sequencing and in the sections 4.5.2 and 5.6.2 for mouse exon arrays. All candidate filtering strategies explained in these sections were performed using filtering and sorting functions of *Excel 2007* (Microsoft). Analyses of exon sequences for the detection of GAA-binding motifs was performed by Dr. Radu Wirth using custom made scripts.

5. Results

5.1. Generation and validation of neuronal-specific Tra2b knock-out mice

Tra2b is a potent modifier of *SMN2* exon 7 splicing, which makes it an extremely promising candidate to ameliorate the SMA phenotype (Hofmann et al. 2000). HDAC inhibitors have been demonstrated to effectively increase *SMN2* exon 7 inclusion via transcriptional upregulation of *Tra2b* ultimately resulting in increased amounts of SMN protein (Brichta et al. 2003; Riessland et al. 2006; Riessland et al. 2010). This makes HDAC inhibitors a potent tool to establish treatment strategies for SMA. However, splicing factors are not restricted to a single target, and increased expression levels of *Tra2b* might alter the splicing pattern of a multitude of transcripts. This study aims at the identification of physiological Tra2b target transcripts in a murine *in vivo* model system. Ubiquitous ablation of *Tra2b* in the mouse is incompatible with life and results in early embryonic lethality (Mende et al. 2010). In this study a neuronal-specific knock-out mouse was generated to allow the analysis of Tra2b-related splicing processes in neurons, which is the primary affected cell type in SMA.

5.1.1. Identification and initial characterization of neuronal-specific knock-out animals

The generation of neuronal-specific *Tra2b* knock-out mice was performed as described in 4.2.2. By crossbreeding of *Tra2b*$^{fl/fl}$ mice with *Tra2b*$^{fl/+}$*; Nestin-Cre*$^{tg/0}$ mice, control mice (CTRL, *Tra2b*$^{fl/+}$ *and Tra2b*$^{fl/fl}$), heterozygous neuronal-specific knock-out mice (HET, *Tra2b*$^{fl/+}$*; Nestin-Cre*$^{tg/0}$) and homozygous neuronal-specific knock-out mice (KO, *Tra2b*$^{fl/fl}$*; Nestin-Cre*$^{tg/0}$) were obtained in a single litter (Figure 8 A). Genotypes were expected to split up according to Mendelian law with each genotype contributing to approx. ¼ of the offspring.

Offspring of this breeding usually comprised few pups that were found dead in the cage in the morning after birth or have already been (partly) devoured by the mother. DNA was isolated from dead pups and genotyping analysis was performed using oligonucleotides as described in 3.9.2 (Mende et al. 2010). The genotype of the *Tra2b* allele and the presence of the *Nestin-Cre* transgene were determined in separate PCRs. Genotyping indicated that the dead animals were KO mice of the genotype *Tra2b*$^{fl/fl}$*; Nestin-Cre*$^{tg/0}$, indicating that these mice were either born dead or died shortly after birth.

To narrow down the time point of death, the mating was performed as a timed breeding. Embryos were isolated from pregnant females at late developmental stages (18.5 dpc or later). All isolated embryos were viable and did not show differences in head morphology or outer appearance (Figure 8 B). Genotyping indicated that those litters also included KO mice (Figure 8 B). Analysis of 113 prenatal pups at late developmental stages

showed that KO mice were viable during late embryonic development and contributed as expected roughly ¼ to the total offspring (Figure 8 C). Later observations showed that KO animals were born alive. In rare cases KO mice survived up to 36 h. CTRL and HET mice were fully viable at birth, developed normally and had a normal life expectancy.

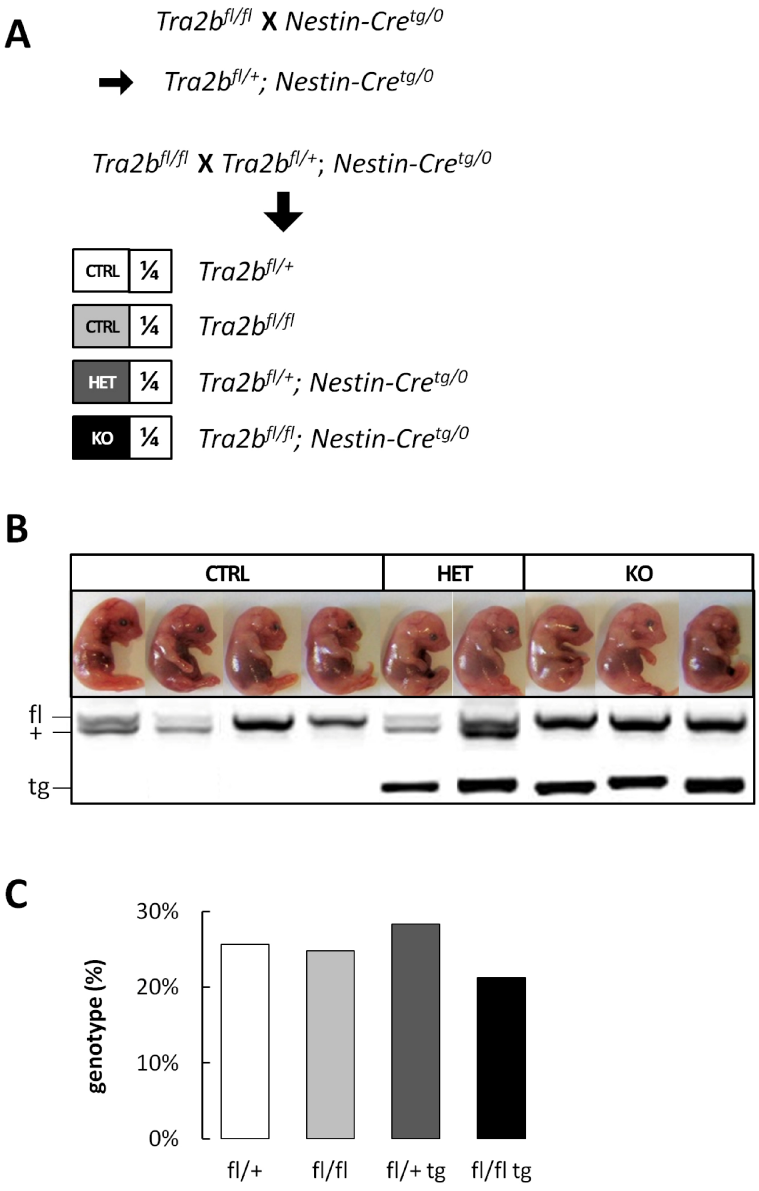

Figure 8 | Crossbreeding of Tra2b mouse lines and identification of knock-out animals.
(A) Breeding scheme. Cross breading of *Tra2b*$^{fl/fl}$ with *Tra2b*$^{fl/+}$*; Nestin-Cre*$^{tg/0}$ mice allowed generation of neuronal-specific knock-out (KO) animals as well as controls (CRTL) and heterozygous knock-out animals (HET) in one litter. (B) General development of conditional knock-out mice is not impaired as there are no gross morphological differences in embryo appearance. The genotype of the *Tra2b* allele (upper bands) was determined using primers 3482 and 2340 for detection of the floxed (fl) allele (950 bp) and the wild type (+) allele (821 bp). The *Nestin-Cre* transgene was detected using the primers 3478 and 3479 yielding a 500 bp product (lower bands). (C) Analysis of prenatal pups at late developmental stages (>18.5 dpc). KO animals were born alive and all possible genotypes were detected according to Mendelian law (N=113).

5.1.2. Analysis of Tra2b expression levels in neuronal-specific knock-out animals

To validate the knock-out of *Tra2b* in conditional knock-out mice the expression of *Tra2b* was analyzed on mRNA and on protein level. Prenatal embryos were isolated on day 16.5 of gestation. Embryonic whole brains were used to isolate total RNA and proteins. RNA was used for quantitative real-time PCR using specific primers to detect *Tra2b*. All data was normalized to *Gapdh* expression. Brains of KO mice showed highly significantly reduced *Tra2b* levels compared to all other genotypes (Figure 9 A). KO mice showed a more than 80 % reduction of *Tra2b* expression compared to *Tra2b*$^{fl/+}$ mice, but not a total loss of expression. This was expected, as *Cre-recombinase* under control of the *Nestin* promoter is not expressed in all cells of the brain. The remaining *Tra2b* expression likely originated from *Nestin*-negative cells. Brains of *Tra2b*$^{fl/fl}$ mice maintained high expression levels that are only slightly but significantly reduced compared to *Tra2b*$^{fl/+}$ mice. HET mice showed an approx. 40 % reduction

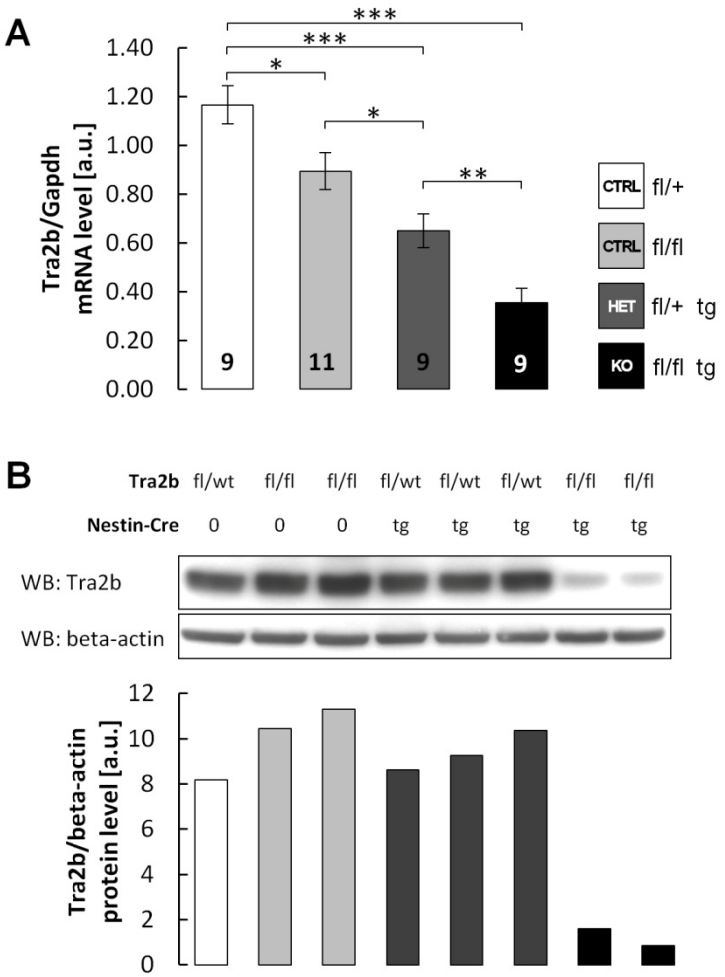

Figure 9 |Analysis of *Tra2b* expression levels in neuronal-specific knock-out mice.
(A) *Tra2b* expression on RNA level. *Tra2b* was detected using the specific primers 3218 and 3219 in a quantitative real-time PCR. Data was normalized to *Gapdh* (primers 3999 and 4000). Numbers within the bars indicate the number of biological replicates (N). (B) Tra2b expression on protein level. Tra2b (38 kDa) and beta-actin (42 kDa) were detected by specific antibodies. Bars show Tra2b expression normalized to beta-actin.

of *Tra2b* expression compared to *Tra2b*$^{fl/+}$ mice. This was significantly different to CTRL and KO mice.

Whole brain protein lysates were analyzed by Western Blotting. Immunostaining of Tra2b showed a clear reduction of protein levels in KO mice compared to controls (Figure 9 B). Both control genotypes (*Tra2b*$^{fl/+}$ and *Tra2b* $^{fl/fl}$) and HET mice showed similarly high expression levels of Tra2b. Despite *Tra2b* transcript levels in HET mice were reduced by half (Figure 9 A), protein levels were maintained high. These data indicate that despite the loss of one *Tra2b* gene copy in HET mice normal protein levels are maintained in the brain. This suggests that some regulatory mechanism might compensate lower transcript concentrations on a post-transcriptional level. Loss of both gene copies in KO animals resulted in a drastic reduction of *Tra2b* levels. Despite Cre-mediated recombination did not occur in *Nestin*-negative cells, the very low Tra2b expression in KO mice indicates an effective downregulation of Tra2b in the majority of the brain.

5.2. Generation of Tra2b deficient murine embryonic fibroblasts

Murine embryonic fibroblasts (MEFs) are an easy to generate cell culture model that finds application in a multitude of experimental strategies. This study aimed at the generation of a murine embryonic fibroblast cell line that is depleted of *Tra2b*. Such cell line would allow the study of Tra2b-related splicing processes in a homogeneous cellular system in a reproducible environment.

MEFs are typically generated from murine embryos at 13.5 days of gestation. The early lethality of mice ubiquitously depleted of *Tra2b* (Mende et al. 2010) precludes MEF generation from this mouse line. Here MEFs were generated from *Tra2b*$^{fl/fl}$ pups. Deletion of *Tra2b* was achieved by stably transfecting cells with *Cre-recombinase* including a nuclear localization signal (NLS-Cre). Transient transfections were unsuccessful generating *Tra2b*-depleted cell lines, as transfection efficiencies in fibroblasts were only low to moderate and cells depleted of *Tra2b* appeared to have a selective disadvantage which became apparent by decreased proliferation rates. Thus, *Tra2b*-depleted cells were quickly repressed by faster proliferating *Tra2b*$^{fl/fl}$ cells.

A stable transfection approach allowing for permanent selection for cells expressing *Cre-recombinase* using antibiotics to prevent survival of *Tra2b*$^{fl/fl}$ cells was followed. Therefore, cells were immortalized by stable transfection of the SV40 (simian vacuolating virus 40) large-T antigen (Zhu et al. 1991; Tevethia et al. 1998). Immortalization allows cells to evade senescence and to keep proliferating extensively, thereby evading limitations of murine

embryonic fibroblasts. The immortalization vector (pBABE-zeo-largeT) was obtained from *Addgene* and was originally provided by the lab of Robert Weinberg (Hahn et al. 2002b).

5.2.1. Generation and characterization of murine immortalized fibroblasts (MIF)

Primary murine embryonic fibroblasts were transfected with pBABE-zeo-largeT as described in 4.1.5. pBABE-zeo propagates resistance to Zeocin. To allow optimal selection of stably transfected cells, MEFs were previously treated with different concentrations of Zeocin to determine a low but effective concentration (Figure 10 A). After 7 days of Zeocin treatment cells were effectively suppressed by 200 µg/ml Zeocin. Lower concentrations failed to effectively suppress cell growth.

Transfection efficiency was surveyed by GFP-fluorescence using transfection of pMAX-GFP. Average transfection efficiency was usually lower than 50 %. Zeocin selection using 200 µg/ml was started 48 h after transfection of pBABE-zeo-largeT. While replenishing Zeocin in a 48 h interval cells were cultivated for 6 weeks. Expression of the largeT-antigen was analyzed by RT-PCR and transfected cells but not controls showed clear largeT expression.

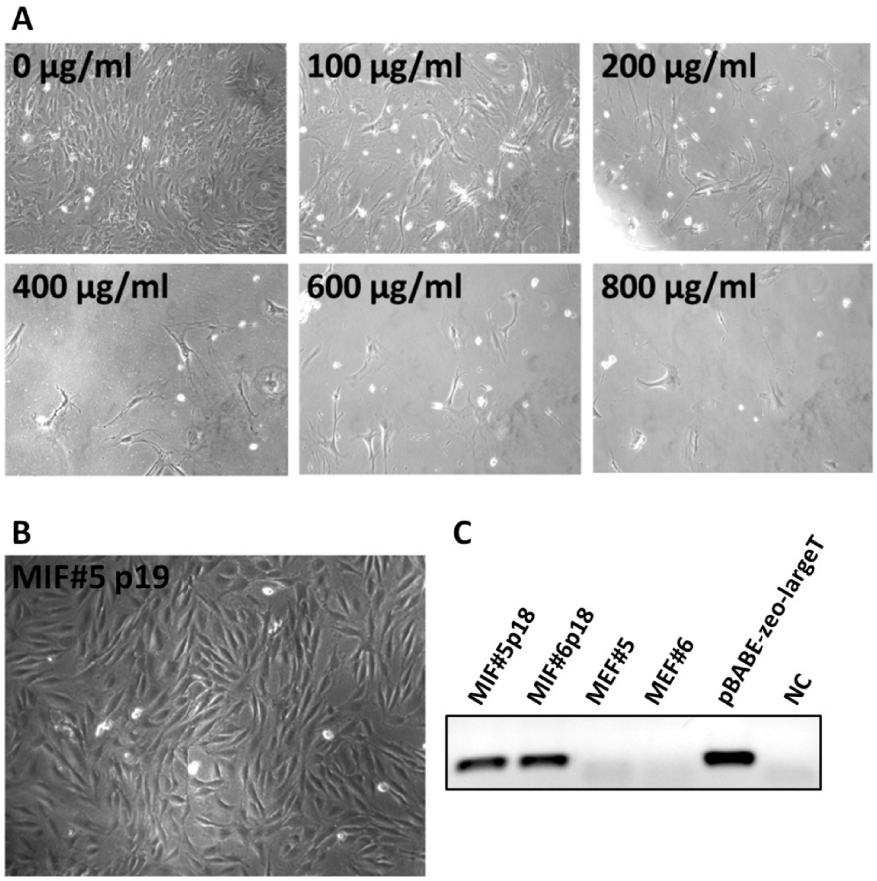

Figure 10 | Generation and characterization of murine immortalized fibroblasts (MIFs).
(A) Determination of the optimum Zeocin concentration for stable transfection. MEFs were treated for 7 days with different concentrations of Zeocin. (B) MIF cell morphology 19 passages after immortalization. (C) Detection of largeT antigen expression via RT-PCR. Primers 4052 and 4053 were used to specifically detect transcripts of the largeT antigen. MIF, murine immortalized firboblasts; MEF, murine embryonic fibroblast; NC, negative control

Zeocin selection was suspended and cells were cultivated for further passages. Immortalized murine firbroblasts (MIFs) showed an increased proliferation rate far superior to that of non-treated MEFs and tolerated splitting ratios of 1:40. MIFs acquired a more pronounced spindle-like shape that is usually not observable in MEFs (Figure 10 B). Furthermore, MIFs did not show any alterations in morphology or proliferation speed at high passage number (p>50). Non-treated MEFs may enter crisis early after 5 passages, but were sometimes observed to keep proliferating up to 20 passages. Testing expression of largeT antigen several passages after Zeocin withdrawal showed solid largeT expression in MIFs (Figure 10 C). Expression of the largeT antigen was not detected when using cDNA from the same cell lines before pBABE-zeo-largeT transfection. These analyses indicate that MIFs show solid expression of the largeT antigen and display enhanced morphological and proliferative properties. This suggests that MIFs can be classified as immortalized cell lines.

5.2.2. Cloning of NLS-Cre-recombinase to pcDNA3.1.TOPO-CT-GFP

Cre-recombinase including an N-terminal nuclear localization signal (NLS-Cre) was amplified from pTriEx1-HTNC, which was a gift from the group of Dr. Frank Edenhofer (Peitz et al. 2002). PCR was carried out using Taq DNA polymerase with the specific primers 4083 and 4084. The NLS-Cre amplicon was cloned into pcDNA3.1-CT-GFP using the pcDNA 3.1 GFP-TOPO TA Expression Kit. The cloned vector encodes a NLS-Cre-GFP, i.e. Cre-recombinase with an N-terminal nuclear localization signal and a C-terminal GFP-tag. The insert was validated by sequencing.

5.2.3. Generation and characterization of Tra2b-depleted MIFs

Murine immortalized fibroblasts (MIFs) were transfected with pcDNA3.1-NLS-Cre-GFP as described in 4.1.5. pcDNA3.1 propagates resistance to Geniticin (G418). Transfection efficiency was surveyed by GFP-fluorescence using transfection of pMAX-GFP and was usually far lower than 50 %. Geniticin selection using 2 mg/ml was started 48 h after transfection of pc DNA3.1-NLS-Cre-GFP. While replenishing Geniticin in a 4 day interval cells were cultivated for 6 weeks.

To functionally evaluate Cre-recombinase expression, *Tra2b* levels were determined on mRNA level using quantitative real-time PCR. Upon Cre-mediated disruption of the *Tra2b* locus the expression of *Tra2b* should drastically decrease. RNA was isolated from cells starting from 3 passages after NLS-Cre-GFP transfection onward. Expression analysis 3 passages after transfection demonstrated a reduction of *Tra2b* expression down to ~20 % of its initial values (Figure 11). By continuation of Geniticin selection *Tra2b* transcript levels decreased further to

background levels. This suggests that the transfected Cre-recombinase construct is functional and effectively abolished *Tra2b* expression in MIFs.

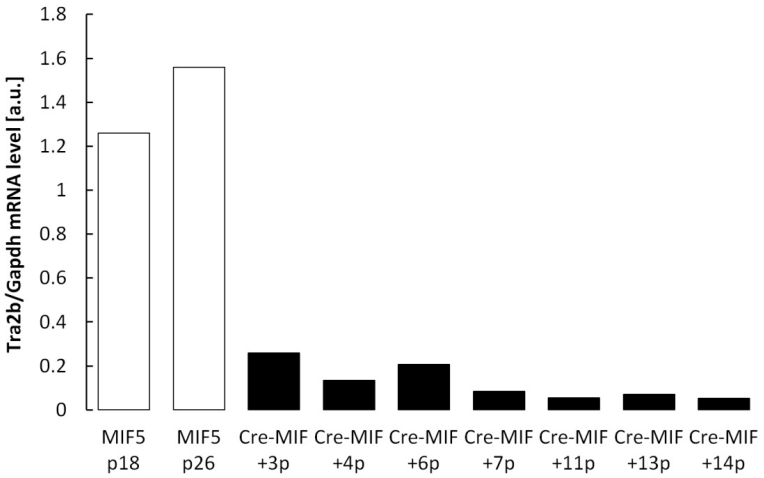

Figure 11 | Tra2b expression analysis in NLS-Cre-GFP transfected MIFs.
RNA was isolated from MIFs and MIFs transfected with NLS-Cre-GFP at various passages after transfection. *Tra2b* expression levels were analyzed using quantitative real-time PCR with the specific primers 3218 and 3219. Expression levels were normalized to *Gapdh* (primers 3999 and 4000). After transfection of Cre-recombinase and persistent selection using Geniticin, *Tra2b* transcript levels were effectively downregulated.

5.3. Characterization of the brain phenotype and analysis of neuronal development

Conditional *Tra2b* knock-out mice die shortly after birth but do not show gross morphological alterations in outer appearance. As the knock-out mice are depleted of *Tra2b* in neuronal and glial precursor cells it is self-evident to investigate the developmental progression of the brain. To allow comparison of different developmental stages and genotypes timed breedings were performed to isolate embryonic brains. Histological and immunohistological analyses were carried out using paraffin-embedded sections.

5.3.1. Histological analysis of neuronal-specific knock-out brains

To understand the underlying cause of perinatal lethality brains of KO mice were first histologically analyzed at various developmental stages (14.5 dpc until birth) using hematoxylin and eosin staining on paraffin-embedded brain sections. Compared to control mice KO mice showed a dilation of both the third and lateral ventricles (ventriculomegaly) (Figure 12). The dilation of the ventricles had its onset at around 14.5 dpc and progressed until birth. Dilation at 14.5 dpc was still moderate and cortical patterning was comparable to control animals. At 16.5 dpc the lateral ventricles occupied a big part of the brain hemispheres and the cortical layers were no longer distinguishable. The ependyma which constitutes the ventricular lining

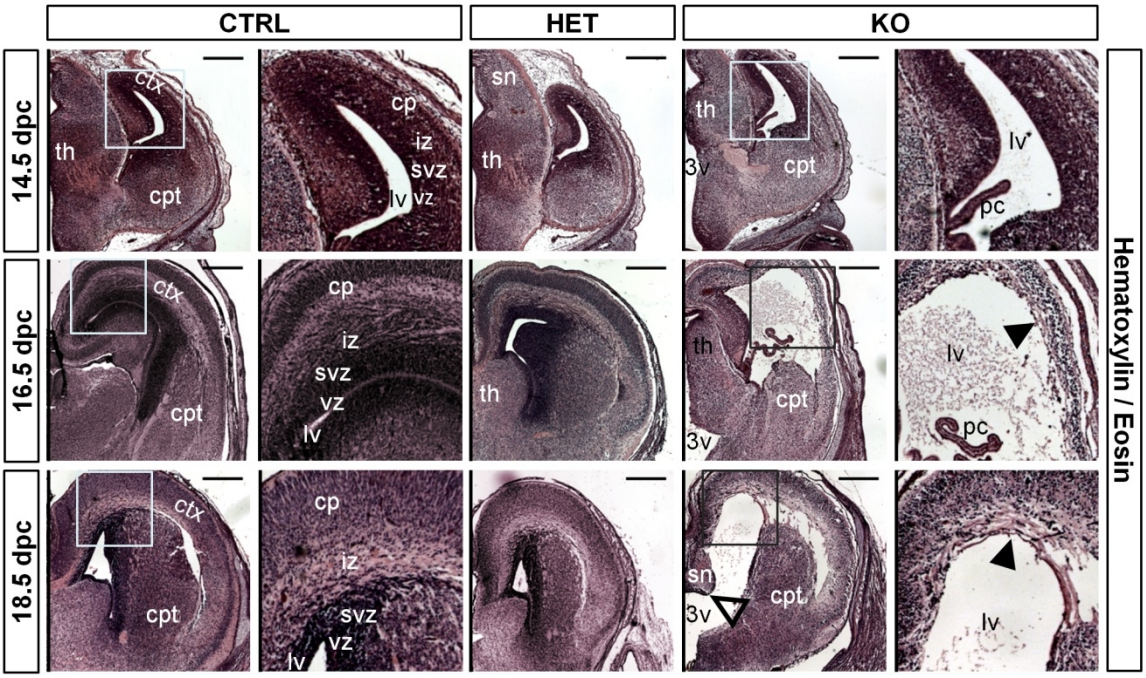

Figure 12 | Conditional ablation of Tra2b causes ventriculomegaly and disturbed cortical patterning.
Hematoxylin/Eosin staining of paraffin-embedded coronal sections at indicated developmental stages. KO animals but not controls or HET animals show ventriculomegaly of the third and lateral ventricles starting at around 14.5 dpc. Cortical layers are largely distinguishable at 14.5 dpc but cortical patterning and the ependymal lining of the lateral ventricle appears highly disturbed at later stages(black arrowheads). 18.5 dpc KO mice show anomalies at the caudal septum or the anterior thalamic area (empty arrowhead). Scale bar equals 400 µm; ctx, cortex; th, thalamus; cpt, caudoputamen; cp, cortical plate; iz, intermediate zone; svz, subventricular zone; vz, ventricular zone; sn, septal nuclei; 3v, third ventricle; lv, lateral ventricle; pc, choroid plexus

appeared to be fully absent or destroyed in KO animals, even though the choroid plexus was present but being detached as a consequence of ventriculomegaly. Furthermore, late stage brain morphology (18.5 dpc) showed irregularities in the caudal regions of the septum or the anterior thalamic area. Heterozygous knock-out animals retained normal cortical patterning with normal ventricle sizes at all stages of development. This is in line with the finding of Tra2b expression being comparable to controls on Western Blots (Figure 9) (Grellscheid et al. 2011a). This suggests that a single gene copy of *Tra2b* can functionally compensate protein function without producing an obvious phenotypic effect. These first histological results indicate that *Tra2b* is essential for brain development and normal cortical patterning.

5.3.2. Immunohistological analysis of neuronal-specific knock-out brains

5.3.2.1. Immunohistochemical analysis of Tra2b expression

To validate efficient downregulation of Tra2b in KO mice and to identify regions depleted of Tra2b or preserving Tra2b, paraffin-embedded brain sections were immunohistochemically stained. The ventricular and subventricular zones of the developing mouse brain showed notably high Tra2b expression (Figure 13). Further cell populations with

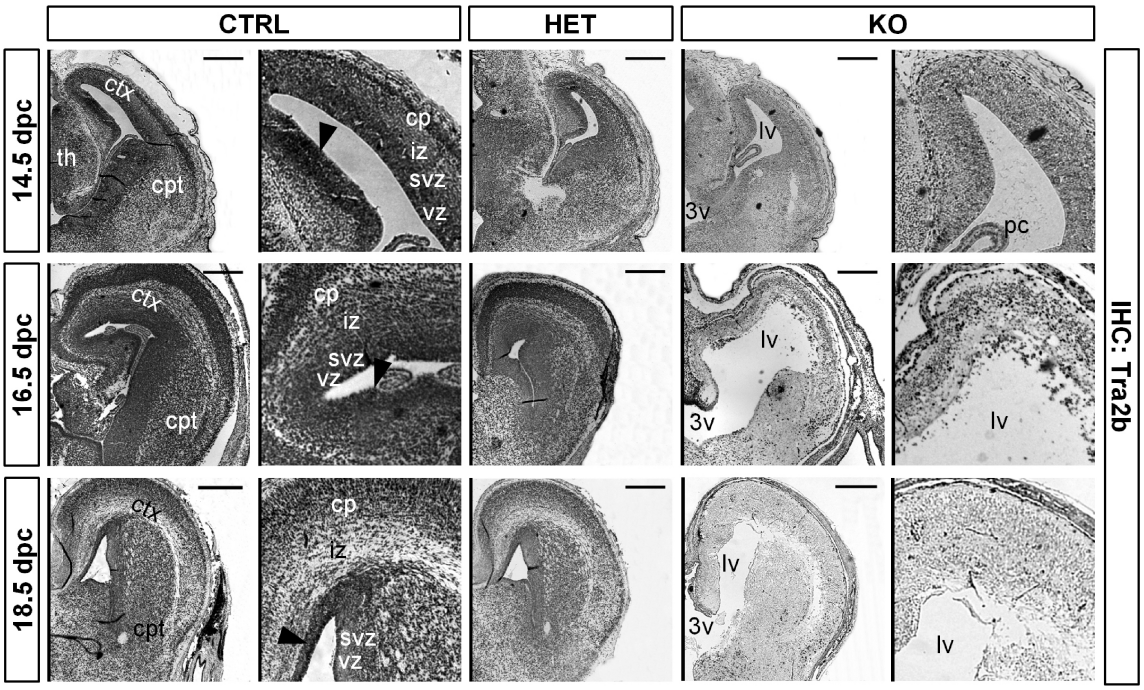

Figure 13 | Tra2b expression is drastically reduced throughout the brains of KO mice.
Immunostaining of Tra2b on paraffin-embedded coronal sections shows efficient downregulation of Tra2b protein in knock-out brains compared to CTRL and HET animals. Cells of the ventricular and subventricular zones of the cortex show strongest Tra2b expression in controls (black arrowheads). Expression in the according areas of KO brains is severely reduced. Scale bar equals 400 µm; ctx, cortex; th, thalamus; cpt, caudoputamen; cp, cortical plate; iz, intermediate zone; svz, subventricular zone; vz, ventricular zone; sn, septal nuclei; 3v, third ventricle; lv, lateral ventricle; pc, choroid plexus

prominent Tra2b expression were found in the outermost layer of the cortex. In contrast, Tra2b expression was strongly reduced in the ventricular and subventricular zones of KO animals. In general, brains of KO animals showed reduced levels of Tra2b expression but clearly retained localized centers of expression that likely comprise Nestin-negative cells (Figure 13). This was in so far expected as the mosaicism of *Cre* expression is based on the natural expression pattern of the *Nestin* promoter. Stainings of KO brains at 14.5 dpc appeared slightly more intense than at later stages, which could be indicative for residual Tra2b expression at this developmental stage. In contrast, cortical cell layers that were still present at 18.5 dpc did not show any remaining Tra2b expression. The absence or decay of the ventricular cortex layers in KO animals and the fact that these layers are loci of neuronal differentiation and neuronal outgrowth leading to cortex formation (Rakic 1988; Noctor et al. 2001) suggest that *Tra2b* plays a pivotal role in cortical development.

5.3.2.2. Immunohistochemical analysis of Ki67 and Caspase-3 expression

To further elucidate the course of the pathological progression we analyzed apoptotic events and the proliferative potential in KO versus control brains. This was in so far important as the observed brain phenotype including ventriculomegaly and failure in cortical patterning

can be caused by aberrant development or by the death of already developed structures. The latter could be caused by increased intracranial pressure that might lead to a normal pressure hydrocephalus based on reduced resorption of cerebrospinal fluid or duct obstruction in the ventricular system (Adams et al. 1965).

The proliferative potential and apoptosis were investigated using immunostainings for cleaved Caspase 3 (*Casp3*) and Ki-67. Localized presence of Casp3 indicating apoptotic events is a typical hallmark of neuronal development as each proliferation, differentiation and apoptosis orchestrate neuronal restructuring (Ikonomidou et al. 1999; Roth and D'Sa 2001). KO brains showed drastically elevated levels of apoptosis at 14.5 and 15.5 dpc but not at later stages of development (Figure 14 A). This primarily affected the cortical layers close to the lateral ventricles and the anterior thalamic region between the cavities of the third ventricle. Accordingly, the thalamic area presented morphological anomalies at later stages of development (16.5 dpc). The more dorsally located septal nuclei appeared largely unaffected, however (Figure 14 A, KO 16.5 dpc).

In concordance with these apoptotic events the proliferative potential of the ventricular and subventricular zone started to decrease at around 14.5 dpc and was fully lost at 16.5 dpc and all later stages (Figure 14 B). This temporal occurrence of events suggests proliferation loss being a direct consequence of initial apoptotic events ultimately leading to tissue loss and ventriculomegaly. The brain development of HET animals was not affected. In these brains neither an increase of apoptosis, nor a loss of proliferative potential was detected.

This demonstrates that haploinsufficiency of *Tra2b* has no phenotypic consequence and rules out any secondary effect of the *Nestin-Cre* transgene on brain development (Forni et al. 2006). In contrast, depletion of *Tra2b* in neuronal precursors leads to ventriculomegaly and aberrant cortical patterning. This appears to be a direct result of massive apoptotic events in ventricular and subventricular cortical layers and subsequent loss of the proliferative potential in these neurogenic areas. These findings clearly point out a pivotal role of *Tra2b* in neuronal development and survival.

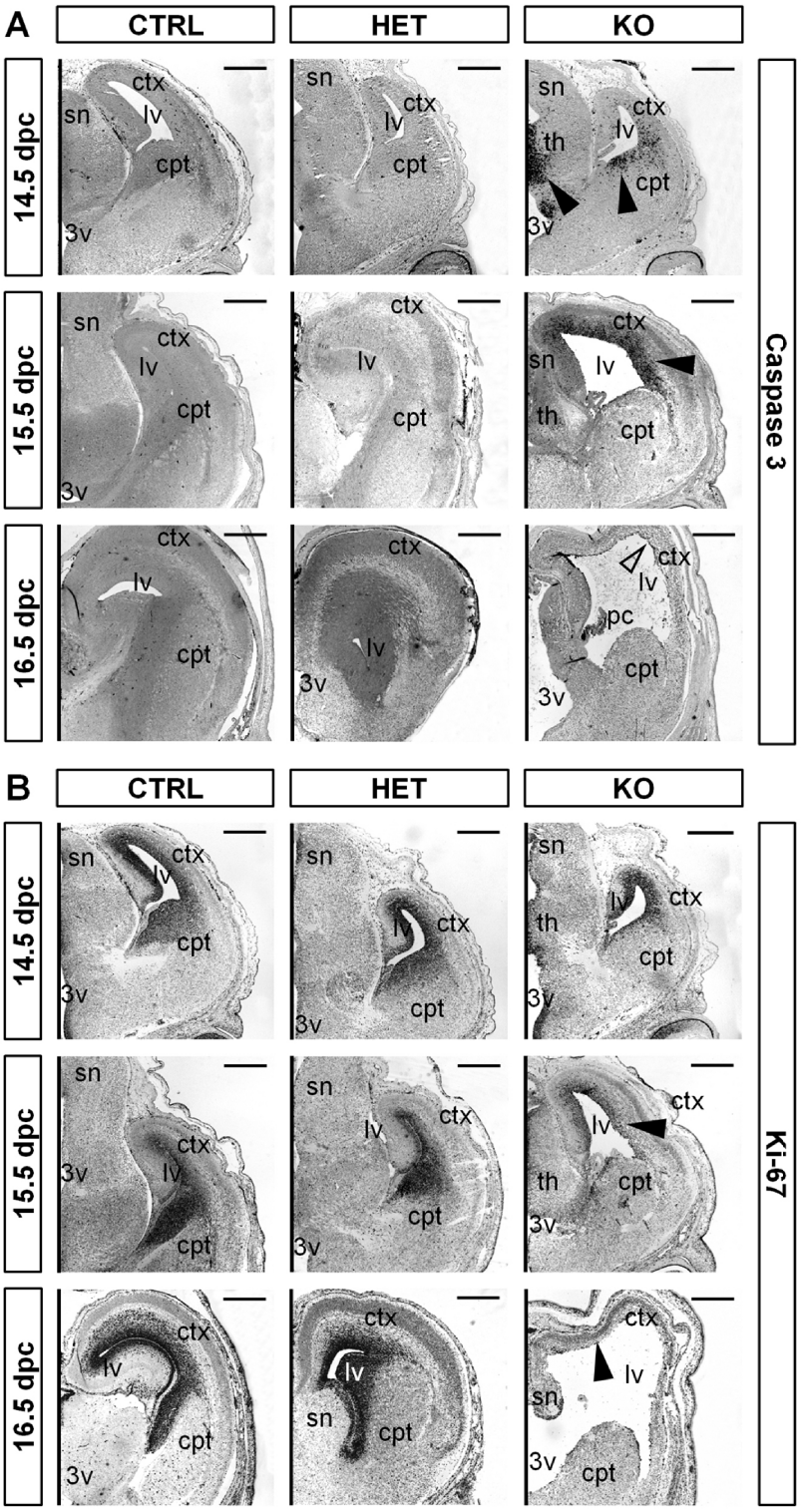

Figure 14 | Brain malformations are initiated by massive apoptosis in the cortex.
(A) Immunostaining for Caspase-3 on paraffin-embedded coronal sections indicates prominent apoptosis in the proximal cortical layers and in the thalamic area of 14.5 dpc and 15.5 dpc KO embryos (black arrowheads). Remaining cortical tissue does not show apoptosis at 16.5 dpc and later stages (light arrowhead). (B) Immunostaining for Ki-67 shows initial decrease of proliferation at 14.5 dpc which is fully lost at 16.5 dpc in KO animals (black arrowheads). CTRL and HET animals retain strong Ki-67 signals in the proximal cortical layers at all indicated developmental stages. Scale bar equals 400 µm; ctx, cortex; th, thalamus; cpt, caudoputamen; sn, septal nuclei; 3v, third ventricle; lv, lateral ventricle

5.4. Molecular biological analysis of neuronal-specific knock-out brains

In preparation for high throughput screening approaches as an effort to detect target transcripts of Tra2b, a detailed molecular biological analysis of *Tra2b*-depleted brains was carried out. Splicing targets of Tra2b that are known, have mostly been identified in *in vitro* systems. Such experiments have a limited power in stating the physiological relevance of a splicing process, as standard splicing assays were performed using minigenes while overexpressing *Tra2b*. This kind of approach might certainly allow a functional validation of a splicing process, but does not allow a physiological assessment. This chapter will focus on the physiological assessment of the splicing processes of well-known Tra2b target transcripts using conditional *Tra2b* knock-out mice. Validation of known targets *in vivo* would be a proof of concept for the identification of physiological splicing targets *in vivo* using conditional *Tra2b* knock-out mice.

Initial analyses showed correct tendencies in confirming the splicing of known targets, but these results were not significant. It was assumed that only cell populations in the brain that were depleted of *Tra2b* could display an effect in splicing. Cell populations that do not express *Cre-recombinase* continue expressing *Tra2b* and thus retain normal splicing. As whole brain RNA was used in these experiments, normal splicing in *Tra2b* expressing cells would mask altered splicing patterns in depleted cell populations. To assess the extent of mosaicism in KO brains, *Tra2b* expression was measured by quantitative real-time PCR and compared between single embryos and between different developmental stages. Measuring low levels of *Tra2b* using whole brain RNA might indicate knock-out in the vast majority of brain areas, while high levels would indicate a high degree of mosaicism with larger brain fractions retaining *Tra2b* expression. Indeed, though all observed KO mice displayed ventriculomegaly and loss of cortical patterning, the severity of the described brain phenotype appeared to vary slightly. Mosaicism of *Cre-recombinase* expression and retained *Tra2b* expression could well explain that phenomenon.

5.4.1. Expression analysis of Tra2b during development

Whole brain RNA was first analyzed for expression levels of *Tra2b* at different developmental stages by real-time PCR and normalized to the expression of *Gapdh* (Figure 15). In control animals the total expression of *Tra2b* in the brain was largely constant between 16.5 dpc and birth. Expression levels at 18.5 dpc were slightly but non-significantly lower than at 16.5 dpc or at birth. HET animals showed an approx. 50 % reduction of *Tra2b* transcript levels, while *Tra2b* was effectively down-regulated in brains of 16.5 dpc, 18.5 dpc and PND 1 KO animals (Figure 15 A). Though the reduction of *Tra2b* expression was significant, 16.5 dpc

brains of KO mice showed highest residual expression levels. As whole brain RNA was used, residual *Tra2b* expression is likely a result of *Nestin*-negative cells that do not express *Cre*-recombinase. To answer whether this high retention of *Tra2b* expression was a general phenomenon at 16.5 dpc, individual expression levels were compared (Figure 15 B). Though average *Tra2b* downregulation was significant, expression levels were individually different between single animals. Despite transcript levels were on average gradually reduced according to the genotype (horizontal lines in Figure 15 B), single animals within each genotype group showed untypical *Tra2b* expression levels. In the KO group three animals were found with expression levels equally high as seen in HET animals. For these animals also the Tra2b protein expression in the brain was expected to be accordingly higher, which makes it likely that such pups with high levels of residual *Tra2b* expression account for the sometimes observable moderate brain phenotypes.

In an attempt to validate this hypothesis RNA was isolated from paraffin-embedded brain sections of KO mice displaying severe and moderate brain phenotypes. We aimed to find a correlation between moderate phenotypes with retained *Tra2b* levels and between severe brain phenotypes and very low *Tra2b* levels. Analysis by quantitative PCR was strongly hampered by highly impaired quality of RNA obtained from paraffin-embedded sections. *Tra2b* transcripts were detected but their amounts varied drastically (data not shown). Therefore, a precise quantification was not possible, which was likely due to pronounced RNA fragmentation based on formalin fixation and paraffin embedding.

Aiming at the validation of previously reported Tra2b-related splicing processes *in vivo*, quantitative RT-PCRs specifically detecting alternatively spliced exons were performed. This is discussed in detail in chapter 5.4.3. Initial analyses showed only a tendency of well-known Tra2b-dependent splicing processes, which was not significant, however. Knowing that a fraction of KO brains retains significant expression levels of *Tra2b* this is not surprising, as splicing processes in *Tra2b* expressing cells are not impaired. For that reason all KO brains used in this study were analyzed for *Tra2b* transcript levels and only brains showing a strong reduction in *Tra2b* expression were used for subsequent analyses. To test this, an arbitrary threshold value of 0.4 in *Tra2b*/*Gapdh* expression was selected (dashed line in Figure 15 B). Using only mice with relative expression levels lower than 0.4 resulted in significant results when assaying for alternative splicing events (5.4.3). For that reason all KO brains with relative *Tra2b* expression levels of higher than 0.4 were excluded from analyses, as incomplete or highly mosaic knock-out hampers the analysis of Tra2b-related splicing processes.

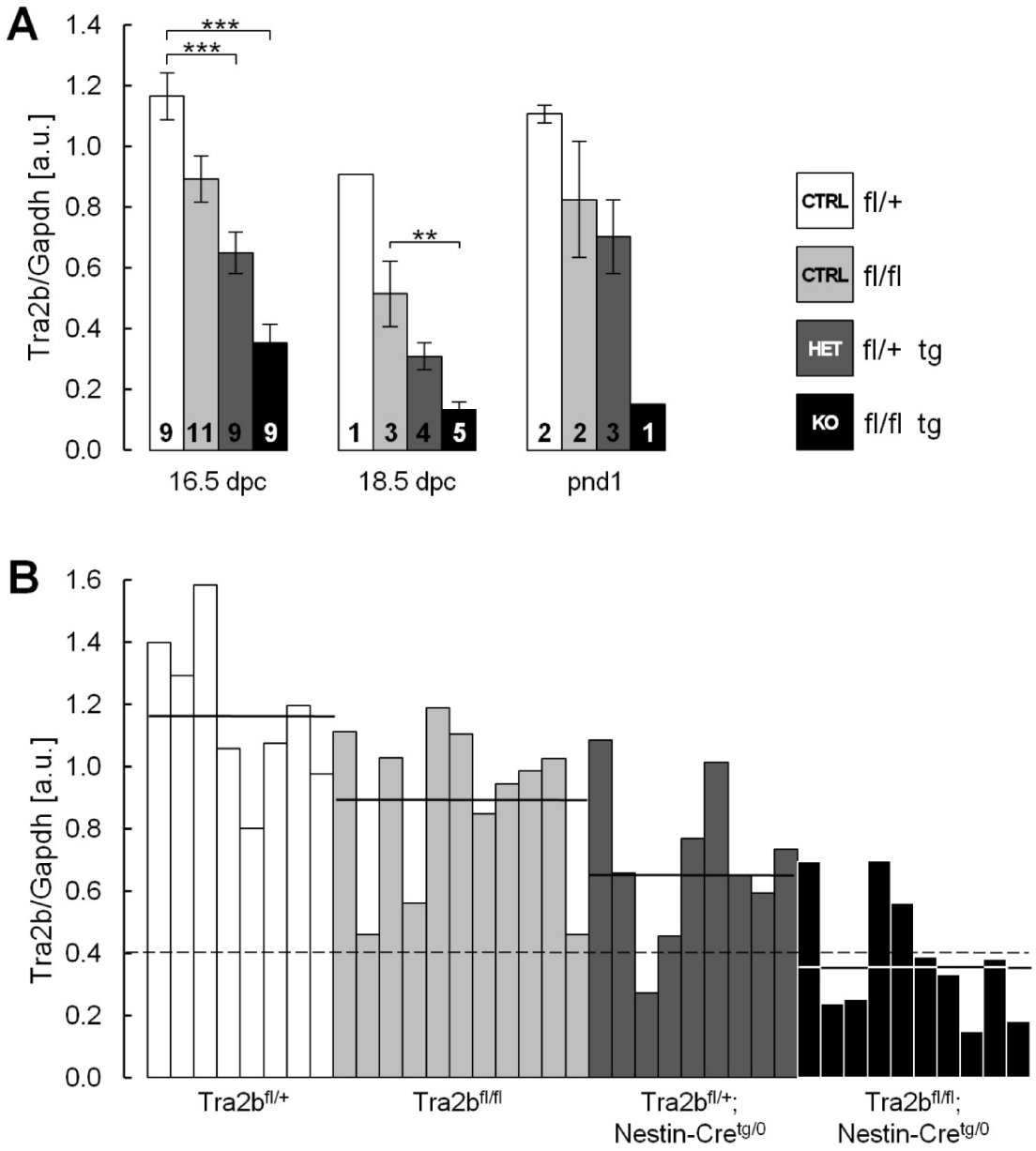

Figure 15 |Expression analysis of *Tra2b* during development.
Quantitative real-time PCR using whole brain RNA of control, HET and KO embryos at 16.5 dpc, 18.5 dpc and postnatal day 1. All mRNA levels have been normalized to *Gapdh* expression. Total *Tra2b* levels were specifically detected using the primers 3218 and 3219. Primers 3999 and 4000 were used to detect *Gapdh*. (A) Total *Tra2b* expression levels are effectively reduced in KO animals at 16.5 dpc, 18.5 dpc and PND1. (B) Individual expression levels of total *Tra2b* in 16.5 dpc brains largely behave accordingly to the respective genotype but are subject of variation within each group. Horizontal lines indicate the mean of each genotype group. The dashed line indicates a threshold value of 0.4. Numbers within bars represent the number of biological replicates (N).

5.4.2. Expression analysis of Tra2b isoforms

Tra2b is known to autoregulate its own expression level via a splicing feedback-loop by including cassette exon 2 into its own transcript (Stoilov et al. 2004). Exon 2 contains a premature termination codon rendering the transcript non-functional and thus decreasing the Tra2b protein concentration. This means that Tra2b balances its own concentration. High amounts of Tra2b protein facilitate effective inclusion of exon 2, thus rendering the transcript non-functional and as a consequence decrease the protein concentration. However, Tra2b auto-splicing might be regulated not solely by its own concentration but also by its phosphorylation. It has been suggested that phsophorylation of Tra2b regulates interactions with other splicing factors and also influences the cellular localization of Tra2b (Stoilov et al. 2004; Li et al. 2013).

Auto-splicing of Tra2b has hitherto been demonstrated in *in vitro* approaches like minigene experiments, which use artificial constructs and high levels of Tra2b protein. The following experiment aimed at analyzing the Tra2b self-splicing mechanism *in vivo* using conditional knock-out mice. In previous *Tra2b* quantitifcation experiments of this study the whole transcript amount of *Tra2b* was determined without differentiating between the functional (exon 2 skipped) and non-functional isoform (exon 2 included) (Figure 15). Here the inclusion of exon 2 was specifically detected using primers individually amplifying the functional or the non-functional isoform.

Using isoform-specific real-time PCRs on 16.5 dpc brains a significant down-regulation of both the functional (-78 %) and non-functional (-88 %) *Tra2b* isoforms was detected (Figure

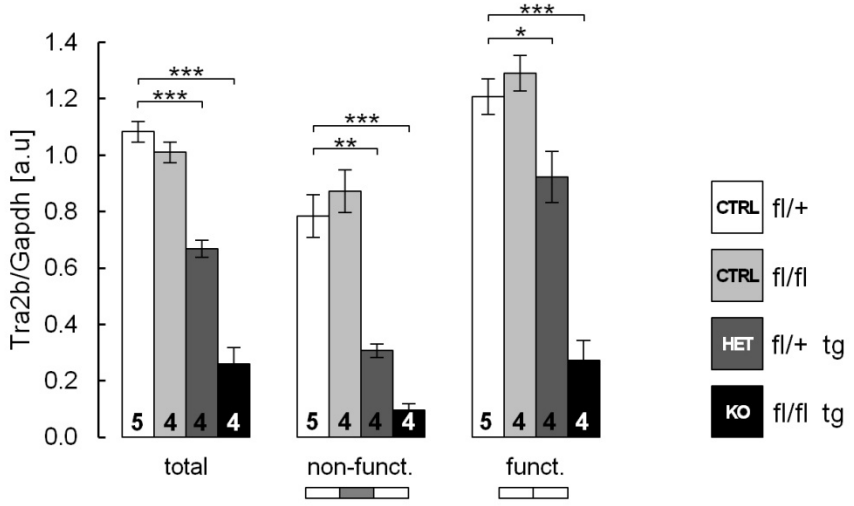

Figure 16 | Expression analysis of *Tra2b* isoforms.
RNA from 16.5 dpc brains was analyzed using real-time PCR. Expression levels of both the functional and non-functional *Tra2b* isoforms were measured using the primers 4271/4272 and 4269/4270. Expression levels were normalized to *Gapdh* (primers 3999/4000). Both isoforms were reduced in KO and HET mice. In HET mice the non-functional isoform was strongly reduced, while the functional isoform was maintained at almost normal levels. Numbers within bars represent the number of biological replicates (N).

16). In general the down-regulation of the non-functional isoform was stronger than for the functional isoform. Interestingly, in HET mice the levels of functional transcripts were only mildly but significantly reduced (-24 %) while non-functional transcripts were very strongly and significantly reduced (-61 %). These data suggest that the auto-regulatory splicing feedback-loop is functional in a physiological environment in mice (Grellscheid et al. 2011a). The only mildly altered levels of functional *Tra2b* transcripts in HET mice are fully consistent with survival, normal brain development (Figure 12) and Tra2b protein levels being comparable to controls (Figure 9). Loss of a single *Tra2b* gene copy resulted in an approx. 50 % reduction of total *Tra2b* transcript levels (Figure 15 A), while protein levels were maintained stable (Figure 9). Analysis of *Tra2b* isoforms demonstrated that both isoforms are not coordinately downregulated but instead being shifted to preserve higher amounts of functional transcripts. This suggests that loss of a single *Tra2b* gene copy can be compensated to a great extent via this auto-regulatory mechanism *in vivo*.

5.4.3. Proof of concept splicing analyses in Tra2b knock-out brains

Neuronal-specific *Tra2b* knock-out mice mostly showed a robust downregulation of *Tra2b* in the brain. However, the knock-out efficacy was low in some pups and *Tra2b* transcript levels were insufficiently reduced to validate well-known Tra2b-dependent splicing processes. To allow efficient validation of splicing events, KO mice with relative *Tra2b* expression levels of higher than 0.4 were excluded from analyses (as described in 5.4.1). Inclusion of samples with insufficient *Tra2b* knock-out prevented a reliable detection of known splicing events.

Besides the already validated Tra2b auto-splicing the following experiments were aiming at the validation of known Tra2b target transcripts *in vivo* under physiological conditions. This analysis included transcripts of the microtubule associated protein tau (*Mapt*) and the vesicle coating protein clathrin light chain b (*Cltb*) (introduced in chapter 1.5). Knowledge about Tra2b-related exon splicing of these transcripts has been obtained in minigene or exon trapping approaches (Stamm et al. 1992; Stamm et al. 1999; Jiang et al. 2003; Kondo et al. 2004) but correlations of enhanced *Tra2b* expression and *Mapt* exon 10 inclusion have been reported in post mortem Alzheimer's disease brains as well (Glatz 2005; Glatz et al. 2006). The validation of known Tra2b-dependent splicing targets would constitute a proof of concept for the identification of alternative splicing events *in vivo* using neuronal-specific KO mice. This would pave the way for the identification of novel Tra2b splicing targets using high throughput screening approaches.

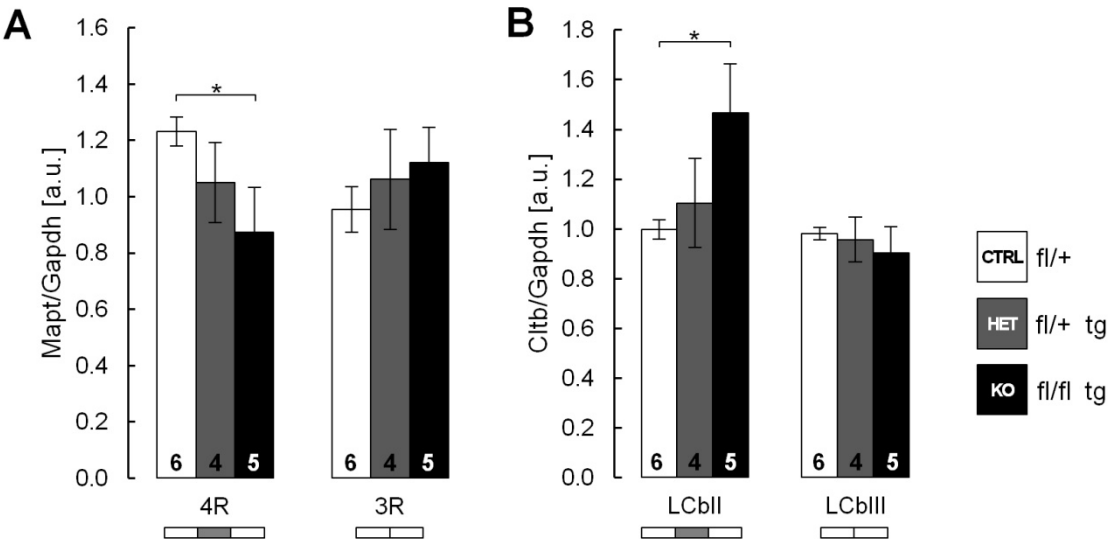

Figure 17 | Splicing analysis of *Mapt* exon 10 and *Cltb* exon 5.
RNA from 16.5 dpc brains with *Tra2b*/*Gapdh* expression lower than 0.4 was analyzed using real-time PCR. Expression levels of both *Mapt* and *Cltb* isoforms were measured using isoform-specific primers: *Mapt* 4R (3990/3991), *Mapt* 3R (3931/3934), *Cltb* II (3939/3941), *Cltb* III (3939/3940). Expression levels were normalized to *Gapdh* (primers 3999/4000). Tra2b promotes inclusion of *Mapt* exon 10 and promotes skipping of *Cltb* exon 5.

Using isoform-specific real-time PCRs the splicing of microtubule associated protein tau (*Mapt*) exon 10 and the vesicle coating protein clathrin light chain b (*Cltb*) exon 5 was analyzed in 16.5 dpc brains. In *Mapt* inclusion of the cassette exon 10 was reported to be promoted by Tra2b (Kondo et al. 2004). Accordingly, quantification of *Mapt* isoforms showed a moderate but significant decrease of the 4R isoform containing exon 10 in KO mice (Figure 17 A). HET mice only showed an insignificant decrease in exon 10 inclusion, suggesting that Tra2b levels previously found to be comparable to controls can also functionally fully compensate splicing processes. In contrast the 3R isoform displayed a slight but non-significant upregulation. These data indicate that Tra2b positively regulates *Mapt* exon 10 splicing under physiological conditions. These results are in good concordance with increased *Mapt* 4R levels in postmortem AD brains, which displayed an increased expression of *Tra2b (Glatz et al. 2006)*.

Furthermore, the splicing of *Cltb* exon 5 was analyzed. Here Tra2b has been reported to promote skipping of exon 5 (Stamm et al. 1992; Stamm et al. 1999). Accordingly, absence of Tra2b promoted inclusion of exon 5 and caused a solid and significant increase of isoform LCbII expression (Figure 17 B). No changes in expression were detected for the LCbIII isoform, which as a response to Tra2b depletion was expected to decrease. Also for *Cltb* no changes in splicing were detected in HET mice.

To summarize, selection of mice with *Tra2b*/*Gapdh* expression in brain lower than 0.4 allowed the *in vivo* validation of previously described Tra2b target exons in a physiological situation. Tra2b physiologically promotes inclusion of *Mapt* exon 10 and promotes skipping of *Cltb* exon 5.

5.4.4. Validation of splicing targets identified in HITS-CLIP experiments

In frame of collaboration with the workgroup of Prof. David Elliott novel Tra2b splicing targets have been identified by HITS-CLIP (high throughput sequencing cross-linking immunoprecipitation) on murine testis lysates (Grellscheid et al. 2011a). Analysis has shown a significant increase in exon inclusion of a poison exon of the *Tra2a* transcript and of a testis-specific exon of the nuclear autoantigenic sperm protein *Nasp*. *Tra2a* and *Tra2b* are paralogs and similar to the *Tra2b* transcript, inclusion of its exon 2 also introduces a premature termination codon in the *Tra2a* transcript. This might constitute a post-transcriptional regulatory mechanism similar to *Tra2b* (Grellscheid et al. 2011a). *Nasp* encodes the histone chaperone nuclear autoantigenic sperm protein, which encodes two isoforms: the somatic isoform *sNasp* and the testis-enriched isoform *tNasp*, which comprises the roughly 1 kb long T-exon (Richardson et al. 2000). Besides its high expression in testis *tNasp* is reported to be highly abundant in embryonic and tumor cells, which are both highly proliferative cell types (Richardson et al. 2000; Richardson et al. 2006). O'Rand and colleagues have demonstrated that selective depletion of *tNasp* in tumor cells abolished proliferation (Alekseev et al. 2011) and ablation of *Nasp* causes early embryonic lethality in mice (Richardson et al. 2006). This data implicates *Nasp* playing a fundamental role during embryonic development and probably in highly proliferative cells in general.

In a collaborative effort experiments were aiming at the physiological validation of the Tra2b-dependent splicing processes of *Tra2a* exon 2 and the *Nasp* T-exon (exon 7) in neuronal-specific knock-out mice. RNA from 16.5 dpc brains was analyzed by real-time PCR using isoform-specific primer pairs each detecting a single isoform of the *Tra2a* and *Nasp* transcripts. Both the T-*Nasp* and the somatic isoform of *Nasp* were detected in the developing murine brain (Figure 18 A). HITS-CLIP experiments suggested a Tra2b-dependent inclusion of the *Nasp* T-exon. Accordingly, expression of the T-*Nasp* isoform was significantly more than 5-fold decreased in KO mice compared to controls (Figure 18 A). HET mice displayed a minor and non-significant downregulation of the T-isoform. These data clearly demonstrate that Tra2b is an *in vivo* activator of *Nasp* T-exon inclusion in the murine developing brain. Given the more than 5-fold reduction of *tNasp* expression in *Tra2b*-depleted brains, Tra2b likely has a prominent physiological role in regulating *Nasp* T-exon inclusion. Interestingly, no changes were detected for the *sNasp* isoform, which was expected to increase as a response to *tNasp* downregulation in KO mice. It cannot be excluded that other transcriptional or post-transcriptional mechanisms of regulation impact total *Nasp* expression as a response to the interference in splicing administered in this experiment.

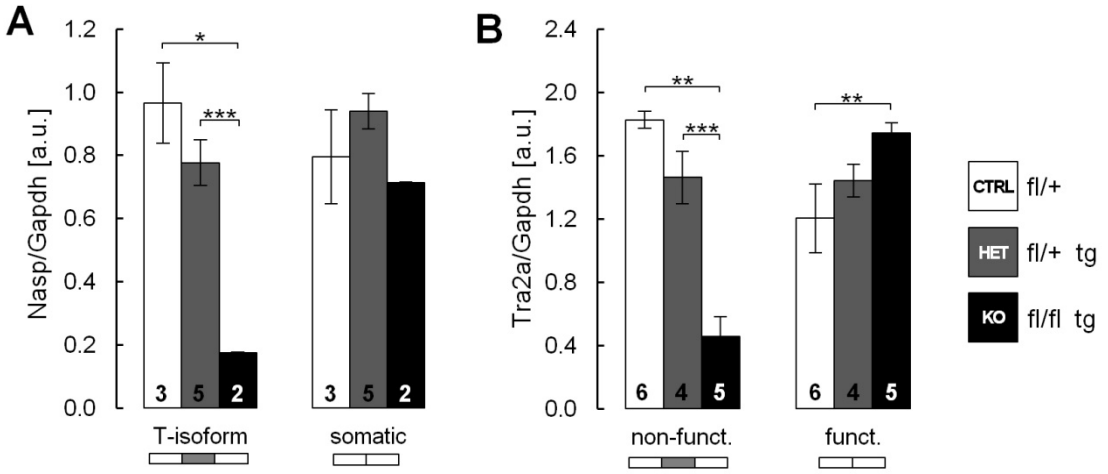

Figure 18 | Splicing analysis of *Nasp* exon 7 and *Tra2a* exon 2.
RNA from 16.5 dpc brains with *Tra2b*/*Gapdh* expression lower than 0.4 was analyzed using real-time PCR. Expression levels of both *Nasp* and *Tra2a* isoforms were measured using isoform-specific primers: *sNasp* (4037/4043), *tNasp* (4049/4042), non-functional *Tra2a* (4040/4044), functional *Tra2a* (4040/4046). Expression levels were normalized to *Gapdh* (primers 3999/4000). Tra2b promotes inclusion of *Nasp* exon 7 and *Tra2a* exon 2.

Tra2a exon 2 inclusion has been determined to be equally promoted by Tra2b (Grellscheid et al. 2011a). Supporting this finding, exon 2 inclusion was highly significantly reduced in KO mice, which resulted in a coordinate approx. 4-fold downregulation of the non-functional *Tra2a* isoform (Figure 18 B). HET mice showed a weak and non-significant downregulation. In contrast and as a response to massive exon 2 skipping in KO mice, a significant increase of the functional *Tra2a* isoform was detected. These data indicate that *Tra2a* is a physiological *in vivo* target of Tra2b. Depletion of Tra2b in the murine brain caused a physiological splicing shift promoting expression of the functional *Tra2a* isoform. Given that the two homologous splicing factors Tra2a and Tra2b have a similar domain structure and bind both to GAA-rich sequence elements (Tacke et al. 1998) a similar function and probably even an overlap in the spectrum of specific target transcripts can be expected. Thus, the *Tra2a* splicing shift promoting its functional isoform might constitute a compensatory mechanism to counteract the loss of Tra2b. Such mechanism might be true for some common splicing targets. However, the clearly determined missplicing events caused by absence of Tra2b (5.4.3 and 5.4.4) as well as the severe brain phenotype and perinatal lethality of KO mice indicate that Tra2b has unique splicing targets that are not compensated by Tra2a or other splicing factors.

5.5. Whole transcriptome sequencing analysis for the identification of Tra2b target transcripts

Mice ubiquitously depleted of *Tra2b* are early embryonically lethal. Here a neuronal-specific *Tra2b* knock-out mouse was generated. Mice homozygously depleted of *Tra2b* in the central nervous system die perinatally but are viable throughout embryonic development. This provides a unique chance to study Tra2b-related alternative splicing processes in a physiological setting during brain development. Conditional *Tra2b* knock-out mice have proven to be a potent *in vivo* model allowing the identification and quantification of alternative splicing processes in a physiological environment. This study aims at the identification of novel Tra2b target transcripts using high throughput screening approaches as whole transcriptome sequencing and exon array analysis. Identification of such targets will disclose putative off-targets that might be deregulated in treatment strategies using HDAC-inhibitors (1.6) and will shed light on splicing processes that are essential for neuronal development.

5.5.1. Data summary, analysis and statistics

The transcriptome sequencing was performed at the Cologne Center for Genomics using 3 RNA samples obtained from control mice and 3 obtained from KO mice. All RNAs were isolated from 16.5 dpc brains and well known Tra2b-dependent splicing processes have been reconstituted using these samples. Furthermore, all KO samples had *Tra2b*/*Gapdh* levels of less than 0.4 (as described in 5.4.1). This validation ensured that the used RNA samples can yield valid data of known splicing processes and are therefore presumably potent in identifying novel splicing targets.

Whole transcriptome sequencing yielded an average coverage of approx. 53-fold assuming a 100 Mb transcriptome. The total number of reads obtained and the resulting coverage were different between the samples, however (Table 5). The main objective of the whole transcriptome analysis was an unbiased assessment of relative exon abundances and their changes between CTRL and KO group to allow detection of alternative splicing events. Different coverage (i.e. different number of total reads) between individual samples and varieties in coverage depending on the transcriptome region exert significant influence on the quantification of individual exons. To overcome this bias, RPKM (**r**eads **p**er **k**ilobase and 1 **m**illion of reads) analysis was performed allowing normalization of the absolute read number with regard to the total transcript length and the total number of reads in the respective region (Figure 19). This allows compensation of variable coverage between samples and makes the read number independent of the respective transcript length. Moreover, exon abundances

were normalized to the total expression levels of a transcript. This prevents changes in total expression between CTRL and KO to influence the assessment of exon inclusion. To achieve that, RPKM was calculated for each whole transcript ($RPKM_T$) and for each of its exons ($RPKM_E$). The ratio of $RPKM_E/RPKM_T$ gives the normalized abundance of a respective exon. As shown in Figure 19 the mean of $RPKM_E/RPKM_T$ ratios was calculated for the control and for the KO group, which allowed a direct comparison of the average exon abundances (x-fold change). The x-fold change value gives the alteration of exon abundance as observed in the KO group relative to CTRL group. The significance level (p-value) of the relative change was calculated using Student's t-test comparing the 3 ratios of the CTRL group with the 3 ratios of the KO group.

In this transcriptome sequencing approach a total of 203,324 exons was identified. Figure 20 gives an overview about the significance and x-fold change distribution. The absolute majority of identified exons did not display significant changes in their abundances between CTRL and KO group (Figure 20 A). 198,900 exons were determined with p-values larger than 0.05 classifying them as non-significant; 4,424 exons with a p-value smaller than 0.05 (Figure 20 A, gray pie chart). The latter group of significantly differently abundant exons split up into 50 very highly significant exons with p-values smaller than 0.001 and into 615 highly significant exons with p-values between 0.001 and 0.01 (Figure 20 A, red pie chart). Most of the significantly different abundant exons had p-values between 0.05 and 0.01 (3,759 exons).

Table 5 | Whole transcriptome sequencing read coverage.
During whole transcriptome sequencing the average coverage was 53-fold assuming a 100 Mb transcriptome. The total read number was variable between samples.

sample	reads $x10^6$	coverage (100 Mb transcriptome)
CTRL A	45.2	42.2x
CTRL B	27.9	53x
CTRL C	39.4	74x
KO D	31.0	58.9x
KO E	26.2	49.8x
KO F	22.2	42.2x

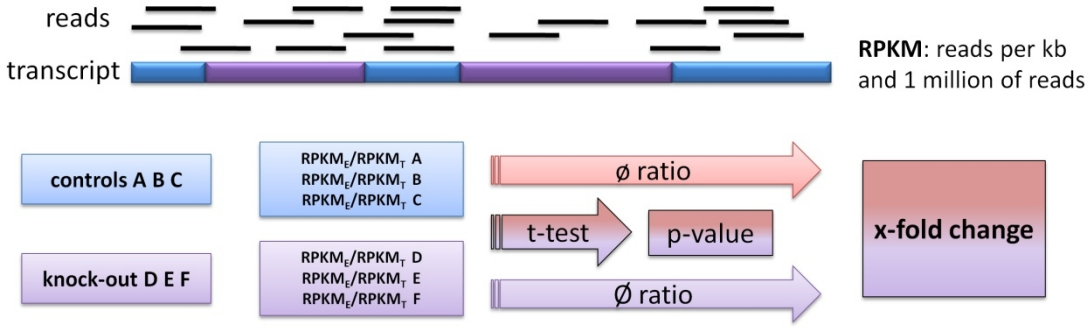

Figure 19 | RPKM Analysis of whole transcriptome data.
Sequence reads were mapped to the transcriptome database at Ensembl. RPKM was calculated to compensate for differences in the total number of reads, differences in the local distribution of reads on the transcriptome and for variable lengths of transcripts. RPKM was calculated for whole transcripts ($RPKM_T$) and for every of their exons ($RPKM_E$). Exon abundance was measured by the ratio of $RPKM_E/RPKM_T$ to allow an assessment unbiased by differential gene expression. The x-fold change of exon abundance was determined by the mean ratios obtained from the CTRL and the KO group. The significance of the difference in exon abundances was calculated using Student's t-test.

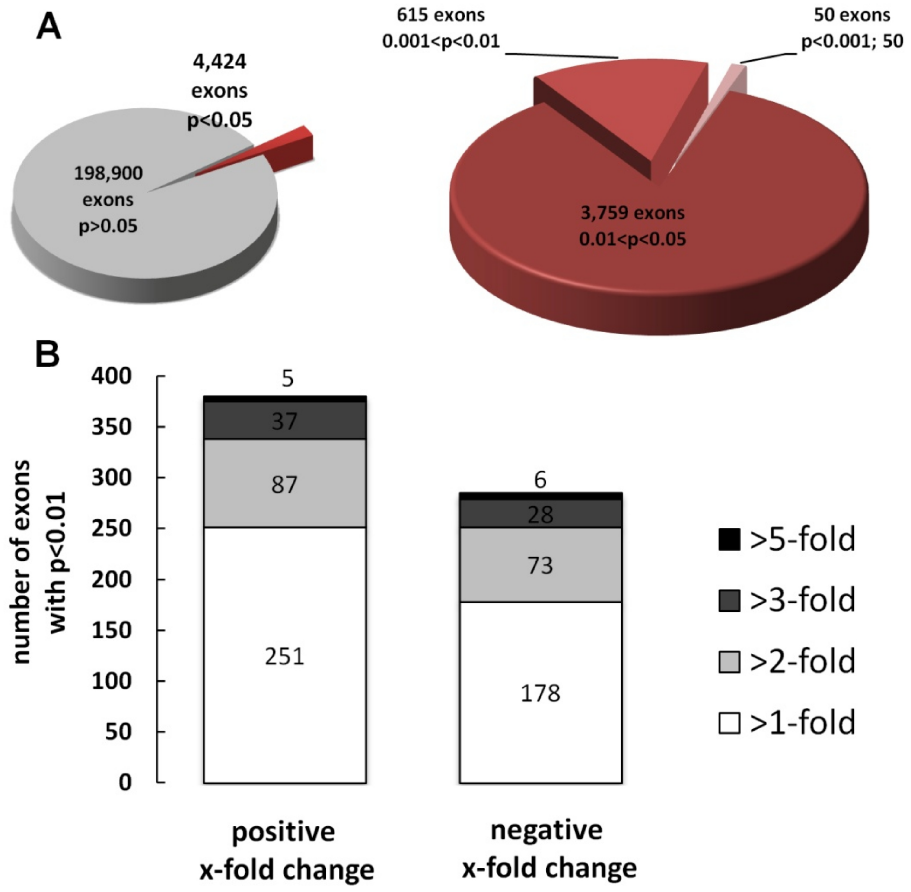

Figure 20 | Statistical distribution of significance levels and differential exon abundances.
Mapping of sequence reads obtained from transcriptome sequencing identified 203,324 exons. (A) Distribution of statistical significance for differential exon abundance between CTRL and KO group. Most exons were non-significantly different (gray pie chart, p>0.05). 4,424 exons were identified with p<0.05. These split up into 3,759 exons with 0.01<p<0.05, 615 exons with 0.001<p<0.01 and 50 highly significant exons with p<0.001 (red pie chart). (B) Distribution of x-fold changes in exon abundances of significantly different exons with p<0.01. 380 out of 665 exons displayed an upregulation in KO vs. CTRL group (i.e. putative inhibition by Tra2b) and 285 out of 665 exons showed a downregulation (i.e. putative activation by Tra2b). Most exons displayed moderate changes in abundance; only few showed higher than 3-fold change.

Exons from the group with p<0.01 (665 exons) mostly showed moderate x-fold changes between the KO and CTRL group. Figure 20 B gives an overview of changes in exon abundances. 380 exons were detected to be upregulated in KO mice indicating a putative inhibition of these exons by Tra2b. 285 exons were found to be negatively regulated in KO mice indicating a putative activation of these exons by Tra2b. Only a relatively small number of exons showed changes of larger than 3-fold.

Given that whole brain RNA of *Tra2b*-depleted mice and controls have been used for transcriptome sequencing and given the effect that *Tra2b* ablation has on brain development, it was expected that genes involved in neuronal function or neuronal development should be enriched among the identified candidates. All genes identified with alternatively spliced exons were cross-referenced with the gene ontology database at Ingenuity Systems. Exons that were detected as being non-significantly different (0.05<p<1.00) were found either associated or not associated to neuronal processes at highly similar frequencies (Figure 21). Interestingly, significant exons with p-values lower than 0.05 displayed a 3-fold enrichment of gene ontologies related to neuronal function or development. This indicates that genes or exons that were significantly detected to be alternatively spliced, have a 3-fold increased probability of having a neuron-related function.

To select candidate exons for validation by real-time PCR, identified exons were sorted according to their significance levels. As exons of high significance levels were shown to have an increased probability of serving a neuronal-related function, exons with best significance levels were selected for validation analysis. Table 6 gives an overview of the top 10 ranking

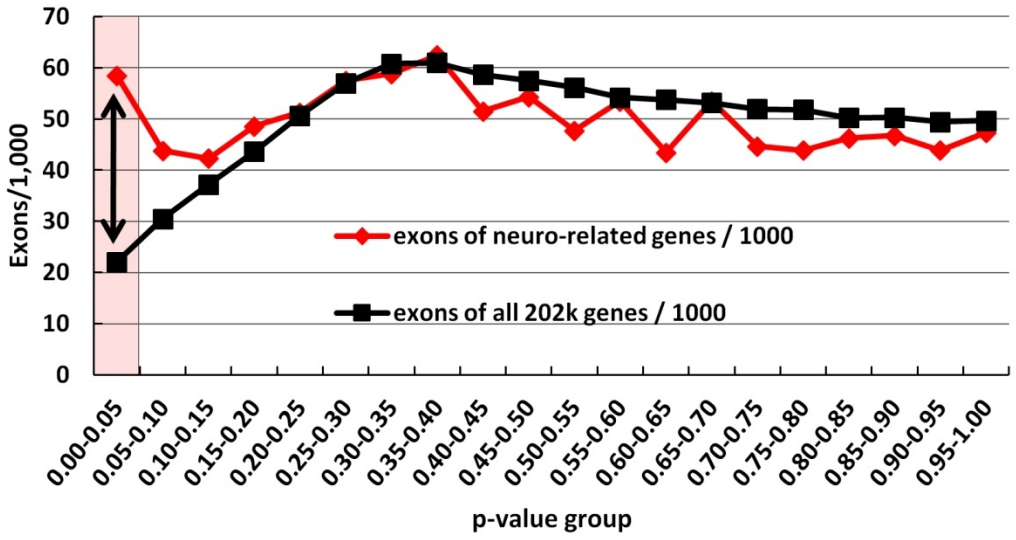

Figure 21 | Enrichment of genes related to neuronal function and development.
Genes identified with alternatively spliced exons were cross-referenced with the gene ontology database at Ingenuity Systems. Exons are displayed according to their significance levels and the exon number was normalized per 1,000 exons. Genes related to neuronal function or development were found 3-fold enriched among the group of significant different exons (p<0.05). There was no enrichment of neuronal-related genes in non-significant groups.

candidate genes containing putative alternatively spliced exons. The top ranking 100 candidates are listed in the appendix (Appendix, I).

Previously validated targets were not convincingly identified by whole transcriptome sequencing. The testis-specific exon of the *tNasp* transcript was identified as non-alternatively spliced and also the splicing of *Tra2a* exon 2 was not detected. To understand why established targets showing robust alternative splicing *in vivo* were not detected in transcriptome analysis, the local coverage was displayed as a custom track using the UCSC genome browser. As an example Figure 22 shows coverage tracks for *Nasp* T-exon for all analyzed samples. Inclusion of the *Nasp* T-exon is promoted by Tra2b. As expected, KO samples generally displayed fewer reads mapping to the T-exon, while controls showed many reads mapping to that region. In contrast to CTRL samples A and C, CTRL sample B showed a drastically reduced number of reads. This phenomenon was not restricted to the T-exon of the *Nasp* transcript but was observed for multiple targets. It is likely that a reduced coverage of CTRL B accounts for the strongly reduced significance. As the RPKM-based normalization should compensate deviating coverage of a sample, it appears likely that CTRL B did not suffer from generally low coverage per se (Table 5). Localized deviations in coverage could probably explain the failure of RPKM-based normalization. This idea is further supported as top ranking candidates did not show a comparably strong reduction of the local coverage in CTRL B (Figure 23).

The top ranking candidates *Appbp2* (Amyloid beta precursor protein binding protein) and *Htt* (Huntingtin) are both related or probably related to neuronal function and disease (Cisbani and Cicchetti 2012). Whole transcriptome analysis suggested that Tra2b promotes inclusion of *Appbp2* exon 4 and promotes skipping of *Htt* exon 60 (Table 6). Moreover, local coverage of these candidate exons did not appear to be severely affected by coverage anomalies of CTRL sample B (Figure 23). Given the adequate data quality for these candidate exons and their strong implication in neuronal function *Appbp2* exon 4 and *Htt* exon 60 were selected for validation experiments using quantitative real-time PCR.

Table 6 | Top 10 candidates obtained by whole transcriptome sequencing

gene designation	gene symbol	index	p-value
amyloid beta precursor protein binding protein 2	Appbp2	-3.24	0.000025
renin binding protein	Renbp	-4.11	0.000039
Huntingtin	Htt	1.52	0.000055
(embryonic lethal, abnormal vision, Drosophila)-like 2	Elavl2	17.30	0.000064
ribosomal protein L34	Rpl34	-1.71	0.000066
expressed sequence AI413582	AI413582	1.25	0.000103
transmembrane 7 superfamily member 2	Tm7sf2	1.28	0.000121
nuclear autoantigenic sperm protein	Nasp	5.20	0.000130
zinc finger protein 507	Zfp507	1.47	0.000199
BAH domain and coiled-coil containing 1	Bahcc1	-2.44	0.000210

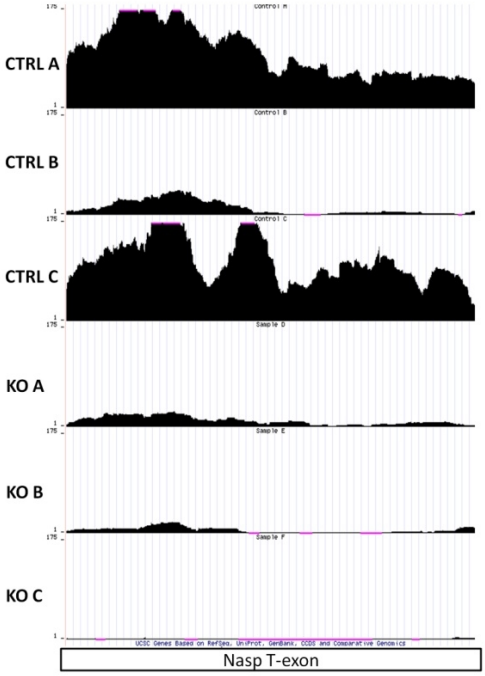

Figure 22 | Local coverage data of the *Nasp* T-Exon.
Coverage data from the whole transcriptome analysis was loaded as a custom track to the UCSC Genome Browser. *Nasp* T-exon inclusion is promoted by Tra2b. Accordingly, CTRL samples show a high coverage of that exon, while KO samples show rather low coverage. CTRL sample B deviates drastically, as its local coverage of the *Nasp* T-exon is strongly reduced compared to CTRL samples A and C.

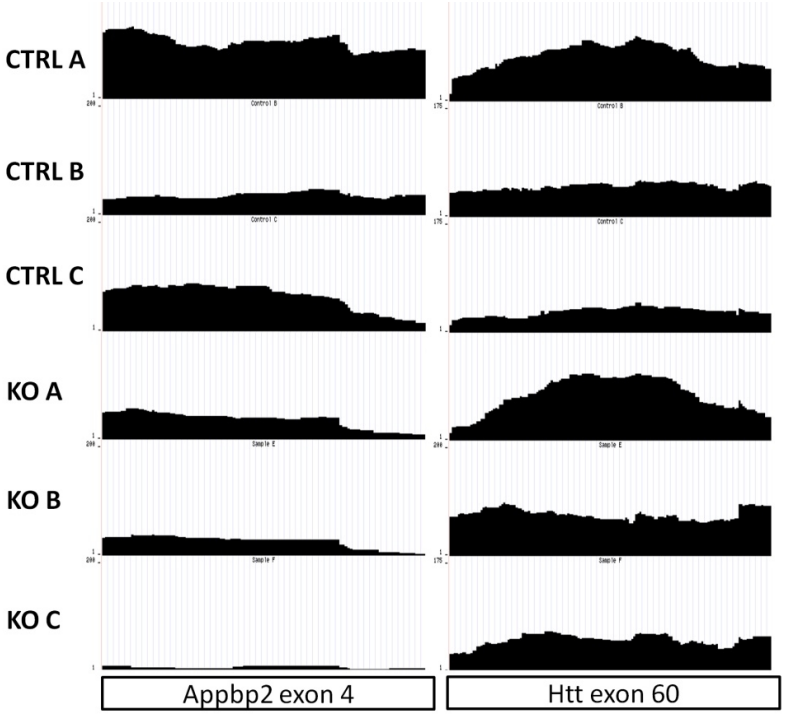

Figure 23 | Local coverage data of *Appbp2* exon4 and *Htt* exon 60.
Coverage data from the whole transcriptome analysis was loaded as a custom track to the UCSC Genome Browser. *Appbp2* exon 4 was predicted to be promoted by Tra2b, while *Htt* exon 60 was predicted to be blocked by Tra2b. The local coverage of each of these exons supports this idea. CTRL B did not appear to be severely affected by very low coverage in the depicted regions.

5.5.2. Validation of top ranking candidates

Quantitative real-time PCRs were carried out using RNA isolated from 16.5 dpc brains to validate top ranking candidates obtained from whole transcriptome sequencing. Using isoform-specific PCRs, isoforms of the *Htt* transcript (FL and Δ60) and of the *Appbp2* transcript (FL and Δ4) were specifically detected. In contrast to the predictions from whole transcriptome sequencing, neither of the two candidate exons was confirmed to be differentially spliced in KO mice (Figure 24). *Appbp2* exon 4 inclusion was predicted to be promoted by Tra2b. For *Htt* exon 60 Tra2b was predicted to promote exon skipping. Determination of isoform expression levels by real-time PCR was highly precise and specific as validated by PCR product sequencing. Both measured isoforms of each transcript are expressed at highly similar levels regardless of the genotype or *Tra2b* transcript levels. These results indicate that splicing of *Htt* exon 60 and *Appbp2* exon 4 is not altered upon *Tra2b* depletion and suggest that these exons are very likely not targets of Tra2b.

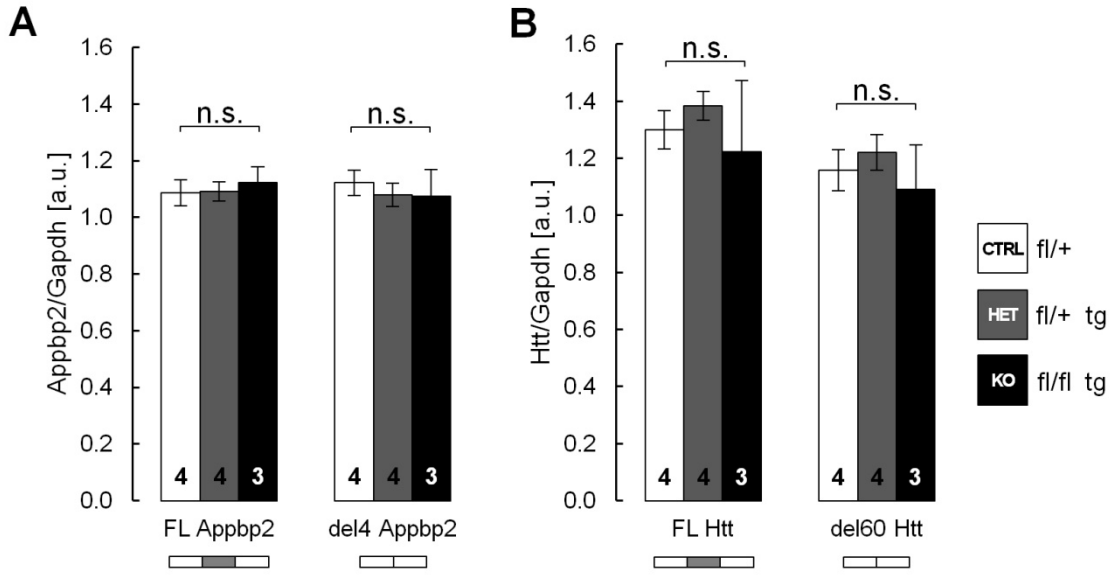

Figure 24 | Splicing analysis of *Appbp2* exon 4 and *Htt* exon 60.
RNA from 16.5 dpc brains with *Tra2b*/*Gapdh* expression lower than 0.4 was analyzed using real-time PCR. Expression levels of both *Appbp2* and *Htt* isoforms were measured using isoform-specific primers: FL *Appbp2* (4277/4278), Δ4 *Appbp2* (4279/4280), FL *Htt* (4273/4274), Δ60 *Htt* (4275/4276). Expression levels were normalized to *Gapdh* (primers 3999/4000). Tra2b is very likely not a slicing regulator of *Appbp2* exon 4 and *Htt* exon 60.

5.5.3. Further proceedings in WTS data analysis

First experiments aiming at the validation of putative Tra2b target exons by real-time PCR were fully ineffective. To exclude the possibility of erroneous data analysis and detection of false candidates based on sequence coverage anomalies in CTRL sample B, the whole transcriptome sequencing of this CTRL sample was repeated. Data analysis was repeated as described in 5.5.1. Substitution of the old by new CTRL data caused a drastically increased number of significantly detected exons. The number of alternatively spliced exons with $p<0.001$ increased from 50 to 415, exons with $p<0.05$ increased from 4,424 to 21,298. This strongly suggests that data originating from the initial control sample B contained indeed numerous outliers that were dissimilar from the other control samples. This caused generally low significance levels and prevented many exons from being significantly detected.

In the initial analysis most changes in splicing were positive (Figure 20 B), meaning that the majority of significant exons was detected to be repressed by Tra2b as they on average had higher levels in KO mice. After substitution of CTRL sample B most of the significant splicing effects were found negative with lower levels in KO mice indicating splicing inclusion by Tra2b. This shift clearly demonstrates the strong impact that one sample of poor data quality can exert on the analysis outcome.

The newly obtained data was filtered to increase the probability of picking a true target exon of Tra2b. Criteria for selecting candidate exons in this second approach included as before a highly significant detection. Moreover, candidates with a high x-fold change were preferred. Sequences of the detected exons were analyzed for the total number and the frequency of AGAA or GAA binding sites. Exons with a high frequency of putative Tra2b binding sites were preferred. More candidates were selected for validation experiments including Las1-like (*Las1l*) exon 6, the cAMP-dependent protein kinase *Prkar2b* exon 9, Renin binding protein (*Renbp*) exon 6 and Asparaginyl-tRNA synthetase (*Nars*) exon 4. Splicing of these exons in CTRL and KO mice was analyzed by real-time and semi-quantitative RT-PCR. To summarize, none of these identified splicing processes could be validated: Of each transcript only the isoform including the respective exon could be detected. This was equally true in both CTRL and KO mice. This suggests that isoforms skipping the respective exons do either not exist or are expressed at extremely low levels. These and previous failures in validation of splicing targets indicate that neither a highly significant detection, high x-fold change in splicing, or enrichment of Tra2b binding sites, nor a combination of all criteria yield true Tra2b splicing targets.

RNA sequencing has become a potent tool to determine whole spectrum gene expression profiles and it also finds increasing application in the identification and

quantification of alternative transcript isoforms (Churko et al. 2013; Costa et al. 2013). However, quantitative expression analysis from whole transcriptome data makes great demands on sequencing technology and its accuracy. The quantitative accuracy of high throughput sequencing approaches is the coverage. The higher a region or transcript sequence is covered, the more precise is the assessment of its expression. Toung and colleagues have demonstrated that a total number of 100M reads has been sufficient for the sole detection of the majority of expressed genes. However, for an accurate quantitative assessment of gene expression about 500M reads or more have been required (Toung et al. 2011). Read numbers achieved in this study reach from 45.2M down to 22.2M with an average coverage of 53-fold, which is far below the required coverage to allow solid evaluation of expression levels. With regard to the failure in validating splicing targets suggested by transcriptome sequencing, very low coverage as observed here, might be accounting for the detection of false candidates. Transcripts that are weakly expressed should be even more prone to such effect as low expression combined with a low coverage can make the smallest difference in expression appear big and significant. This is in line with the finding, that many isoforms tested during the validation process could not at all or only very weakly be detected. Furthermore, brains of KO mice showed a very severe phenotype that included the loss of whole brain areas associated with apoptosis (Figure 14). This causes changes in the cell-type composition when comparing CTRL with KO brains. These inert differences between these groups might also be accounting for a big fraction of artifact results obtained in the whole transcriptome analysis. Using whole brain RNA for analysis, such artifacts would then not be based on alternative splicing but be a result of secondary effects and alterations in gene expression based on selective tissue loss in the brain. This possibility will be revisited and evaluated in the following chapters.

5.6. Mouse exon array analysis for the identification of Tra2b target transcripts

Failure of the whole transcriptome sequencing approach to identify novel Tra2b splicing targets was likely caused by a too low sequencing coverage that made it impossible to reliably quantify transcript and isoform levels. Moreover, solid and significant identification of targets was very susceptible to irregularities in a single sample, as due to the high cost of a whole transcriptome sequencing run three years ago a rather small 3 CTRL versus 3 KO comparison was performed. As an alternative attempt to identify novel Tra2b splicing targets quantification was repeated using mouse exon arrays. Compared to whole transcriptome sequencing exon arrays are rather inexpensive and allow comparison of a larger sample number. In this approach 5 CTRL RNA samples were compared to 5 KO samples. Procedures carried out in preparation and during mouse exon array analysis are described in chapter 4.5.2 of the methods section.

5.6.1. RNA quality control assessment

In preparation for exon array analysis RNA samples were checked for good integrity and quality on the Experion gel electrophoresis system as described in 4.4.2.6. Total RNA as it was isolated comprises primarily ribosomal RNAs (rRNAs) and mRNAs only contribute a very small fraction to the total number of RNA molecules. Visualization of rRNAs on agarose gels or by capillary gel electrophoresis allows assessment of the general RNA integrity. In RNA samples isolated from eukaryotic cells or tissues the 28S rRNA and the 18S rRNA should be detectable at sizes of 4,718 nt and 1,874 nt, respectively. Sharp bands or signals of well defined size are indicative of intact rRNA, which likely is transferable to mRNAs. Blurry bands or smear are indicative of degraded or fragmented rRNAs.

Results obtained from RNA analysis are shown in Figure 25. All RNA samples isolated from whole brains of CTRL and KO mice showed sharp bands of the 28S and 18S rRNA (Figure 25 A). For ratio analysis the results of capillary gel electrophoresis were automatically plotted. This is exemplarily shown for the first control sample in Figure 25 B. Samples displayed clear peaks for both rRNAs and no or very weak smear and the automated quality assessment program of the gel electrophoresis system indicated all samples suitable for downstream analyses. Size deviations observed for the 28S rRNA (it appears smaller than it theoretically is) are likely due to inaccuracies in ladder calibration of the capillary electrophoresis system that should not have impaired analysis in any way.

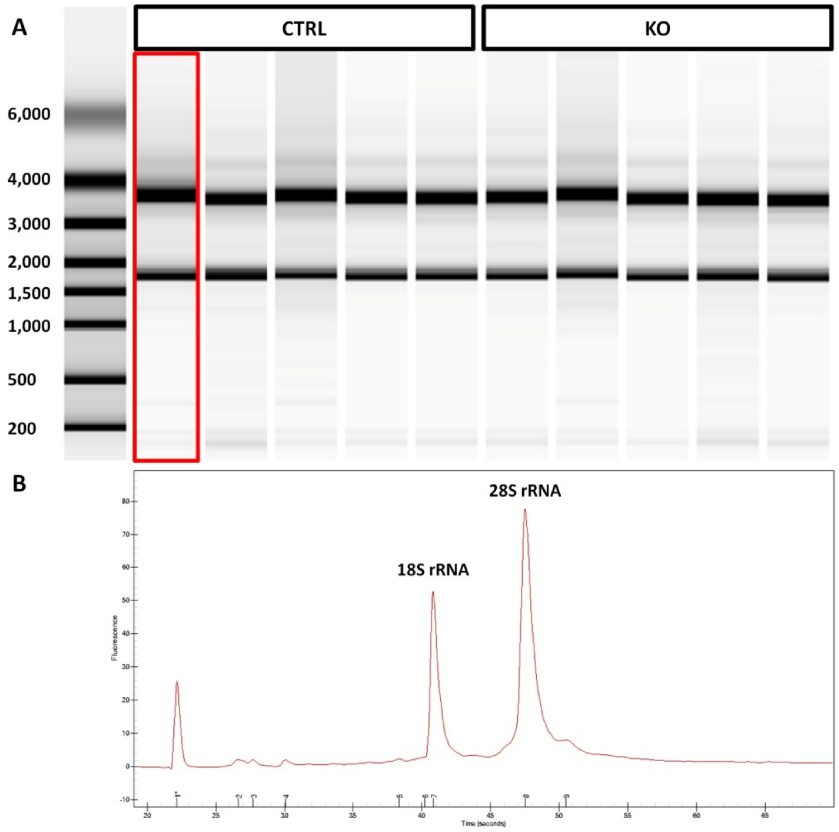

Figure 25 | Analysis of RNA integrity.
RNA samples were analyzed on the Experion capillary gel electrophoresis system (BioRAD). (A) The virtual gel output shows sharp bands of the 28S and 18S rRNA at defined sizes indicating good RNA integrity. (B) The detection plot of the capillary electrophoresis shows distinct peaks for both ribosomal RNAs. Automated quality assessment by the Experion software indicated high quality for all tested RNAs.

5.6.2. Data summary, statistics and filtering

To identify novel splicing targets for Tra2b, analyses of 5 CTRL and 5 KO whole brain RNAs was performed using Affymetrix mouse exon 1.0 ST arrays. This array type detects single exons via antisense probes and accordingly requires labeled sense-strand targets for detection. The array represents a total of 1.2 million probe sets including 835,897 probe sets supported by gene prediction and 266,791 probe sets that are supported by the Ensembl transcript database (Affymetrix product information). Each probe set represents a single exon and mostly comprises 4 distinct probes to specifically hybridize to sequences of the respective exon. RNA samples used for this analysis were tested for integrity as described in 5.6.1 and *Tra2b*/*Gapdh* transcript levels were assured to be smaller than 0.4 as described in 5.4.1. Preparation, labeling and hybridization were performed as stated in 4.5.2.

Before proceeding with the statistical and bioinformatic analyses of candidate exons, the quality of the array read out data was assessed by comparing the dispersion of cell intensities between the arrays. A summary of this analysis is shown in Figure 26. The box plot

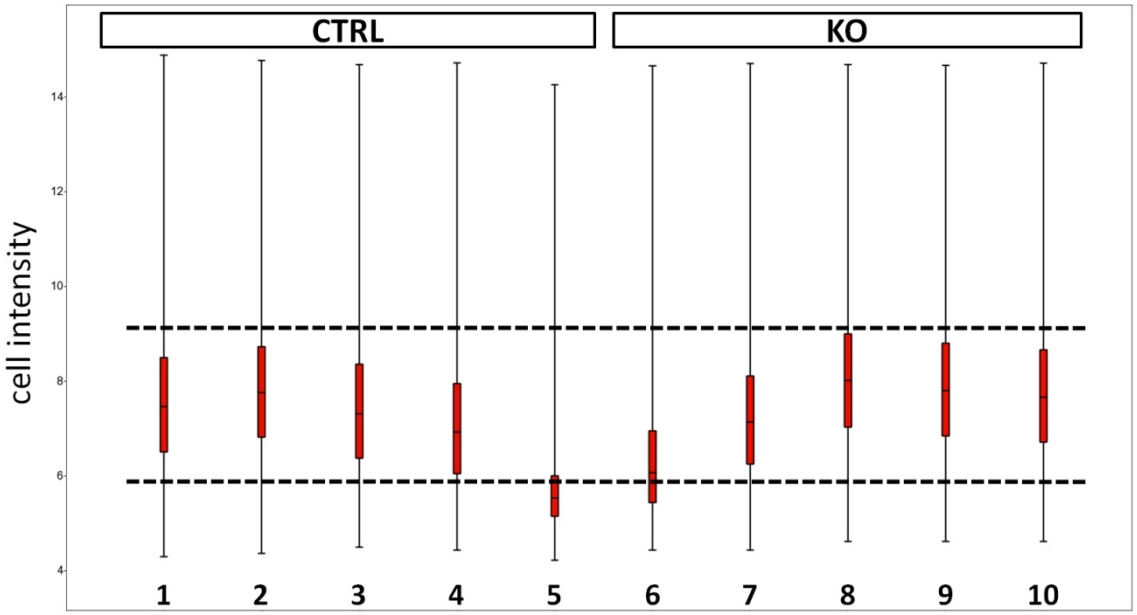

Figure 26 | Comparative analysis of array cell intensities.
The dispersion of cell intensities of 10 exon arrays was compared. 8 out of 10 arrays showed very similar dispersion of intensities. Arrays 5 and 6 displayed median intensities drastically reduced compared to all other arrays surveyed. Data obtained from the arrays 5 and 6 was excluded from analysis. Whiskers indicate the lower and upper quartiles of distribution; the band in the box represents the median.

indicated that 8 of the 10 arrays performed had a similar dispersion of signal intensities. The arrays number 5 and 6 displayed median cell intensities that were drastically lower than intensities of all other arrays. This deviation might have been a result of in part ineffective hybridization or labeling. To prevent that these deviating samples bias the determination of exon abundances in CTRL versus KO group, the arrays number 5 and 6 were excluded from the analysis. Consequently, estimation of changes in exon abundances was based on 4 CTRL versus 4 KO samples.

Data obtained from the exon array was processed in two different ways: First, to analyze changes in gene expression (i.e. changes in the abundance of whole transcripts) and second, to analyze changes in splicing (i.e. changes in the abundance of single exons relative to whole transcripts). To analyze changes in gene expression in general, intensity data of a whole transcript cluster was combined. A transcript cluster contains all exon clusters (i.e. groups of probe sets) belonging to a certain transcript. Thus, global changes of intensity within the whole cluster are indicative of changes in gene expression. To determine changes in splicing, exon clusters were directly compared between the CTRL and KO group. If differences in cell intensity within single exon clusters between CTRL and KO were detected, but were not detected for the whole transcript cluster in general, the probability of exon specific regulation was considered high.

To filter out all hits (expression and splicing analysis) that did not present any difference between the CTRL and KO group only such transcripts or exons that showed at least

a change of ±1.5-fold in abundance were considered. Furthermore, the Student's t-test was applied to analyze the difference between samples of CTRL and KO group which had to yield p-values smaller than 0.05. This initial filtering resulted in a list of 86 differentially expressed transcripts (Appendix, II). 24 transcripts were found to be upregulated and 62 transcripts were identified as being downregulated in KO brains.

Most of the originally identified exons showed no or very little changes in splicing inclusion (Figure 27 A, gray bars). Filtering out exons with less than ±1.5-fold up- or downregulation (gray bars) excluded about 97 % of the identified exons. Further, elimination of exons with significance levels of worse than 0.05 yielded a list of still 1,006 differentially included exons. The majority of 622 identified exons was found to be downregulated in KO mice indicating a positive regulation by Tra2b (Figure 27 B). 384 were upregulated in KO mice. In total most of these differentially regulated exons were changed not more than ±1.5-fold. Only 31 and 81 exons were more than 2-fold up- or downregulated in KO mice.

Among the 1,006 identified exons, multiple exons from one and the same gene were identified as being alternatively spliced (132 exons in total). Genes included in the list of 86 differentially expressed transcripts (Appendix, II) were compared with those genes called with multiple exons. From the list of 132 alternatively spliced exons, 54 matched with the list of 86 differentially expressed genes. A subset of the double-identified transcripts/exons was analyzed including the transcription factor *Eomes* exon 4, Kinsesin-like protein 11 (*Kif11*) exon 12 and Transcription factor 4 (*Tcf4*) exon 18. CTRL and KO samples were analyzed by semi-quantitative RT-PCR using primers that either detect the full length isoform or the isoform in which the exon of interest is skipped. All PCRs were performed as multiplex PCR using *Hprt* as a housekeeping gene. Bands were densitometrically quantified and expression levels of the respective isoforms were normalized to *Hprt* expression. Analysis revealed general changes in expression but no exon-specific changes were detected (Figure 28). This became apparent by equally strong and coordinate downregulation of both isoform levels. The transcription factor *Eomes* (Eomesodermin) was analyzed in detail. Cohorts of CTRL, HET and KO samples were used to quantify levels of total *Eomes,* the full-length and the Δ4 isoform using isoform specific primers. Primer positions for the quantification of total *Eomes* were independent of exon 4. Quantification demonstrated that both isoforms and the total *Eomes* transcript were equally and significantly downregulated in KO mice (Figure 29). This strongly suggests that these exons being identified to be alternatively spliced are true artifacts caused by misinterpretation of altered transcript levels. Thus, the 132 exons belonging to this group were excluded from the candidate list (Figure 27 C).

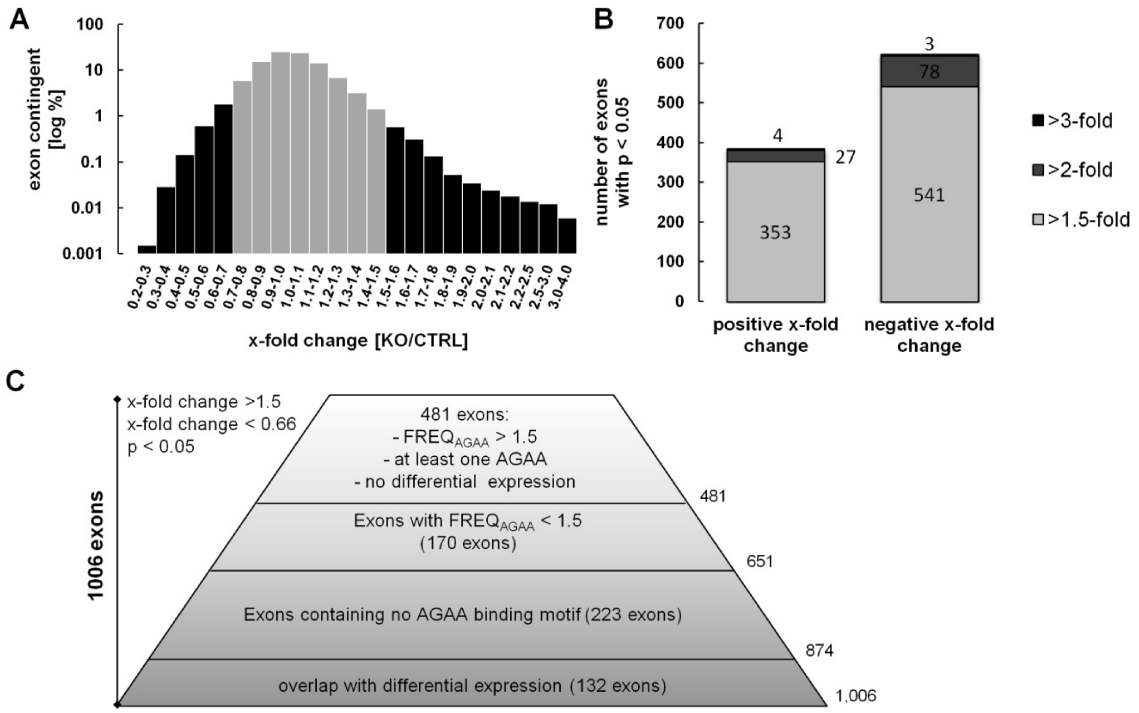

Figure 27 | Statistical analysis and filtering of differentially spliced exons from exon arrays.
(A) The x-fold change of exon inclusion is defined as [$KO_{INCLUSION}$] / [$CTRL_{INCLUSION}$]. Values were distributed from ~0.2 until ~4.0. Initial filtering strategies comprised exclusion of exons with x-fold changes between 0.66 and 1.5 (grey bars) as well as restriction to p-values smaller than 0.05 which yielded a total of 1,006 exons. (B) Distribution of positive and negative inclusion ratios in KO mice. (C) Exons associated with transcripts identified as being transcriptionally up- or downregulated were excluded from analysis. Ranking of those was further refined using large x-fold change values and considering presence of putative Tra2b binding sites (AGAA-motifs). Thereby, exons had to contain at least a single AGAA-site and an AGAA-frequency higher than 1.5.

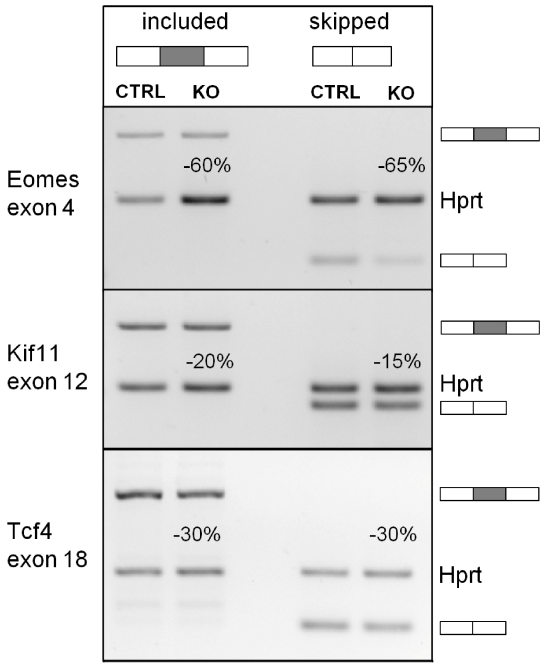

Figure 28 | Analysis of double candidates for alternative splicing and differential expression.
RNA from 16.5 dpc brains with *Tra2b*/*Gapdh* expression lower than 0.4 was analyzed using semi-quantitative RT-PCR. Expression levels of all shown isoforms was measured using isoform-specific primers: FL *Eomes* (4494/4495), Δ4 *Eomes* (4496/4497), FL *Kif11* (4498/4499), Δ12 *Kif11* (4500/4501), FL *Tcf4* (4516/4517), Δ18 *Tcf4* (4526/4518). Expression levels were normalized to *Hprt* (primers 2811/2812).

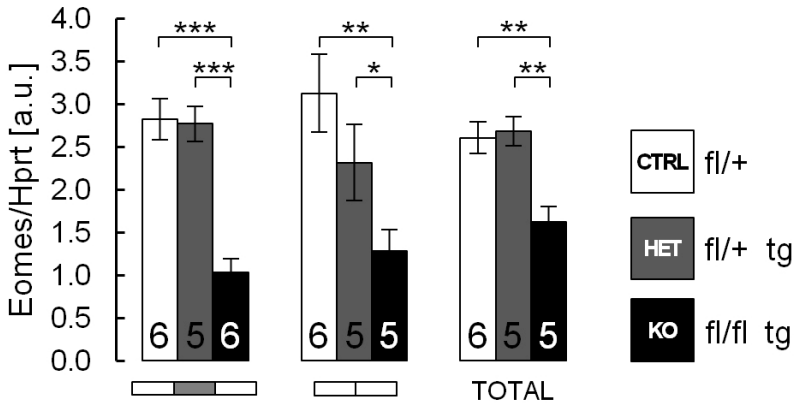

Figure 29 | Quantification of total Eomes and individual isoform levels.
RNA from 16.5 dpc brains was analyzed using semi-quantitative RT-PCR. Expression levels of total *Eomes,* full-length and Δ4 *Eomes* were measured using isoform-specific primers: total *Eomes* (4502/4503), FL *Eomes* (4494/4495), Δ4 *Eomes* (4496/4497). Primers used for the quantification of total *Eomes* were located independently of exon 4. Expression levels were normalized to *Hprt* (primers 2811/2812). Numbers within the bars indicate the number of biological replicates (N).

The overall high number of alternative events and the strong overlap with differentially expressed transcripts is probably indicative of poor data quality and a high error rate which strongly increases the number of false positive exons. It cannot be excluded that the strongly altered brain morphology and misdevelopment in the KO mice and the according changes in cell type and transcriptome composition cause a large number of artifact results. Especially whole transcripts that were found to be downregulated in KO mice are extremely prone to error. Assuming that a transcript is strongly expressed in cortical neurons and is found downregulated in KO mice, it is very likely that the downregulation is a result of absence of many cortical neurons in KO mice but is not a result of real transcriptional regulation.

Furthermore, mosaicism of *Cre*-recombinase expression leads to two distinct cell populations in the KO brain. First, surviving cells in which *Tra2b* was depleted and show altered splicing patterns and, second, cells that express *Tra2b* and show normal splicing. In such a mosaic brain the latter cell population masks the splicing effect present in KO cells. That phenomenon appears to hamper the analytical efforts to reliably call alternative splicing events.

As an attempt to counteract this problem, exons that do not contain any AGAA binding motif were excluded from analysis (Figure 27 C). Considering that such motif occurs statistically once in a 256 bp stretch, the binding site frequency (F) was calculated as the number of binding sites multiplied with the exon length in bp divided by 256. Exons with F_{AGAA} < 1.5 were excluded resulting in 481 remaining exons (Figure 27 C). The remaining 481 exons after applying the described filtering conditions are listed in the Appendix (Appendix, III).

5.6.3. Validation screening of top ranking splicing candidates

Detection of putative alternatively spliced exons on exon arrays required validation by PCR. Even when using simple semi-quantitative PCRs it was due to the sole quantity of candidates not feasible to test all or a big fraction of the 481 remaining candidate exons. Therefore, candidates were individually selected for downstream analysis according to their x-fold change of splicing inclusion and AGAA binding site frequency. Exons with a big change of inclusion in KO mice and a high binding site frequency were preferred. Furthermore, exons that contained binding site clusters were preferred. Additionally, the conservation of the binding sites in mammals was manually surveyed using the UCSC genome browser. As a high conservation of binding sites might point to a conserved biological function of that exon, exons with conserved binding sites or binding site clusters were preferred to others.

As a feasible fast forward screening approach semi-quantitative PCR was performed using isoform specific primer pairs that either detected the isoform including the candidate exon or detected the isoform skipping the candidate exon. Up to 2 forward and 2 reverse primers allowing for 4 possible PCR combinations were used in parallel PCRs. As a template, cDNA was generated from each a representative CTRL and KO RNA. Semi-quantitative PCR was carried out for the screening of 11 candidates matching the properties mentioned above. These were candidate exons of the following transcripts: *Casp1*, *Ezh2*, *Gca*, *Nrg1*, *Prkd3*, *Rcor2*, *Sgol2*, *Slu7*, *Sox6*, *St18* and, *Tubd1*. All these tested candidates and their scores are also highlighted in the candidate list in the Appendix (Appendix, III).

Analysis of these candidates was mostly not able to reproduce the tendencies of exon inclusion detected on the array. In many cases the isoform lacking the exon in question was not detected at all, suggesting that it either does not exist or is expressed at extremely low levels. Figure 30 shows representative examples of the candidate screening. Most PCRs were optimally amplifying the product using multiplex conditions with *Hprt* as the housekeeping gene. Normalization and densitometric quantification showed similar downregulation for both isoforms of the *Slu7* and *Nrg1* transcripts. Coordinate downregulation is indicative of general changes in the amounts of transcripts but rather not of an alternative splicing event. The full length isoforms of the *Prkd3* and *Sox6* transcripts did not demonstrate any clear changes of expression between CTRL and KO. The isoforms of both transcripts lacking the respective candidate exon were not detected at all. However, possible differences in splicing were detected for the Shugoshin-like2 and Tubulin delta chain transcripts (*Sgol2* and *Tubd1* are introduced in chapter 5.6.4). The expression of the *Sgol2* isoform including exon 4 was reduced by approx. 70 % in KO mice compared to CTRL. The isoform skipping exon 4 was not detected in CTRL mice but was detectable in KO mice. This inverse correlation of isoform expression in

CTRL and KO mice strongly suggests that *Sgol2* exon 4 splicing is a Tra2b-dependent splicing process. This appeared to be equally true for the splicing of *Tubd1* exon 4. Expression of the full-length isoform including exon 4 was reduced by approx. 20 % in KO mice compared to CTRL mice. The isoform lacking exon 4 was not detected in either CTRL or KO mice. Computational analysis showed that removal of exon 4 from the *Tubd1* transcript causes a frame-shift and generates multiple premature termination codons. Therefore, the *Tubd1* transcript lacking exon 4 probably undergoes nonsense-mediated decay and might not be detectable for that reason. Of the 11 initially surveyed putative splicing events only the exons 4 of the *Sgol2* and *Tubd1* transcripts showed indications for alterations in splicing based on the absence of Tra2b. Other tested candidates displayed either coordinate up- or downregulation of both isoforms, which could be an indicator of transcriptional regulation but is most likely an artifact result based on the severe changes in brain morphology and tissue composition as it is present in the brains of KO mice. This idea is further supported by the fact that in the absolute majority of cases a coordinate downregulation but not an upregulation was observed. Failure in detection of isoforms lacking the respective candidate exon in both CTRL and KO mice suggests that this isoform is not expressed and is likely not a physiological isoform – at least not in brain. The sudden appearance of the shorter isoform in KO mice – as it is the case for *Sgol2* exon 4 – indicates that inclusion of this exon at least in part might depend on Tra2b.

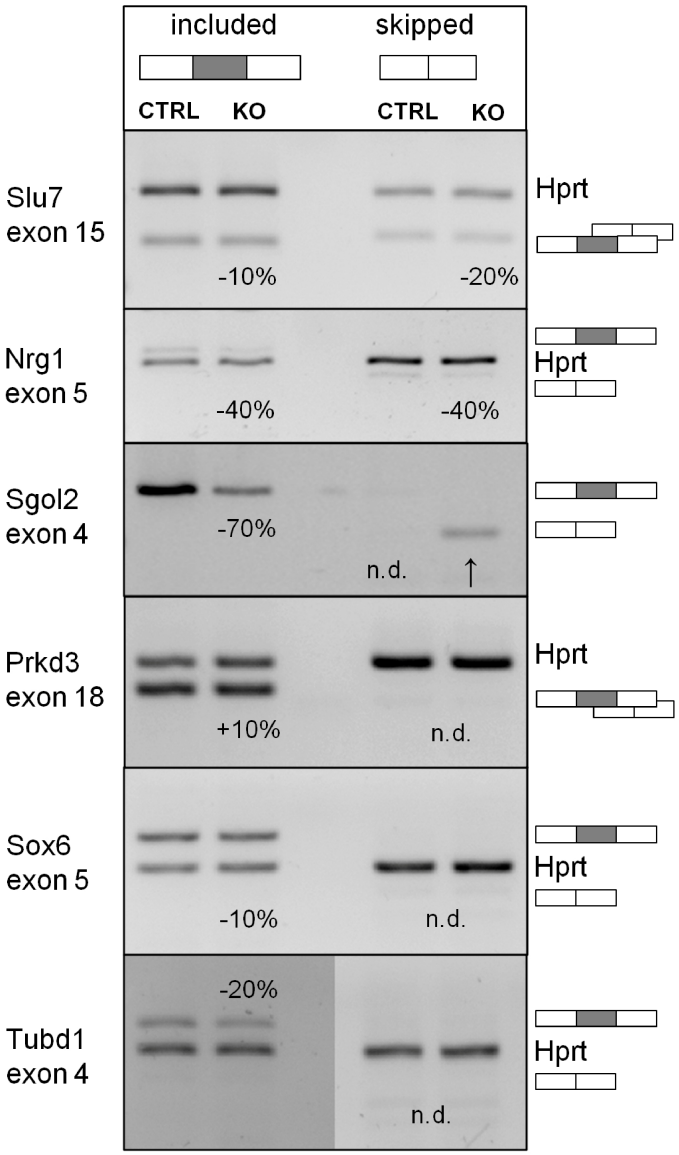

Figure 30 | Fast forward screening of candidate exons for alternative splicing.
RNA from a representative 16.5 dpc brain with *Tra2b*/*Gapdh* expression lower than 0.4 was analyzed and compared to CTRL using semi-quantitative RT-PCR. Gel images show PCR results of the most effective primer combination used. Isoform-specific PCRs were carried out in parallel using the primers: FL *Slu7* (4512/4513), Δ15 *Slu7* (4522/4523), FL *Nrg1* (4504/4505), Δ5 *Nrg1* (4506/4507), FL *Sgol2* (4510/4511), Δ4 *Sgol2* (4510/4521), FL *Prkd3* (4577/4578), Δ18 *Prkd3* (4580/4579), FL *Sox6* (4514/4515), Δ5 *Sox6* (4524/4525), FL *Tubd1* (4529/4531), Δ4 *Tubd1* (4532/4530). Expression levels were normalized to *Hprt* if multiplex PCR was applicable (primers 2811/2812).

5.6.4. Validation of positive splicing candidates

Mouse exon array analysis and candidate screening identified exons of the *Sgol2* and *Tubd1* transcripts as likely targets of Tra2b, as both exons showed differential inclusion ratios in CTRL versus KO mice. *Sgol2* encodes the Shugoshin-like 2 protein that protects the integrity of the cohesin-complex, which regulates sister chromatid cohesion during meiosis I (Llano et al. 2008; Clift and Marston 2011). It has been demonstrated to be essential for meiosis but not for mitosis and *Sgol2*-deficient mice are viable but infertile. The *Tubd1* gene encodes the tubulin delta 1 protein. It is highly expressed in testis and is a component of the perinuclear ring of the manchette, which helps to translocate and elongate the nucleus during sperm cell maturation (Smrzka et al. 2000; Kato et al. 2004). The somatic function of delta tubulin is poorly understood and expression levels in brain are rather low. Still it has been detected on exon arrays using brain RNA and the transcript was identified in subsequent screening PCRs.

To validate splicing of the respective candidate exons of the *Sgol2* and *Tubd1* transcripts *in vivo*, semi-quantitative RT-PCRs were performed using cohorts of CTRL, HET and KO whole brain RNAs for a comparative analysis. As before, primers were selected to specifically amplify either the isoform including or skipping the respective exon. All PCRs were performed as multiplex PCRs using *Hprt* as a housekeeping gene. Band intensities were densitometrically determined and expression levels were normalized to *Hprt* expression.

The isoform of *Sgol2* including exon 4 showed a modest and non-significant decrease in KO and HET mice compared to controls (Figure 31 A,C). This is in contrast to the results shown in Figure 30 in which full-length *Sgol2* appeared to be drastically downregulated in KO mice. However, this result was based on a single KO sample compared to a CTRL sample and was therefore statistically non-significant. Quantification of the *Sgol2* Δ4 isoform showed weak but consistent expression in CTRL and HET mice (Figure 31 B,D). Strikingly, *Sgol2* Δ4 expression was significantly increased by 2.75-fold in KO animals. This quantification fully resembles the results obtained in exon array analysis, where a 2.7-fold loss of exon inclusion was detected in KO mice (Appendix, III). These data demonstrate that the inclusion of exon 4 into the *Sgol2* transcript is promoted by Tra2b. The mild decrease of full-length *Sgol2* transcripts might be a consequence of *Sgol2* Δ4 upregulation. Very likely the decrease of FL *Sgol2* expression is minimal, because - regardless of the presence of Tra2b – FL *Sgol2* is the major isoform and even drastic changes of the inferior Δ4 isoform do not strongly affect the prominent FL isoform. Thus, *in vivo* validation of *Sgol2* exon 4 splicing constitutes a proof of concept for this *in vivo* candidate approach to detect Tra2b-related splicing processes under physiological conditions.

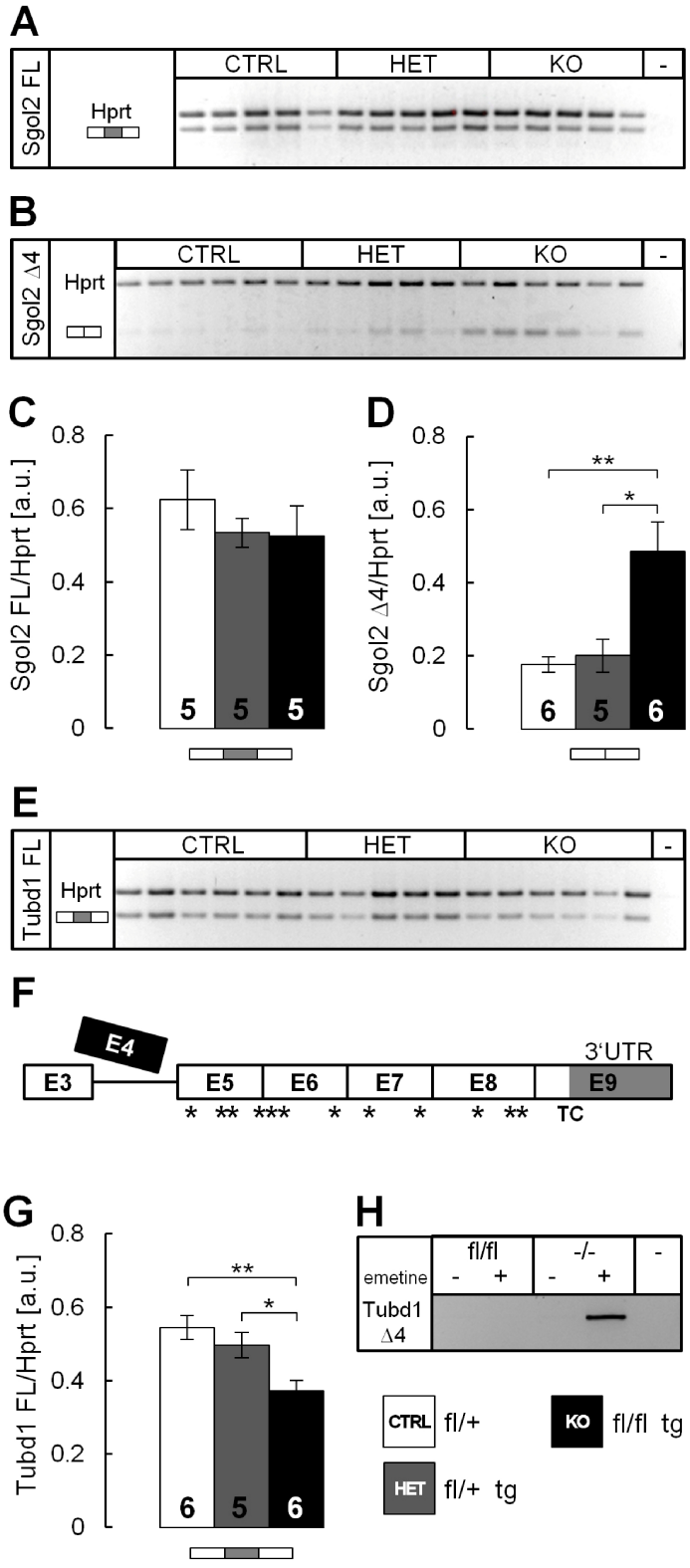

Figure 31 | Validation of *Sgol2* exon 4 and *Tubd1* exon 4 splicing.
RNA from 16.5 dpc brains with *Tra2b*/*Gapdh* expression lower than 0.4 was analyzed using semi-quantitative RT-PCR. Expression levels of all shown isoforms was measured using isoform-specific primers: (A,C) FL *Sgol2* (4510/4511), (B,D) Δ4 *Sgol2* (4510/4521), (E,G) FL *Tubd1* (4529/4531), (F) Δ4 *Tubd1* (4532/4530). Δ4 *Tubd1* was not detectable *in vivo*. Skipping of exon 4 causes a frame-shift generating numerous premature termination codons. (H) Analysis of *Tubd1* Δ4 expression in emetine treated fibroblasts that either express or do not express *Tra2b*. FL, full length; Δ4, transcript lacking exon 4, (-) PCR negative control; numbers within the bars indicate the number of biological replicates (N), asterisks indicate premature termination codons.

Tubd1 exon 4 was further demonstrated to be dependent on Tra2b by semi-quantitative RT-PCR (Figure 31 E,F). The *Tubd1* isoform including exon 4 showed a significant decrease by 1.46-fold in the brains of KO animals compared to CTRL. Expression of the full-length isoform in CTRL and HET animals was at similar levels and non-significantly different. The reduction of full-length isoform expression in KO mice verifies the 1.7-fold loss of exon inclusion detected on the microarray (Appendix, III). Similar to the initial candidate screening approach, the ∆4 isoform of *Tubd1* was not detectable in brains of either CTRL or KO mice. *Tubd1* exon 4 is 217 nucleotides long and thereby comprises 72 nucleotide triplets plus a single nucleotide. Computational analysis showed that skipping of exon 4 from the *Tubd1* transcript causes a frame-shift that results in at least 10 premature termination codons that would likely lead to transcript degradation by nonsense-mediated decay (Figure 31 F). This hypothesis was tested using previously generated murine immortalized fibroblasts that were either depleted of *Tra2b* or expressed *Tra2b*. The generation and characterization of these cell lines is described in 5.2. Cells were treated with emetine to block NMD in *Tra2b*-depleted and *Tra2b*-expressing cells (as described in 4.1.6). RNA isolated from treated and non-treated fibroblasts was used to detect the *Tubd1* Δ4 isoform using isoform specific primers. Strikingly, the *Tubd1* Δ4 isoform was exclusively detectable in *Tra2b*-depleted cells that were treated with emetine (Figure 31 H). *Tra2b*-depleted cells that were not treated with emetine did not show expression of the Δ4 isoform, strongly suggesting that the *Tubd1* Δ4 transcript is highly effectively degraded via the NMD-mechanism. Furthermore, fibroblasts expressing *Tra2b* (*Tra2b*$^{fl/fl}$) did not allow detection of *Tubd1* Δ4 neither without, nor with emetine treatment. This clearly indicates that ablation of *Tra2b* is directly causative for the generation of this isoform. Moreover, *Tubd1* exon 4 inclusion can obviously not be fully compensated by other factors, an idea which is supported by reduced full-length *Tubd1* levels in KO mice (Figure 31 G).

Which functional or physiological implications the observed Tra2b-dependent splicing processes have *in vivo* cannot be answered with certainty. The question whether the *Tubd1* Δ4 isoform is expressed in presence of Tra2b under physiological conditions cannot be resolved as the inability to abolish NMD *in vivo* precludes this analysis. *In vivo* splicing regulation of *Tubd1* exon 4 could constitute a mechanism to regulate the amount of Tubd1 protein levels on a post-transcriptional level. Whether and which functional role the skipping of exon 4 from the *Sgol2* transcript has and which consequences this has for the protein function and stability remains to be determined.

To summarize, splicing processes of two candidate exons of the *Sgol2* and *Tubd1* transcripts identified on mouse exon arrays were successfully verified *in vivo*. This proofs the

neuronal-specific *Tra2b* knock-out mouse to be a functional model to study Tra2b-related splicing processes *in vivo* under physiological conditions. High throughput analyses have also uncovered weaknesses of the system. These will be discussed in detail in section 6.3.

5.7. Functional splicing analyses

The Shugoshin-like 2 exon 4 and the Tubulin delta 1 exon 4 were shown to be spliced in dependency of Tra2b in the murine developing brain. Both were identified on exon arrays and were validated by RT-PCR using RNA from murine brains. For the identification of splicing targets most previous studies have been focusing on approaches involving artificial or biased systems like cross-linking of RNA-protein interactions, overexpression of splice factors and minigene experiments (Stamm et al. 1999; Hofmann et al. 2000; Kondo et al. 2004; Stoilov et al. 2004; Grellscheid et al. 2011a; Grellscheid et al. 2011b). The identification of the *Sgol2* and *Tubd1* exons 4 using murine whole brain RNA is a proof of concept for the *in vivo* identification of novel splicing relationships.

Though the *in vivo* analysis allowed detection of these splicing events on a physiological background, it cannot be excluded that results obtained from the neuronal-specific knock-out system are inertly biased by massive changes in brain physiology. It was demonstrated that *Tra2b*-depleted brains develop a very severe phenotype including ventriculomegaly and loss of cortical cells and patterning. Thus the final ratio of cell types is different between CTRL and KO brains. This is in so far important as different cell types with diverse function might have different patterns of global gene expression but also different patterns in alternative splicing of transcripts.

To assure that the splicing effects of *Sgol2* exon 4 and *Tubd1* exon4 are truly Tra2b-related but not a product of a shift in cell type composition, minigene experiments were performed. Even though minigene approaches are non-physiological with regard to splice factor concentrations and substrate mRNA concentrations, experiments are performed in a controlled environment that allows direct comparison. The human and murine versions of *Sgol2* exon 4 and *Tubd1* exon 4 were tested for splicing inclusion as a response to increased or abolished *Tra2b* expression in HEK293T cells using the pSPL3 exon trapping system as described in chapter 5.7.1.

5.7.1. Generation of Sgol2 and Tubd1 minigene constructs

For testing of *Sgol2* exon 4 and *Tubd1* exon 4 in minigene experiments the human and murine versions of the exons with approx. 0.5 – 1 kb flanking intronic regions were cloned into the pSPL3 exon trapping vector. Genomic regions were amplified from human and murine DNA using restriction-overhang primers as listed in 3.9.1. The PCR amplified insert was cloned to the multiple cloning site of pSPL3 using 5'XhoI and 3'BamHI restriction ligation. The whole genomic insert was validated by sequencing. In total 4 distinct minigene constructs were generated (Figure 32 B): the human and murine versions of each *Sgol2* exon 4 and *Tubd1* exon 4. Their genomic inserts are somewhat variable in size as to the length of their flanking introns. These were chosen to be shorter for the Tubd1 genomic insert due to a GC-rich genomic vicinity which made specific amplification by PCR laborious. Once transfected into eukaryotic cells the pSPL3 vector produces mRNA containing the genomic insert with the respective exon of interest flaked by the *tat a* and *tat b* exons of HIV. These flanking exons are constitutively spliced and have the potential to link with foreign exons resulting in a specific mRNA (Buckler et al. 1991). Inclusion of the exon of interest results accordingly in a longer mRNA that can be detected and compared by size differences using RT-PCR (Figure 32 A).

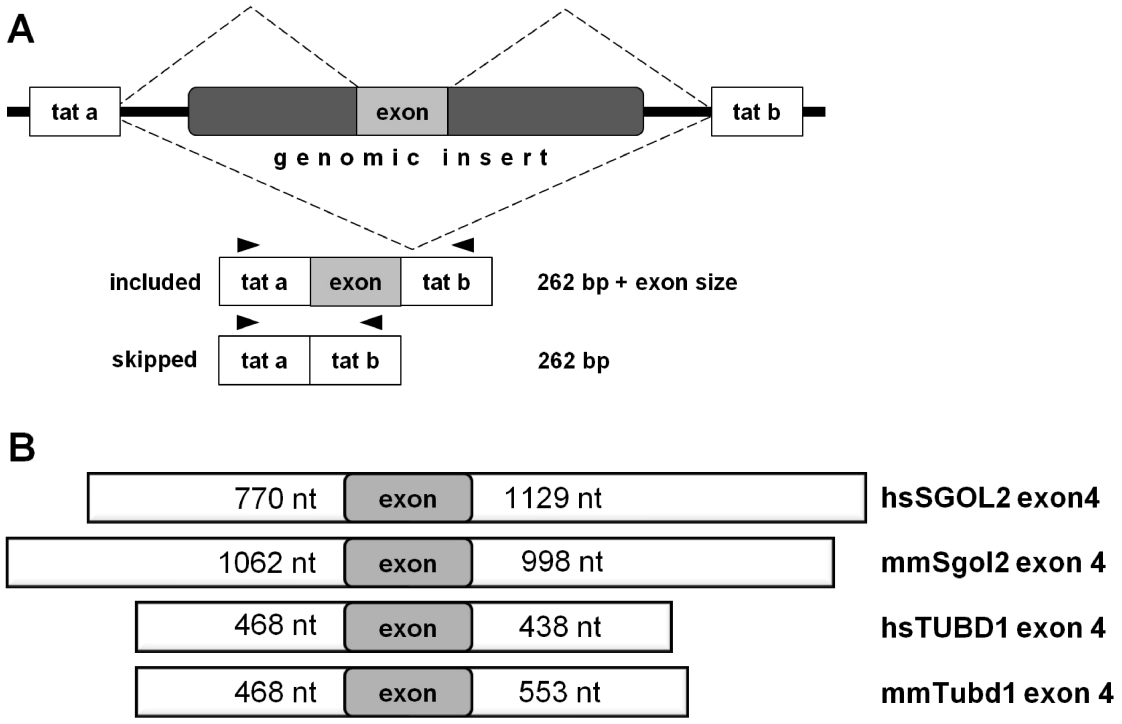

Figure 32 | The pSPL3 exon trapping system: Sgol2 and Tubd1 minigene constructs.
(A) Functional structure of transcripts generated from the pSPL3 vector. Tat a and tat b are constitutive exons of HIV that form a 262 bp transcript. The genomic insert containing the exon of interest was cloned in between them. Depending on exon recognition and the used conditions the exon of interest will be spliced or skipped. The transcript can be detected using specific primers within the tat a and tat b exon producing a 262 bp product. Inclusion of the exon of interest results in an accordingly longer product. (B) Genomic inserts containing human and murine versions of *Sgol2* exon 4 and *Tubd1* exon 4. Exons are surrounded by their flanking intronic sequences. The length of the intronic sequences is given.

5.7.2. Minigene splicing assay

The minigene splicing assays using the previously described pSPL3 exon trapping constructs were carried out in HEK293T cells. HEK293T cells were transfected using the DharmaFECT I transfection reagent to deliver plasmid DNA, siRNA or both. Splicing of the human and murine versions of *Sgol2* exon 4 and *Tubd1* exon 4 was tested while overexpressing *HTRA2B* from pEGFP or depleting *HTRA2B* using siRNAs in parallel to transfection with a minigene in pSPL3. As control experiments proteins were isolated from HEK293T cells. Immunological detection of TRA2B indicated that TRA2B expression is effectively abolished upon siRNA treatment and the TRA2B-GFP fusion-protein is strongly expressed upon pEGFP-TRA2B transfection (Figure 33 A). RNA was isolated 48 h after transfection and pSPL3-derived transcripts were detected by semi-quantitative RT-PCR using the tat a and tat b specific primers 4508 and 4509. The typical transcript product of 262 bp was specifically detected. PCR products with the *Sgol2* or *Tubd1* exon included were accordingly longer (Figure 33 B).

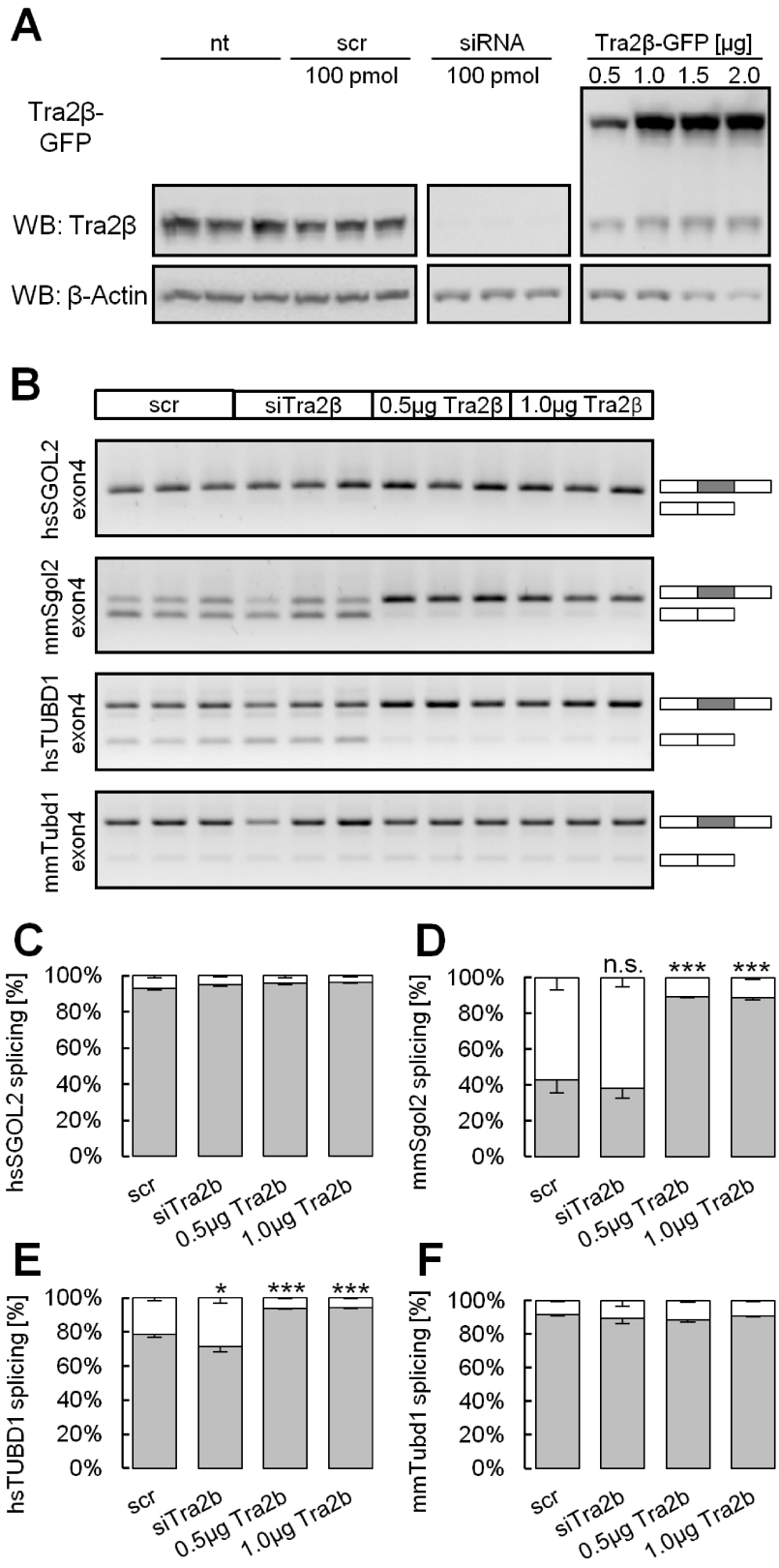

Figure 33 | Minigene splicing assay of *Sgol2* exon 4 and *Tubd1* exon 4.
HEK293T cells were co-transfected with the pSPL3 minigene vector and siRNA specific for *Tra2b* or a *TRA2B-GFP* expression vector. (A) Western Blot analysis shows efficiently reduced TRA2B protein levels and solid overexpression of TRA2B-GFP. (B) RNA was analyzed for exon inclusion after 48h by semi-quantitative RT-PCR using the primers 4508/4509. (C-F) RT-PCR results were densitometrically quantified. Exon 4 of murine but not human *Sgol2* was responsive to increased concentrations of TRA2B as splicing inclusion significantly increased from 43 % to 89 %. Knock-down of *TRA2B* was insufficient to reduce *Sgol2* exon4 splicing inclusion. Exon 4 of human but not murine *Tubd1* was responsive to increased concentrations of TRA2B as splicing inclusion significantly increased from 79 % to 94 %. Knock-down of *TRA2B* decreased inclusion of exon 4 from 79 % to 71 %. nt, non-treated; scr, scrambled siRNA; si, siRNA against *Tra2b*; n.s., non significant; all transfections were performed in triplicates

PCR results indicate that only the murine *Sgol2* but not the human *SGOL2* is responsive to *HTRA2B* overexpression in HEK293T cells (Figure 33 C,D). Under normal conditions (scr) *Sgol2* exon 4 was included in 43 % of the transcripts. After overexpression of *HTRA2B* exon 4 was included in 89 % of the transcripts suggesting that *HTRA2B* effectively promotes inclusion of murine *Sgol2* exon 4 (Figure 33 D). On the contrary, *HTRA2B* overexpression had no effect on the human *SGOL2* transcripts (Figure 33 C). The human exon 4 is included in 93 % of the transcripts under normal conditions (scr) (Figure 33 C). This suggests that *HTRA2B* overexpression is unable to further enhance *SGOL2* exon 4 inclusion. As human embryonic kidney cells (HEK293T) were used for minigene splicing assays the intrinsic splice site strengths of the tested exons was analyzed by comparison to the primate splice site consensus (http://ibis.tau.ac.il/ssat/SpliceSiteFrame.htm). Splicing machineries of rodent organisms might process exons differently than in primate organisms with regard to their optimal consensus sequence of the major spliceosome. Indeed, scores of both the 3' and 5' splice sites of *Sgol2* exon 4 were higher for the human sequence compared to the murine version (Figure 34). Furthermore, ablation of endogenous *HTRA2B* in HEK293T cells by RNAi did not decrease

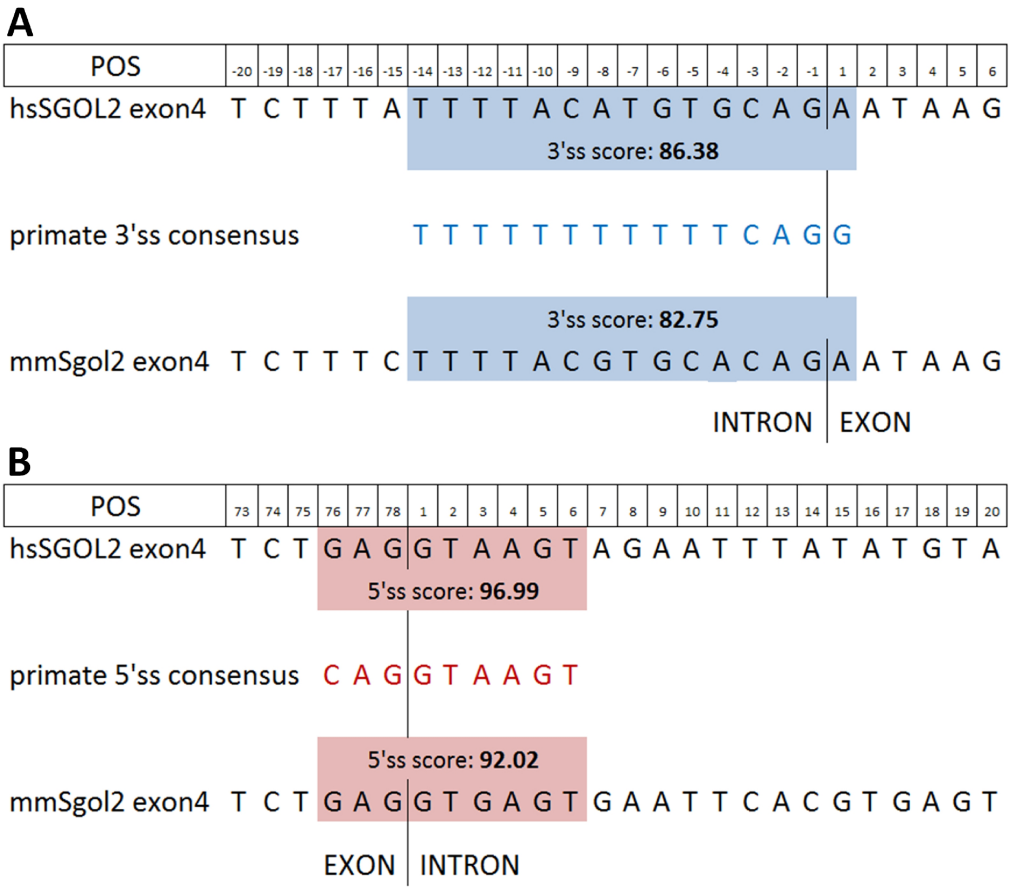

Figure 34 | Comparative splice site analysis of human and murine *Sgol2* exon 4.
3' and 5' splice sites of human and murine *Sgol2* exon 4 were rated by comparing to the primate consensus sequence using the AnalyzerSpliceTool available online (http://ibis.tau.ac.il/ssat/SpliceSiteFrame.htm). (A) Alignment of 3' splice sites. The relevant nucleotide stretch from -14 (intonic) to +1 (exonic) is labeled blue. (B) Alignment of 5' splice sites. The relevant nucleotide stretch includes the last 3 exonic nucleotides and the first 6 intronic nucleotides (labeled red). Both the 3' and 5' splice site of human *SGOL2* achieve higher scores than the murine version.

exon 4 inclusion (Figure 33 C,D). These results suggest that the high intrinsic splice site strength of the human *SGOL2* exon 4 is sufficient to constitutively splice that exon. Therefore, results obtained for *Sgol2* in the mouse might probably not be transferable to humans.

The alternative splicing of *Tubd1* exon 4 in the minigene is *HTRA2B* dose-dependent (Figure 33 E,F). Under normal conditions (scr) 79 % of transcripts show exon 4 inclusion. Ablation of *HTRA2B* slightly but significantly reduced exon 4 inclusion to 71 % while overexpression significantly increased inclusion to 94 % (Figure 33 E). This indicates that *HTRA2B* promotes *TUBD1* exon 4 inclusion, confirming the results obtained in the mouse model. Surprisingly, the murine version of the minigene failed to respond to altered *HTRA2B* concentrations in HEK293T cells (Figure 33 F). Under normal conditions (scr) exon 4 was included in 92 % of the transcripts and splicing inclusion did not change upon knock-down or overexpression of *HTRA2B*. This finding is in so far intriguing as the *Tubd1* exon 4 splicing event was originally detected in mouse (Figure 31 E-H). The splice site strengths of human and murine *Tubd1* exon 4 are very similar and analysis suggests that the murine 3' and 5' splice sites are presumably equally well recognized by the human splicing machinery in HEK293T cells. However, there are minor differences in number and position of well-known binding sites (e.g. SF2/ASF and SRp40) that might be attributable to a more efficient inclusion of exon 4 in a human cellular system compared to native conditions in the mouse.

Analysis of *Sgol2* exon 4 and *Tubd1* exon4 splicing using minigene analyses could in part confirm both exons being responsive to different concentrations of TRA2B. Responsiveness to HTRA2B overexpression clearly demonstrated that both exons are regulated by TRA2B. However, it appears like TRA2B exerts a rather modulating effect on the inclusion of *Sgol2* exon 4 and *Tubd1* exon 4. A basic and often solid recognition of both exons was even detectable in absence of TRA2B, suggesting that these exons are well recognized by the general splicing machinery and / or other factors. The human *TUBD1* minigene was the only construct that was responsive to reduced TRA2B concentrations and accordingly showed less exon inclusion. In how far the result obtained *in vivo* using neuronal-specific knock-out mice can be compared to results from minigene experiments is difficult to determine. HEK293T cells are of human origin and their spliceosome matches best the primate consensus sequence. This consensus sequence in rodents is different, however. For that reason, responsiveness of a minigene construct in human cells might not be fully transferable to a splicing effect observed in mice.

5.8. Functional analysis of Tra2b-mediated p21-upregulation

In frame of exon array analysis of murine brain RNA, total expression levels of genes were studied besides performing analyses regarding differential exon inclusion. As underlined in previous chapters severe morphological changes and alterations in cell type composition take place in brains of neuronal-specific knock-out mice. This impacts the analysis of splicing events and the determination of transcript levels in general (see 5.7). The identification of decreased transcript levels might be a direct result of tissue loss in the brain and might therefore not reflect a real transcriptional downregulation. However, the identification of increased transcript levels is more likely to reflect real and solid transcriptional upregulation, as the increase is counteracted but not supported by the general loss of transcripts caused by tissue loss. Accordingly, many more transcripts have been significantly identified as being downregulated than being upregulated (Appendix, II).

Among the candidates with increased expression levels the cyclin-dependent kinase inhibitor 1A (*Cdkn1a),* also known as p21, attracted most attention for two reasons. First, p21 is an important regulator of cell cycle progression that can inhibit the activity of the cyclin-dependent kinases CDK2 and CDK4, thereby blocking cell-cycle progression (Harper et al. 1993). This was in so far intriguing as massive apoptosis and a loss of the proliferative potential was detected in *Tra2b*-depleted murine brains. As CDK inhibition by p21 leads to cell-cycle arrest at the G_1-S-phase transition (Harper et al. 1993) and inability to proliferate might cause apoptosis in parts of the developing brain, expressional upregulation of p21 might be causative for apoptosis in *Tra2b*-depleted murine brains.

Second, the histone chaperone *Nasp* was shown to be spliced by Tra2b in mouse testis and the developing brain (Grellscheid et al. 2011a) (Figure 18 A). Ablation of *Tra2b* decreased the inclusion rate of the *Nasp* T-exon which resulted in a decreased expression of the *tNasp* isoform (Figure 18 A). Ablation of *Nasp* has been shown to cause early embryonic lethality in mice (Richardson et al. 2006). Astonishingly, selective ablation of *tNasp* via RNAi in tumor cells has previously been demonstrated to induce apoptosis via expressional upregulation of *p21* (Alekseev et al. 2011).

Apoptosis in the *Tra2b*-depleted brains, early embryonic lethality in mice ubiquitously depleted of *Tra2b* (Mende et al. 2010), aberrant splicing of *tNasp* in *Tra2b*-depleted brains, and *tNasp* ablation causing p21-mediated apoptosis in tumor cells (Alekseev et al. 2011) possibly form a common ground to explain *p21* upregulation and apoptosis observed in neuronal-specific knock-out mice. Therefore, *p21* upregulation identified on exon arrays

became the major subject of further investigations. First experiments included the validation of p21 upregulation in brains of neuronal-specific knock-out mice.

5.8.1. p21 is upregulated in brains of conditional knock-out mice

To validate the expressional upregulation of *p21* found on exon arrays semi-quantitative RT-PCR was carried out specifically detecting *p21* transcripts in whole brain RNA. Cohorts of CTRL, HET and KO mice were compared and expression data were normalized to *Hprt* expression levels. In agreement to the 1.5-fold upregulation of *p21* detected on exon arrays (Appendix, II), *p21* was found significantly upregulated by 1.4-fold in KO mice compared to CTRL (Figure 35). HET mice did not show any changes in *p21* expression levels compared to controls. Due to *Cre* recombinase-related mosaicism and the spatial limitations of apoptotic areas in the brain of KO mice (Figure 14) it was expected that apoptotic cells likely contribute only with a small fraction to the whole brain. Still the detection of *p21* upregulation was possible in whole brain RNA indicating a solid upregulation in the affected cell populations.

Expressional upregulation of *p21* likely occurred in brain regions that showed increased Caspase 3 levels indicative of apoptosis. These regions were primarily the ventricular and subventricular zone of the cortex (Figure 14), which are neurogenic areas and origin of neuronal outgrowth to form different cortical layers (Lopez-Bendito and Molnar 2003). P21-mediated cell-cycle arrest in this population of neuronal precursor cells could very likely

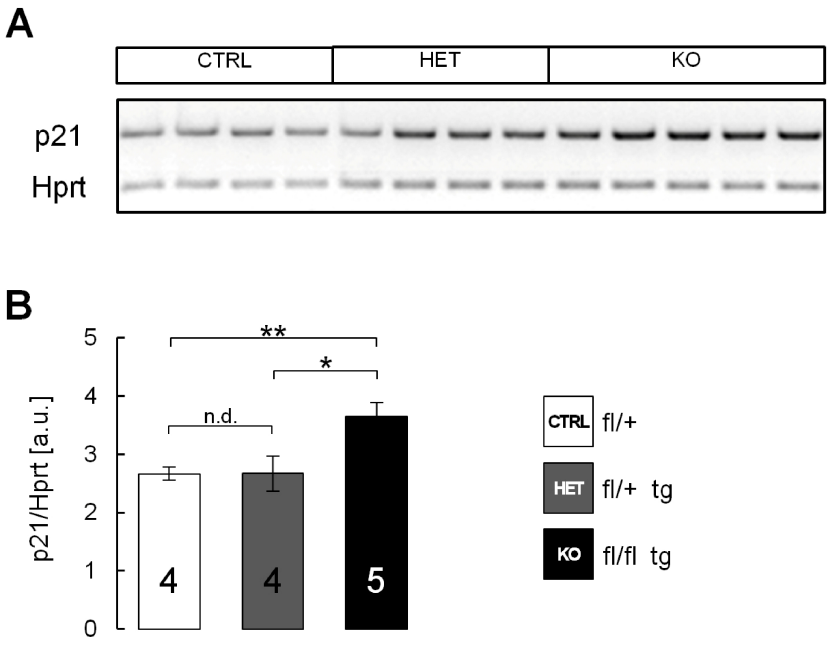

Figure 35 | Expression analysis of p21 in neuronal specific knock-out mice.
(A) RNA from 16.5 dpc brains was analyzed using semi-quantitative RT-PCR. Expression levels of p21 were measured using specific primers (4491/4492). *P21* expression was normalized to *Hprt* (primers 2811/2812). (B) *P21* levels were found 1.4-fold significantly upregulated in KO mice compared to CTRL and HET mice. Numbers within the bars indicate the number of biological replicates (N).

abolish cell differentiation and heavily disturb cortical development. To test whether *p21* upregulation is a direct result of *Tra2b* depletion, NSC34 cells were used to analyze the effects of *Tra2b* depletion in a homogeneous cellular system. The use of this neuronal precursor cell line further allowed assessing the cellular response to *p21* upregulation.

5.8.2. Tra2b knock-down experiments in NSC34 cells

To test the consequences of *Tra2b*-depletion in a homogeneous cellular system murine neuronal stem cells (NSC34 cells) were used, as their neuronal fate resembles the neuronal identity of affected cells found in the ventricular and subventricular cortical layers of KO mice. NSC34 cells were depleted of *Tra2b* using specific siRNAs. RNA and protein levels of Tra2b and p21 were determined by quantitative real-time PCR and immunoblot. RNA levels of the functional *Tra2b* Δ2 isoform (Figure 37 A) and Tra2b protein levels (Figure 36) were markedly reduced within 24, 48 and 72 hours after siRNA transfection. Transfection with a scrambled control siRNA did not cause significant changes of the Tra2b protein level (Figure 36). As expected, depletion of *Tra2b* caused changes in splicing of the *Nasp* transcript. The *tNasp* isoform was significantly downregulated by at least 50 % within 24 to 48 hours after transfection (Figure 37 B) suggesting that Tra2b activity is functionally abolished upon siRNA transfection. To test whether *p21* elevation found in KO mice (Figure 35) can be reconstituted in NSC34 cells, *p21* expression was measured via real-time PCR using the specific primers 4493 and 4492. The p21 expression on protein level was determined via Western Blot. Within 48 hours after transfection, p21 did not increase on protein level (Figure 36). However, a modest but significant increase of *p21* expression was detected on mRNA level (Figure 37 C). To asses long term effects of *Tra2b* ablation on *p21* expression RNA levels were assayed 72 and 96 hours after transfection. Notably, after 72 hours about 20 % of the *Tra2b*-depleted cells had died (data not shown) and *p21* was strongly and highly significantly upregulated by +2.2-fold on mRNA and protein level (Figure 37 C, Figure 36). Strikingly, after 96 hours the majority of transfected cells had died indicating a massive loss of proliferative potential and possible cell cycle arrest and apoptosis. The succession of events suggested that the downstream effects originally triggered by *Tra2b* ablation do not occur instantly but rather require time to establish. This is in line with the findings in KO mice, in which onset of *Cre*-recombinase expression and thus *Tra2b* depletion starts at around 10.5 dpc. Abnormalities in brain development cannot be observed before 14.5 to 15.5 dpc, however. These findings strongly indicate that *Tra2b* deficiency leads to *p21* activation and suggest *p21* activation to be a major contributor to apoptosis and developmental defects observed in neuronal-specific *Tra2b* KO mice.

As depletion of Tra2b reduced the inclusion rate of the *Nasp* T-exon, the expression of the *tNasp* isoform was accordingly reduced in the performed experiment (Figure 37 B). Like mentioned previously, p21 has been demonstrated to be upregulated in tumor cells upon selective depletion of *tNasp (Alekseev et al. 2011)*; a situation that is very similar to the experiment performed in this study. One could argue that *Tra2b*-depletion mediates *p21* upregulation via depletion or downregulation of *tNasp*. This indirect signal transmission finally leading to *p21* upregulation includes the probably time-consuming degradation of tNasp protein, which fits to the delayed cellular response in NSC34 cells. To approach functional testing of this hypothesis, follow-up experiments were aiming to rescue *p21* expression to normal levels by overexpressing *tNasp* from a vector while depleting *Tra2b* using siRNAs.

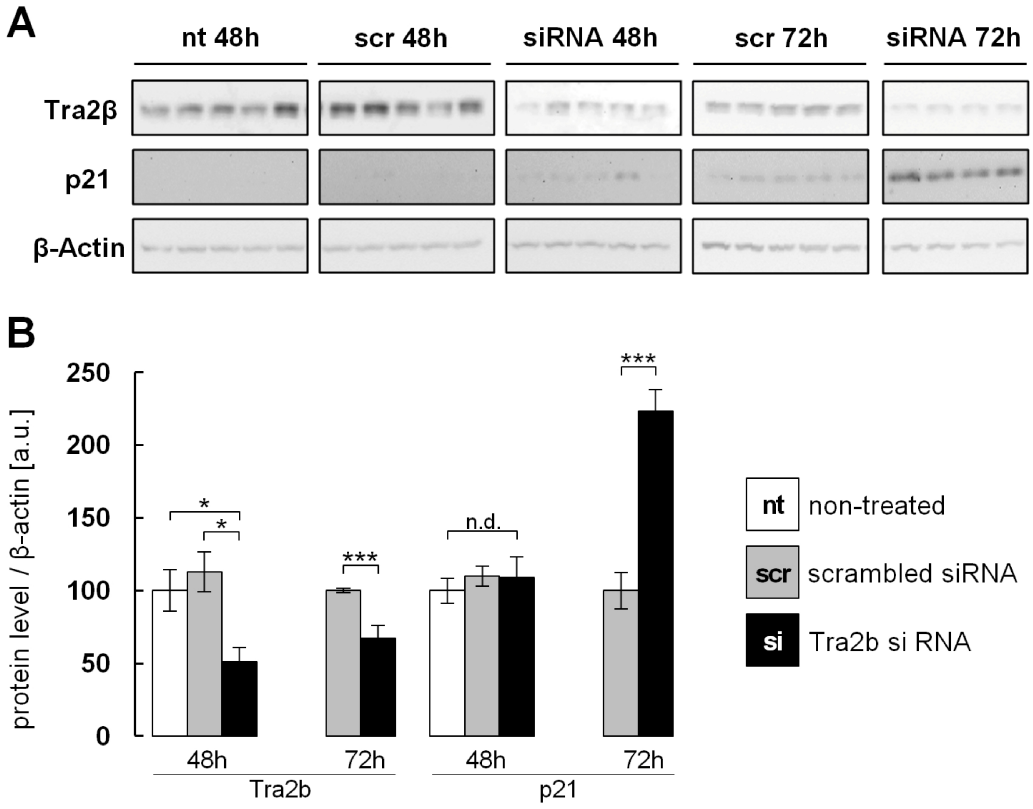

Figure 36 | Tra2b depletion in NSC34 cells: Analysis on protein level.
NSC34 neural stem cells were transfected with siRNAs specific to *Tra2b* or scrambled siRNAs using the DharmaFect I transfection reagent. (A) Immunoblots of NSC34 cell lysates at the indicated time points. (B) Densitometric analysis of immunoblot band intensities. All data were normalized to β-actin expression. siRNA treatment but not scr-treatment effectively reduced Tra2b protein levels 48h and 72h after transfection. P21 protein levels are unchanged after 48h but are significantly increased by 2.2-fold in *Tra2b*-depleted cells 72h after transfection.

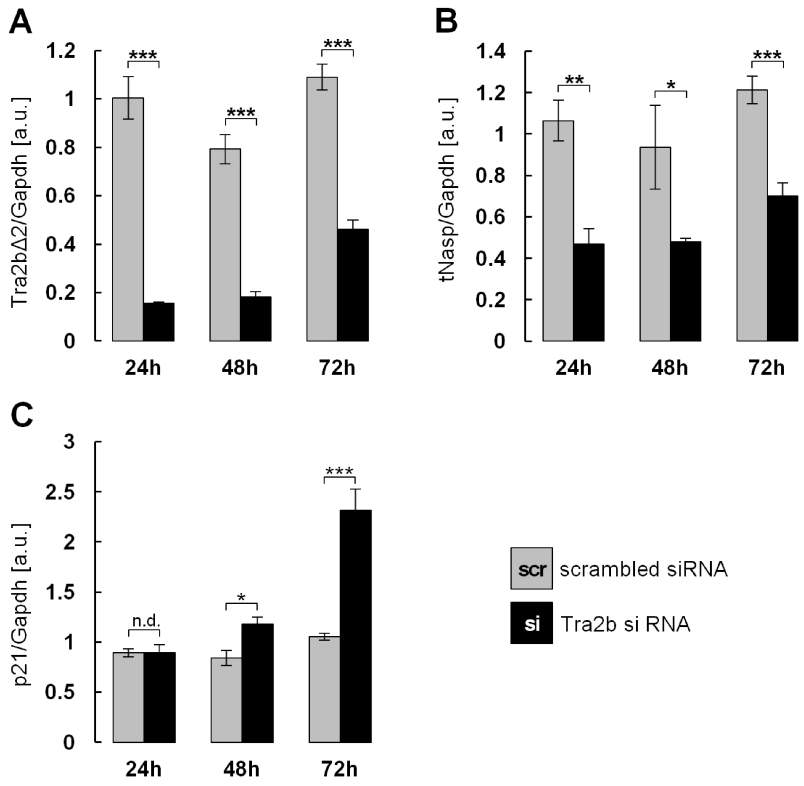

Figure 37 | Tra2b depletion in NSC34 cells: Analysis on RNA level.
NSC34 neural stem cells were transfected with siRNAs specific to *Tra2b* or scrambled siRNAs using the DharmaFect I transfection reagent. Expression levels of the functional *Tra2b Δ2* isoform (A), *tNasp* (B) and *p21* (C) were determined using quantitative real-time PCR. All data were normalized to *Gapdh* expression. Specific primers were: *Tra2b Δ2* (4271/4272), *tNasp* (4049/4042), *p21* (4493/4492), *Gapdh* (3999/4000). siRNA treatment effectively reduced *Tra2b* mRNA levels 24h, 48h and 72h after transfection. *Tra2b* ablation caused reduction of *tNasp* isoform expression by approx. 50 %. *P21* mRNA levels were slightly but significantly increased after 48h but are highly significantly increased by 2.2-fold in *Tra2b*-depleted cells 72h after transfection.

In preparation to rescue experiments a *tNasp* expression vector was produced. *tNasp* was amplified from murine adult testis cDNA using the primers 4594 and 4595 and cloned into pEGFP-C2 using XhoI and BamHI restriction ligation. The pEGFP vector used here has an N-terminal EGFP tag. This was chosen as the tNasp protein physiologically localizes to the nucleus and the nuclear localization signal is predicted at the C-terminal end of the protein (UniProt). Linking the EGFP tag to the N-terminal end of tNasp should prevent any interference of the tag with the nuclear localization signal. Identity of the cloned *tNasp* insert was validated by sequencing. Vectors used for transfection experiments were prepared using low-endotoxin plasmid purification kits.

NSC34 cells were co-transfected with scrambled or *Tra2b*-specific siRNA and pEGFP-tNasp or the pEGFP vector without any insert. RNA and proteins were isolated 48h and 72h after transfection. GFP-immunoblot analysis showed that transfection of the pEGFP or pEGFP-tNasp vector results in solid expression of EGFP (28 kDa) or EGFP-tNasp (111 kDa) with specific bands at the expected molecular weight (Figure 38). This indicated that intact EGFP-tNasp fusion-protein is produced from the generated pEGFP-tNasp vector. However, most cells

subjected to this experiment died 48h to 72h after transfection. This was equally true for cells transfected with pEGFP-tNasp but also for cells transfected with the pEGFP control vector independently of *Tra2b* silencing. This indicated that the sole presence of the pEGFP vector backbone is toxic for the cells probably due to high expression levels of EGFP. Also possible is a contamination with endotoxins from the plasmid preparation. In an attempt to reduce toxicity the amounts of vector were reduced to 0.5 µg pEGFP per 6-well plate. This equally resulted in a very high cell mortality rate.

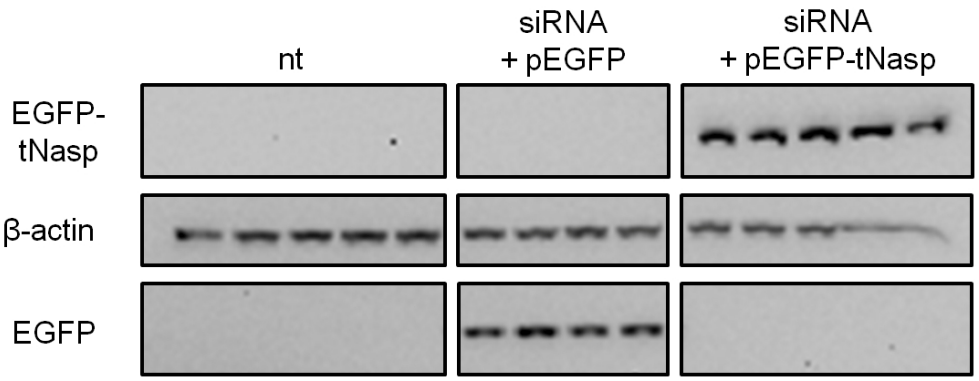

Figure 38 | Expression of EGFP-tNasp in NSC34 cells.
NSC34 cells were co-transfected with siRNA against *Tra2b* and pEGFP or pEGFP-tNasp. Proteins were isolated 48 hours after transfection. NSC34 showed solid expression of EGFP and the EGFP-tNasp fusion-protein, showing bands at the expected molecular weights of 28 and 111 kDa, repectively. β-actin was used as a loading control.

As previously determined (Figure 37 C), upregulated p21 expression levels take up to 72h of *Tra2b* depletion to be clearly detectable. However, due to the high rates of cell death in pEGFP-tNasp rescue experiments it was not possible to accurately determine *p21* levels via quantitative real-time PCR 72h after transfection. At this time point cells were affected to severely by the procedure, that also normalization of transcript levels was not possible due to highly fluctuating *Gapdh* expression. It is not clear whether the overexpressed tNasp protein itself also exerts a cytotoxic effect to NSC34 cells. However, the fact that the sole pEGFP vector even at low dosage is able to inflict great damage to the cells suggests that another expression system for tNasp is recommendable. Employing a tag-free expression system would allow evaluation of inert toxicity of the tNasp protein and would exclude possible interference with tNasp functionality.

Here the cell-cycle-regulating protein p21 was shown to be upregulated on RNA and on protein level as a response to *Tra2b*-depletion. This caused cell-cycle arrest and cell death in murine neuronal stem cells (NSC34 cells). Furthermore, *p21* was demonstrated to be upregulated in brains of neuronal-specific *Tra2b* knock-out mice. Upregulation of p21 in neuronal precursor cells and strongly increased apoptosis and loss of the proliferative

potential in the ventricular and subventricular zone of the developing murine brain strongly suggest that *Tra2b*-depletion causes brain developmental defects by increased p21 levels in neurogenic brain areas. By what signaling pathway it comes to upregulation of *p21* expression could not be elucidated at last. The involvement of the tNasp protein in the chain of events cannot be excluded and will be discussed in detail (6.6). Also other factors that act besides p21 upregulation might impact cell viability in the developing murine brain. However, p21 upregulation appears to be a major contributor in the brain developmental pathology and cell death, as to its direct involvement in cell cycle regulation (Harper et al. 1993)

6. Discussion

Alternative splicing is a complex and extremely potent mechanism to modulate gene function and individual splicing factors can exert an effect on a specified group of transcripts (Nilsen and Graveley 2010; Zheng and Black 2013). Many features of the general splicing machinery are well understood but little is known about the function and target transcripts of non-essential splicing factors. Mice deficient for the splicing activator *Tra2b* are embryonically lethal at early gestational stages (Mende et al. 2010). In this study a neuronal-specific knock-out mouse model for *Tra2b* was generated. These mice are born but die shortly after delivery and display severe developmental defects in the cortical and thalamic brain regions. Degeneration of these regions is preceded by massive apoptotic events resulting in a loss of proliferative potential. These data indicate that *Tra2b* is essential for survival and for neuronal development. Well-known Tra2b-related splicing processes were validated in neuronal specific *Tra2b* knock-out mice and new targets were identified using high throughput screening approaches. After discussing the brain developmental defects observed in KO mice (6.1 and 6.2) it will be considered to what extent the identified aberrantly spliced exons might contribute to cell death and misdevelopment (6.5). The applied high throughput screening strategies will be evaluated as to their potential to identify changes of splicing patterns *in vivo*, while emphasizing their individual advantages and drawbacks (6.3). Further, the mode of action of Tra2b and its interplay with splice sites as well as their total combinatorial effect on exon inclusion will be compared for several of its target exons (6.4). Finally, the impact of p21 upregulation and its possible foundations will be discussed (6.6).

6.1. Characterization of neuronal-specific *Tra2b* knock-out mice: Tra2b is essential for brain development in mice

Neuronal-specific *Tra2b* knock-out mice die within the first 36 h of life. This is in clear contrast to ubiquitously *Tra2b* depleted animals that die very early during embryonic development (Mende et al. 2010). Sole depletion of *Tra2b* in the central nervous system allowed birth of viable pups that were morphologically (outer appearance) indistinguishable from controls. Initial histological analysis of the brain was carried out and revealed gross morphological alterations in the brains of KO mice including severe ventriculomegaly of the third and the lateral ventricles. Brain volume that was occupied by largely cortical tissue in control animals was occupied by the ventricles in KO animals. Cortical structures and the

typical layered patterning of the cortex were not observable in KO animals. This clearly suggests an important role for *Tra2b* in neuronal development.

Furthermore, medial structures of the brain that were close to the third ventricle appeared absent in KO mice. These regions were with high likelihood identified as anterior parts of the thalamus. A clear histological identification of thalamic structures was difficult. In frame of histological analyses coronal sections (frontal sectioning) were analyzed and the alignment of sectioning levels for specimen comparison had to be set by reference points in the brain. In most cases this allowed comparison of sections belonging to sectioning levels very close to each other. Due to gross morphological changes in knock out brains, reference points could not always be clearly allocated. This made it difficult to assess the depth of sectioning from anterior to caudal and to distinguish between structures anterior of the thalamus, or the thalamus itself. Furthermore, the developing mouse brains underlied rapid morphological changes. Within 2 days of brain development, structures of the brain did significantly change their appearance which taken together made the precise identification of structures difficult.

The described brain phenotype including structural aberration and full loss of defined brain structures indicates a prominent role of *Tra2b* during development or maintenance of these structures. Indeed, Tra2b appeared to be strongly expressed in those regions of CTRL animals that are affected in KO animals. Moreover, both the cortex regions and the thalamic regions showed strongly increased occurrence of localized apoptosis measured by Caspase3 immunostaining. Caspase3 expression was found highest in the cortical layers close to the lateral ventricles (i.e. ventricular and subventricular zone) as well as at the thalamus. Massive apoptosis was followed by a drastic reduction in expression of the proliferation marker Ki67. The ventricular zone is a neurogenic area and origin of neuronal migration to form the layered structure of the cortex (Krubitzer and Huffman 2000; Lopez-Bendito and Molnar 2003). Consequently, cell death and the incapacity to proliferate are detrimental for normal cortex formation.

Another approach to explain failure in cortex development or maintenance is via the absence of thalamic input. Thalamic and cortical development depend on each other and occur largely in parallel. Most cortical restructuring takes place in the second to third week of gestation (13-18 dpc). Further, thalamocortical projections are necessary for normal cortical patterning (Lopez-Bendito and Molnar 2003). Occurrence of apoptosis in thalamic regions of KO mice parallels this developmental time window. Absence of thalamic projections targeting defined areas of the cortex might explain aberrant cortical patterning and probably deterioration of cortical structures. However, this study does not provide direct evidence for the absence of thalamocortical projections. From the data obtained in this study it cannot be

deduced which region of the brain was affected first, thalamus or cortex. Apoptosis occurred simultaneously at around 14.5 dpc in both structures. This might probably underline the parallel development and co-dependency of thalamus and cortex but does not necessarily indicate causality.

On the way of indentifying the underlying mechanism for brain malformations the question arose whether brain development itself is disturbed in KO mice – meaning structures are not properly generated – or whether properly developed structures undergo apoptosis causing the observed malformations. First, the gross morphological appearance of the cortex was largely indifferent between CTRL and KO animals at 14.5 dpc and cortical layers were found generally intact in KO animals at this early stage. This suggests that cortical development itself is obviously not impaired before 14.5 dpc. Also the temporal occurrence of apoptosis and loss of the proliferative potential support this idea. Onset of apoptosis was detected at 14.5 dpc continuing on until 15.5 dpc. However, the proliferative potential was not impaired at 14.5 dpc, but started to decrease at 15.5 dpc and was fully lost at 16.5 dpc. This strongly indicates that the loss of the proliferative potential is a direct consequence of apoptosis. Thus, brain development is not necessarily impaired *per se*. Instead properly generated structures die due to apoptosis, which in turn precludes further proliferation and differentiation of neurons.

Results indicated that apoptosis is the primary cause for brain malformations, but the question for the trigger of apoptosis remains. On the one hand side apoptosis could have been induced by the lack of thalamocortical signaling as described previously in this chapter. On the other hand apoptosis could be a secondary effect evoked by increased cerebroventricular pressure. Opposed to the idea that apoptosis causes cortical decay on a cellular level and as a consequence liquid filled ventricles occupy the vacant space, an increased pressure of cerebrospinal fluid within the ventricles may have caused repression of cortical tissue as well. This effect is referred to as normal pressure hydrocephalus (Adams et al. 1965). Increased pressure of cerebrospinal fluid may be caused by impaired liquor production or resorption, as well as duct obstruction preventing circulation and exchange of liquor between the ventricles. Thus, pressure-related decay of cortical tissue might be an alternative explanation for ventriculomegaly. Which molecular and cellular responses lastly mediate apoptosis and which of the previously discussed mechanisms is the basis of such a response cannot be answered at this point. The relevance of Tra2b-related missplicing events occurring in KO brains and their possible effects on the molecular level will be discussed in chapter 6.5.

Molecular biological analyses of the murine brains via quantitative real-time PCR suggested a rather constant expression of *Tra2b* throughout the brain that did not show gross alterations during embryonic development until birth. This suggests that there is no specific time point at which Tra2b is required most and therefore ablation of *Tra2b* is a non-physiological situation at any time point of development. Changes in splicing caused by its absence should thus evoke splicing patterns that are untypical for the developing brain at any given time point.

Even though neuronal-specific *Tra2b* knock-out mice always died within the first 36 h of life, the severity of their brain phenotype was subject of variation. Ventriculomegaly and loss of cortical patterning were always observable, but the extent of ventricular dilation appeared less severe in some KO animals than in others. This might be due to incomplete or strongly mosaic knock out of *Tra2b.* Indeed, some KO mice were demonstrated to retain increased *Tra2b* expression levels. Quantification of *Tra2b* expression demonstrated a robust average reduction of expression in KO animals at all investigated developmental stages, but single KO individuals showed *Tra2b* expression levels that were higher.

As described previously, gradually changed *Tra2b* expression levels (measured on whole brain RNA) are likely not due to real expressional modulation but rather originate from an altered ratio and distribution of *Tra2b* expressing and non-expressing cells in a mosaic brain. *Tra2b* knock-out was mediated by Cre-recombinase and *Cre* expression relied on the activity of the rat Nestin promoter in the different cell populations of the brain (Tronche et al. 1999). However, mosaic expression patterns of Cre-recombinase were theoretically expected to be consistent between individuals. For what reason *Tra2b* knock-out was differently efficient in individual mice remains elusive. The possibility of *Tra2b* expression being variable independently of Cre expression is supported by variable expression levels in control mice. These mice do not harbor the *Nestin-Cre* transgene and carry either the $Tra2b^{fl/+}$ or the $Tra2b^{fl/fl}$ alleles. Variability of *Tra2b* expression in these mice suggests that *Tra2b* expression levels might generally be subject of natural variability between individuals. Attempts to correlate mild brain phenotypes with retained *Tra2b* expression levels largely failed as analysis of RNA obtained from formalin-fixated and paraffin-embedded sections did not yield high quality data. By comparing *Tra2b* expression data of individual mice (Figure 15), *Tra2b* expression levels in KO mice were found to partially exceeded expression found in HET mice by far. Intriguingly, despite sometimes displaying lower expression levels than KO mice, HET mice always had normal life expectancies and never displayed any kind of brain phenotype. This strongly argues for a mosaic effect being the underlying reason for this variability. Not total

Tra2b amounts expressed throughout the brain, but rather the identity of cells or brain regions that retain *Tra2b* expression appears important for neuronal development and survival.

A less severe cortical phenotype has recently been demonstrated in a conditional *Emx1-Cre Tra2b* mouse model (Roberts et al. 2013). Emx1 is a transcription factor that is expressed in the embryonic cortex (Simeone et al. 1992). These knock-out mice were fully viable and did not show gross behavioral abnormalities, suggesting that cortical loss of *Tra2b* can be tolerated. Still these mice showed apoptosis within the cortex and disorganization of the cortical plate zone. The brain phenotype demonstrated for *Emx1-Cre* mice was clearly milder than observed in *Nestin-Cre* KO animals presented in this study. Depletion of *Tra2b* by *Nestin-Cre* in neuronal progenitor cells, obviously affects a broader range of neuronal cells, which showed to be essential for survival and affected regions of the brain beyond the cortex.

The neuronal-specific knock-out of *Tra2b* in this study was achieved by using expression of Cre-recombinase in neuronal and glial precursor cells (Tronche et al. 1999). Expression of Cre-recombinase in these cell populations has been demonstrated to exert genotoxic stress that manifests in hydrocephaly and cell death within the cortex and the ependymal cell layer (Forni et al. 2006). Thus, high expression levels of Cre-recombinase have been implicated to confuse developmental studies in mice.

In contrast to the previously mentioned study, all mice used in the present study were bred to generate exclusively mice carrying a single *Cre*-transgene. Thus, expression levels of Cre-recombinase were expected to be drastically reduced as compared to the above mentioned study (Forni et al. 2006). Furthermore, heterozygous knock-out animals (HET) were always included in analyses conducted in this study. HET and KO mice equally harbor a single *Cre*-transgene, but HET mice never displayed impaired survival or any kind of brain phenotype. This clearly indicates that Cre-recombinase-mediated toxicity does not contribute to the observed developmental defects in neuronal-specific *Tra2b* knock-out mice in any way.

6.2. Molecular analysis of *Tra2b* KO mice: Identification of well-known Tra2b-dependent splicing processes *in vivo*

In preparative experiments for high throughput approaches for the identification of Tra2b splicing targets, well-known Tra2b-dependent splicing processes should be reconstituted *in vivo*. In this study splicing of the *Nasp* T-exon, *Tra2a* exon 2, *Mapt* exon 10 and *Cltb* exon 5 was investigated using whole brain RNA of KO and CTRL mice. Furthermore, the self-splicing

effect of Tra2b regulating its own exon 2 was investigated (Stoilov et al. 2004). These exons have previously been identified as being splicied by Tra2b using *in vitro* experimental strategies including CLIP and minigene experiments (Stamm et al. 1999; Kondo et al. 2004; Grellscheid et al. 2011a). First attempts of identifying these splicing processes *in vivo* failed, as the detected differences were very small and non-significant. As described in 6.1, the residual *Tra2b* expression found in KO mice but also in non-KO genotypes was somewhat variable between individuals. Therefore, KO animals with *Tra2b* levels clearly higher than the average expression of that genotype were excluded from analyses. As mentioned, differential expression levels in KO brains are likely caused by Cre-recombinase mediated mosaicism or naturally occurring variations in *Tra2b* expression. When analyzing whole brain RNA (as it is the case here), retained *Tra2b* expression in some areas of the brain might not only cause amelioration of the brain phenotype but might also result in normal transcript splicing in non-affected tissues. During analysis of the whole brain mRNA mixture, mRNA from non-KO cell populations likely masked the splicing effect that occurred in KO cell populations. Exclusion of KO samples with high *Tra2b* levels minimizes the contribution of non-KO tissue to the whole brain RNA analyzed here. Indeed, analysis of whole brain RNA with *Tra2b/Gapdh* levels strictly lower than 0.4 allowed successful reconstitution of known Tra2b-related splicing events *in vivo*. This did not only prove the physiological relevance of these *in vitro* identified splicing processes in the developing mouse brain, but also provided proof of concept for the identification of altered splicing patterns that are based on absence of Tra2b in an *in vivo* system. This paved the way for using RNA obtained from whole brains of *Tra2b* knock-out mice in high throughput screening approaches to identify transcripts or exons targeted by Tra2b.

In vivo splicing analysis demonstrated that the autoregulatory splicing feedback-loop of Tra2b regulating inclusion of its own exon 2 is active in the developing mouse brain. This autoregulation is likely the primary cause for the absence of any brain phenotype in HET mice and likely also contributes to the dampening of splicing effects in HET mice that – if even detectable – were always much weaker than in KO animals. Heterozygous knock-out of *Tra2b* exerted clearly diverse regulatory effects on the amounts of functional (Δ2) and non-functional (FL) *Tra2b* isoform transcripts. While non-funcional transcript levels were very robustly downregulated, the functional transcripts only decreased by approx. 20 %. A heterozygous deletion of *Tra2b* disables 1 of 2 *Tra2b* gene copies. Still, non-functional transcripts are reduced by more than 50 % and functional transcripts by far less than 50 %. This suggests that the autoregulatory feedback-loop counteracts negative effects of the missing gene copy, resulting in almost non-reduced levels of functional transcripts. This is further supported by indistinguishable Tra2b protein levels between CTRL and HET mice. As mRNA levels of

functional transcripts were reduced by approx. 20 % but protein levels were not, an independently acting translational regulatory mechanism cannot be excluded. This data fully supports the observed absence of any developmental brain phenotype in HET mice and demonstrates that a single *Tra2b* gene copy is sufficient to provide the required Tra2b protein levels which is supported by a splicing-shift towards the functional *Tra2b* isoform. Even though, in HET mice significant splicing effects were never detected as compared to CTRLs, HET mice always showed the correct tendency and constituted the intermediate situation between CTRL and KO mice. Tra2b is known to impact splicing of exons in a concentration-dependent manner (Hofmann and Wirth 2002; Grellscheid et al. 2011a; Grellscheid et al. 2011b). It is likely that slightly reduced concentrations of functional Tra2b protein may exert a splicing effect on those transcripts that are very sensitive to changes in Tra2b concentratons but not to others. What are the characteristics of transcripts affected by such slight reductions in concentrations and what is the physiological relevance for that will be discussed in section 6.4.

Incomplete or mosaic knock-out of *Tra2b* in murine brains prevented the detection of Tra2b-mediated splicing processes *in vivo*. This might correlate with milder brain phenotypes observed in some KO animals and also explains a strong masking of Tra2b-dependent splicing processes. Furthermore, increased Tra2b levels could also be attributable for longer survival of new born mice, even though the variability in survival was minimal and mostly differed by some hours only. Exclusion of brain RNAs with insufficient *Tra2b* knock-out allowed solid detection of Tra2b-related splicing events. Only brains for which known-splicing events have been reconstituted were used for high throughput screening approaches. The general success and applicability of whole transcriptome sequencing and mouse exon arrays for whole brain RNA obtained from in part mosaic brains will be evaluated in the subsequent section (6.3).

6.3. Identification of alternatively spliced exons on murine whole brain RNA: The challenge of heterogeneous transcriptome compositions in complex tissues

A primary objective of this study was the *in vivo* identification of Tra2b splicing targets using high throughput screening approaches on mRNA obtained from neuronal-specific *Tra2b* knock-out mice. Due to mosaic effects, morphological changes of the brain, total expression levels of *Tra2b* and detectable effects on Tra2b splicing targets were variable between animals of the same genotype. To provide a solid basis for these experiments only animals with strong *Tra2b* downregulation were used for screening analyses (6.2).

In the frame of this study whole transcriptome sequencing (5.5) and mouse exon array analysis (5.6) were performed to identify novel Tra2b splicing targets. Both approaches yielded very limited success in the identification of novel splicing targets. The following sections will deal with the problems encountered during analyses and will speculate about possible underlying reasons for the weak performance. Therefore, this section will include specific aspects for (i) whole transcriptome sequencing, (ii) exon array analysis and (iii) aspects that commonly caused problems for both approaches.

(i) Whole transcriptome sequencing: None of the candidate exons obtained by whole transcriptome sequencing could be validated by subsequent PCR analysis. The primary reason for failure in detection of candidates and detection of a magnitude of false positives is likely technological inadequacy. Recent studies have revealed that whole transcriptome sequencing is able to identify most of the expressed mRNA isoforms even at moderate levels of coverage with approx. 100 x 10^6 reads (Toung et al. 2011). However, for a precise determination of transcript abundances as it is needed for splicing-related analyses about 500 x 10^6 reads have been found to be necessary (Toung et al. 2011). Whole transcriptome sequencing performed in frame of this study yielded an average coverage of 53-fold assuming a 100Mb transcriptome. Read numbers obtained were between 22.2 x 10^6 and 45.2 x 10^6 and are therefore far below the required values needed for accurate quantification of transcripts or exons. Due to very low coverage the representation of especially lowly abundant exons can be dramatically disturbed. Minor changes in expression of a lowly abundant transcript can easily be interpreted as a multi-factor up- or downregulation. This effect generates false positive targets whose actual change in expression or splicing is physiologically fully negligible. This effect became apparent during the validation process of candidates, in which isoform specific PCRs were used. Frequently, the isoform lacking the respective candidate exon could not be detected at all, even though transcriptome sequencing suggested highly significant changes in exon abundance. Such candidate exons were considered to be false positives as either the isoform lacking the exon of interest does not exist and was thus an artifact result, or *Tra2b* depletion caused a change in exon abundance, however at a very low expressional level that a physiological impact can virtually be excluded.

Another weakness of analysis that might have led to inaccurate detection and generation of false positive hits is the low number of samples compared. In whole transcriptome sequencing 3 CTRL samples were compared to 3 KO samples. A higher number of samples could have yielded a higher significance and a more powerful detection of true candidates; due to the high cost of whole transcriptome sequencing at the time point of performance the screening was limited to 3 CTRL vs. 3 KO samples, however. Indeed, the low

sample number was demonstrated to make solid calling of splice variants prone to single sample variations. Re-sequencing of a single sample and repetition of analysis changed candidates and statistics dramatically (5.5.3). A higher number of samples will generally be able to compensate for single sample variations, while preserving significance. In the present analysis localized coverage deviations within a single sample hampered the analysis in so far that it did not only cause detection of false positive exons, but also prevented the detection of *in vivo* validated and strongly regulated exons like the *Nasp* T-exon or *Tra2a* exon 2. This suggests that besides multiple false positive candidates, a whole magnitude of real candidates was not detected. The RPKM-based analysis, which is specialized for calculating exon abundances, while normalizing to compensate differences in the total read number and differential total transcript abundances, failed to overcome localized disturbances in sequencing coverage. Whether this weakness is purely attributable to this analysis method or equally to sequencing and read mapping algorithms cannot be determined with certainty at this point.

(ii) Exon array analysis: Multiple candidates were identified as being differentially spliced or differentially expressed between CTRL and KO mice. Despite only two exons were validated to be spliced by Tra2b in this study (*Sgol2* exon 4 and *Tubd1* exon 4), this provides proof of concept for the *in vivo* identification of splicing targets in a knock-out mouse model. However, this also demonstrates the limitations of the system. Like whole transcriptome sequencing also array analysis yielded multiple false positive candidates, but was able to in part recognize bona fide Tra2b splicing targets that were previously validated *in vivo.* Other well known target exons of *Lpin1* (Pihlajamaki et al. 2011) or *Mypt1* (Shukla and Fisher 2008) were not detected as being differentially spliced, however. Reduced *Tra2b* levels have been associated with upregulation of a lipogenic *Lpin1* isoform that is involved in obesity (Pihlajamaki et al. 2011). Splicing of *Mypt1* exon 23 has been shown to be regulated by Tra2b and to be responsible for the diversification of smooth muscles into fast and slow contractile types (Shukla and Fisher 2008). Results for both of these transcripts have been obtained using RNA from liver and muscle, respectively. Both *Lpin1* and *Mypt1* are physiologically relevant in these tissues. Even though expression databases suggest that both are expressed in brain, Tra2b-related splicing processes might be absent in brain and might therefore not be detected.

Despite confirmed targets were not identified as top ranking candidates and definite false positive candidates obtained a better rank, exon array analysis produced data of higher quality compared to whole transcriptome sequencing. The sole quantity of significant hits was smaller as compared to whole transcriptome sequencing and filtering was applicable to reduce candidates to reasonable amounts. This is likely based on a much smaller variability between

samples, as data obtained from arrays did not underlie any kind of coverage fluctuations as described before for whole transcriptome sequencing.

Even though analyses for changes in alternative splicing and changes of expression on a transcript wide level were performed independently, many candidates were identified as being alternatively spliced and transcriptionally regulated. Experimental proof was provided that whole transcript levels but not splicing processes of these candidates were altered. Astonishingly, the majority of double identified transcripts/exons were identified with decreased whole transcript abundances. It cannot be excluded, that the failure in analysis is a result of severe morphological changes in KO brains that via tissue loss caused exons or whole transcripts to appear lower expressed. However, the mode of analysis was chosen to prevent exactly this phenomenon, as exon abundances were normalized to whole transcript abundances. Hybridization problems of individual exons can be virtually excluded as a reason, as this would apply equally for and KO and CTRL samples. As also in whole transcriptome sequencing, weakly expressed transcripts are prone to false positive detection, as minor changes appear bigger than they actually are. This might in part be due to *in vitro* transcription-based amplification of transcripts that was performed during exon array preparation. All these arguments state possible explanations for aberrant splicing detection, but an ultimate disclosure covering the broad majority of aberrantly detected candidates cannot be provided here.

Analysis of exon array data included various filtering strategies including enrichment of putative Tra2b binding sites. Exons that were significantly detected as being alternatively spliced were further filtered for a high abundance of (A)GAA motifs (Clery et al. 2011). This short sequence motif statistically occurs once in a 256 bp stretch. Even candidate exons containing a multitude of their respective statistically expected number of GAA-motifs could not be validated as real Tra2b splicing targets. This suggests that even very high frequencies of putative binding sites are not necessarily indicative of Tra2b-related splicing regulation. Thus, filtering for putative binding sites is not a reliable method to call Tra2b-dependent exons, but can probably be used to complement other filtering approaches in parallel (as in Figure 27). However, both of the validated exons of *Sgol2* and *Tubd1* comprised multiple binding sites (4 and 10, respectively). Also many other *in vivo* validated Tra2b-dependent exons of *Mapt*, *Tra2a* and *Nasp* contain at least one putative binding site. These data indicate that the presence of GAA-binding motifs is not an indicator for Tra2b dependency, but exons that depend on Tra2b in most cases harbor at least a single binding site. However, there are examples of indirect regulation by Tra2b that does not depend on any sequence motif. The *Cltb* exon 5 is negatively regulated by Tra2b. Previous studies have shown that Tra2b indirectly

regulates the *Cltb* exon 5 (Stamm et al. 1992; Stamm et al. 1999) and this exon does not contain any putative binding site. The impact of the presence and the number of Tra2b binding sites on splicing will be further discussed in section 6.4.

(iii) Common problems in whole transcriptome sequencing and exon array analysis: The brain development of KO animals is strongly impaired and mice display severe alterations in the brain phenotype. Large areas of the brain, like major parts of the cortex, are fully absent in KO animals and the free brain volume is occupied by the liquid filled ventricles (ventriculomegaly). As introduced before (6.1 and 6.2), changes of the total cellular composition of the brain directly impact the composition of expressed transcripts, as well as the composition of spliced isoforms. Tissue loss has biggest and most direct impact on general transcript expression levels. Transcripts that are typically expressed in brain regions that are affected or absent in KO brains, were likely identified as being transcriptionally downregulated in KO mice. With very high probability those transcripts are false positive hits, as their downregulation is solely based on tissue loss. Therefore, the severe brain phenotype including massive tissue loss in neuronal-specific *Tra2b* knock-out mice virtually precludes the reliable detection of downregulated transcripts. The decrease in transcript abundance, that is based on tissue loss, might be very strong and is thus easily detectable in both whole transcriptome sequencing and exon arrays. Real physiological transcriptional downregulation is likely weaker than the previously mentioned artificial effect. Even if real downregulation events are properly detected in high throughput screening approaches, the much stronger tissue loss-based effects will appear to be more promising and more significant candidates, which makes real changes even more difficult to detect. This argumentation is also applicable to splicing processes and the quantification of single exons. Changes in splicing patterns indentified *in vivo* using neuronal-specific knock-out mice could also be due to changes in cellular brain composition and needed to be validated in artificial systems to proof the real impact of Tra2b on the respective exon (5.7).

In contrast to that, the situation is different for transcriptionally upregulated transcripts. If a transcript is identified as upregulated in KO brains in a high throughput screening, there is a high probability that this constitutes a real physiological transcriptional upregulation. Tissue loss and thus transcript loss counteracts the detection of upregulated transcripts and upregulated candidates likely are stronger upregulated than they appear. The cell cycle regulator p21 (*Cdkn1a*) was found to be upregulated in KO brains and Tra2b knock-down experiments indicated an even stronger upregulation of p21 in cell culture than in KO brains. This will be discussed in detail in section 6.6.

Even though both methods, whole transcriptome sequencing and exon arrays have each its own technological weaknesses, the major problem precluding solid detection of alternatively spliced exons primarily was the biological material itself. Whole brain RNA was used as starting material for both analyses. Brains of KO mice were too severely affected and whole brain RNAs obtained from CTRL and KO brains were like intrinsically too different to confidently call physiological splicing changes. This study was aiming at the identification of Tra2b splicing targets in neuronal tissue. Data obtained here suggests that a more homogeneous system that is not severely affected by absence of Tra2b *per se* is needed. To directly overcome heterogeneity-based problems that caused failure of solid quantification of exons and transcripts in the presented experiments, laser dissection of distinct brain areas would allow analysis of comparable brain areas of CTRL and KO mice. Indeed, expression patterns and alternative splicing are highly variable in different brain areas and laser microdissection in combination with quantitative PCR has been successfully used to detect such spatial differences (Baj et al. 2013). As only low amounts of starting material are needed for exon array analysis, RNA obtained from microdissected brain specimens would likely suffice to conduct a high throughput experiment while maintaining CTRL and KO samples comparable. Probably usage of whole brain RNA from 14.5 dpc could improve high throughput approaches, as the phenotype of KO brains is still moderate, but origins and foundations leading to the ultimate developmental defect are already manifested. Also the use of a cell culture system could be advisable. Even though this might not fully reflect physiological processes as they occur *in vivo*, a homogeneous cellular system could provide an accurate splicing comparison using high throughput approaches. Overexpression of *Tra2b* in cells as it is used in minigene approaches likely results in detection of non-physiological targets, as exons might be responsive to unnaturally high Tra2b concentrations, which does not reflect the physiological situation, however. Instead, *Tra2b* knock-down in cell culture could probably detect exons, whose inclusion is physiologically strictly Tra2b-dependent.

6.4. The role of Tra2b in alternative splicing processes: From fine tuning to a splicing switch

Tra2b has been demonstrated to promote as well as to repress splicing inclusion of exons (Stamm et al. 1999; Kondo et al. 2004). Thereby its mode of action is different as it can bind target mRNA directly via its RRM thus promoting exon inclusion (Stoilov et al. 2004), or act indirectly via other splicing factors or interacting partners to repress exon inclusion (Stamm

et al. 1992; Stamm et al. 1999). However, this paradigm is not fixed as the effect of Tra2b on exon inclusion is independent of its mode of action. Inclusion of *SMN2* exon7 has been shown to be promoted by Tra2b, which was further facilitated by hnRNP-G, the SR protein SRp30c (Hofmann and Wirth 2002) and TDP-34 (Bose et al. 2008). Furthermore, exons that were effectively included by FL Tra2b via sequence specific binding were strongly repressed by a Tra2b construct lacking the RRM, indicating that different isoforms of Tra2b can adversely either promote or repress exon inclusion (Grellscheid et al. 2011a).

Besides its direct or indirect effect, the strength of impact that Tra2b can exert on splicing of an exon may depend on its concentration or the number of Tra2b binding sites within an exon, which allows multiple molecules to associate sequence specifically. Moreover, the intrinsic splice site strength of an exon appears to strongly determine the maximal modulating impact that Tra2b can exert on exon inclusion (Hertel 2008; Grellscheid et al. 2011b). The effect of splice site strengths on the power of Tra2b became apparent in minigene splicing assays for *Sgol2* exon 4 and *Tubd1* exon 4. Experiments conducted in this study used human and murine minigenes in a human cellular system (HEK293T cells). As the splice site strengths are determined by the effectivity of the interplay of spliceosomal components with the respective splice sites of exons, it needs to be considered that the human/primate splice site consensi are different from rodent consensus sequences. Therefore, the strength of a murine splice site might be different in a human cellular system. Compared to the primate splice site consensus the 3' and 5' splice sites of human *SGOL2* exon 4 are stronger than the splice sites of the murine version. Accordingly, exon 4 of the human minigene was effectively included at physiological TRA2B concentrations in HEK293T cells (Figure 33). Knock-down or overexpression of TRA2B failed to abolish or to promote inclusion of that exon, suggesting that the intrinsically strong splice sites are sufficient to constitutively include exon 4 of the human *SGOL2* transcript. In contrast, the somewhat weaker splice sites of the murine *Sgol2* exon 4 only allowed inclusion of exon 4 in less than 50 % of transcripts at physiological TRA2B levels. Here overexpression of TRA2B promoted exon 4 inclusion up to 90 % supporting the idea that TRA2B has a greater impact on splicing in the context of weak splice sites (Grellscheid et al. 2011b). Both the murine and human version of *TUBD1* have rather weak splice sites. Disregarding the plethora of other splice factor binding sites present in both versions of the *TUBD1* exon 4, TRA2B exerts a positive effect on splicing inclusion at physiological levels in HEK293T cells. This was demonstrated in TRA2B knock-down experiments in which exon 4 inclusion was slightly but significantly reduced. This suggests that TRA2B is not necessary but at least contributes to *TUBD1* exon 4 inclusion at physiological levels. Given the high number of putative binding sites for other splice factors it appears likely that TRA2B acts in concert

with other factors to facilitate splicing and to compensate weak splice site strengths of *TUBD1* exon 4. This is further supported by the fact that TRA2B overexpression causes virtually complete inclusion of exon 4 into the *TUBD1* transcripts.

The *Nasp* T-exon is roughly 1 kb long and is therefore clearly longer than the average eukaryotic exon (Luco et al. 2011). Absence of Tra2b was shown to cause strong downregulation of T-exon inclusion (Figure 18). This suggests not only that Tra2b promotes T-exon inclusion but also that it does so very strongly at physiologically low concentrations. This might be explained by the more than 10 in part clustered putative Tra2b binding sites within the T-exon. Furthermore, spliceosome assembly for the T-exon likely occurs over the exon, as its up- and downstream introns (594 bp and 4,639 bp) are clearly too large to allow spliceosome assembly spanning the intron (Fox-Walsh et al. 2005). For that reason efficient splicing of the *Nasp* T-exon strongly relies on the splice sites and other accessory splice factors to facilitate inclusion, as also the sole length of the T-exon makes recognition more difficult. This is in line with the very strong impact that physiological Tra2b levels exert on T-exon inclusion. Indeed, targeted disruption of some Tra2b binding sites within the T-exon has been demonstrated to dramatically reduce T-exon inclusion at physiological Tra2b levels (Grellscheid et al. 2011a). This has been rescued by increased Tra2b concentration which likely worked via remaining Tra2b binding sites in the T-exon.

Taken together these data and data obtained in frame of this study suggest that the impact of Tra2b on splicing can be modulated by many factors: The splice site strengths and the lengths of the respective exon and the flanking introns, the concentration (i.e. expression) of Tra2b and the number and probably location of Tra2b binding sites. Therefore, the effects that Tra2b exerts on different target exons are very versatile.

The effect of strong versus weak splice sites on the impact of Tra2b-mediated splicing enhancement has been impressively demonstrated for the *HIPK3* T-exon (Grellscheid et al. 2011b). Cooperation of multiple Tra2b binding sites has been shown to be able to assure solid exon inclusion and to overcome an intrinsically weak exon definition based on weak splice sites. Exons of this kind have been proposed to require Tra2b and multiple binding sites for effective inclusion. Binding sites have been demonstrated to become partly redundant with increased splice site strengths. Strong splice sites accordingly have been shown to reduce the positive total effect that Tra2b could exert on exon inclusion (Grellscheid et al. 2011b).

These data obtained for the *HIPK3* T-exon (Grellscheid et al. 2011b) and by *in vivo* splicing analyses of *Mapt* exon 10, *Nasp* T-exon and *Tra2a* exon 2 in this study, demonstrate the two adverse extremes of Tra2b function in terms of positive splicing regulation: Fine tuning

of exon inclusion as for *Mapt* exon 10 and adversely, function as a splicing-switch by drastically changing exon inclusion rates as shown for the *Nasp* T-exon and *Tra2a* exon 2. Especially the latter appear to be of great physiological relevance as knock-out of *Tra2b* compared to physiological levels caused tremendous loss of exon usage. The testis-specific T-exon of *Nasp* constitutes a robust example for tissue specific Tra2b-dependent splicing, as high concentrations of Tra2b – as they are present in testis (Grellscheid et al. 2011a) – are needed to facilitate effective inclusion. Possible physiological roles of *Nasp* splicing will be further discussed in 6.5. Other well-known examples for splicing-switch mechanisms are the splicing of *MYPT1* exon 23 (Shukla and Fisher 2008) and sex determination in *D. melanogaster* which is mediated by Transformer-2-dependent splicing of the *doublesex* gene (Belote and Baker 1982). In both cases the concentration of Tra2b is the critical determining factor for exon inclusion or exon skipping. In *D. melanogaster Transformer-2* is much stronger expressed in females than in males, which allows effective inclusion of *doublesex* exon 4 only in females but not in males (Belote and Baker 1982; Butler et al. 1986). In the *MYPT1* transcript inclusion of exon 23 mediates diversification of fast contractile smooth muscles. TRA2B concentrations have been reported to be 10-fold lower in smooth muscles that are committed to a slow contractile phenotype (Shukla and Fisher 2008). Thus, differential expression of *Tra2b* may contribute significantly to diversity between different tissues (or sexes) by switch-like regulation of splicing processes.

As demonstrated by *in vivo* splicing analyses *Tra2b* depletion had a rather weak effect on *Mapt* exon 10 inclusion (Figure 17). Even in absence of Tra2b, *Mapt* exon 10 was less but still effectively included, suggesting that it does not fully depend on Tra2b. This is in line with multiple splicing enhancing/regulating elements in *Mapt* exon 10 that have been functionally validated due to their association with exon 10 missplicing in neuronal disease (reviewed by (Liu and Gong 2008)). This delicately regulated splicing of *Mapt* exon 10 appears to reflect the physiological requirements as exon 10 inclusion needs to be precisely balanced to assure correct microtubule organization and neuronal functioning (Goedert and Jakes 1990; Wang et al. 2005; Glatz et al. 2006; Wang and Liu 2008; Zhou et al. 2008). Taken together the tight regulation of *Mapt* exon 10 and the moderate impact of Tra2b on its inclusion suggests a role in fine tuning of that exon. Moreover, this argumentation demonstrates that *in vivo* splicing analysis in a KO system as it was conducted in this study, might probably allow making direct conclusions about the physiological importance of Tra2b for a specific exon. Possibly the extent of a splicing-shift caused by *Tra2b* ablation *in vivo* might allow to deduce the physiological impact of Tra2b for the respective splicing process directly. This way of analysis is not possible in Tra2b overexpression approaches. These experiments allow detection of a

general responsiveness of an exon to extreme splice factor concentrations. Such are not physiological, however, and a direct conclusion for a real function in splicing should not be deduced from overexpression experiments alone. This underlines the importance of *in vivo* splicing analyses to determine the real physiological extent of splice-modulating power of a splicing factor.

6.5. Implications of aberrantly spliced transcripts in murine brain development

Splicing targets of Tra2b have previously been identified primarily by *in vitro* approaches like minigenes or crosslinking-immunoprecipitation. Such approaches detect even weak and transient interactions of splicing factors with transcripts and often evaluate the impact of a splicing factor on an exon under non-physiological conditions including artificial exon constructs and overly high splice factor concentrations (i.e. overexpression). In this study several previously detected Tra2b-related splicing events were identified in the developing mouse brain under physiological conditions. Validation of splicing events *in vivo* demonstrates that Tra2b facilitates splicing under naturally occurring conditions, suggesting that the validated splicing events are of biological relevance and likely functionally implicated in neuronal development. Targets that were *in vivo* validated in frame of this study include Clathrin light chain b (*Cltb*), Microtubule associated protein tau (*Mapt*), the *Tra2b* paralog *Tra2a* and the Nuclear autoantigenic sperm protein (*Nasp*). Furthermore, Shugoshin-like 2 (*Sgol2*) and Tubulin delta 1 (*Tubd1*) were identified as novel *in vivo* targets of Tra2b. Whether and to what extent missplicing of these transcripts contributes to aberrant brain development in neuronal-specific KO mice cannot be defined with certainty. However, there are strong implications as some targets are closely related to neuronal function and disease. In the following, possible effects of aberrant splicing on murine brain development, but also general implications will be discussed with regard to individual alternatively spliced Tra2b *in vivo* targets. The possible physiological impact of *Nasp* missplicing will be discussed in a dedicated chapter (6.6).

Shugoshin-2, exon 4

By filtering and enrichment of exons containing Tra2b binding sites Shogoshin-2 (*Sgol2*) exon 4 was identified as a novel Tra2b splicing target and was furthermore confirmed to be responsive to increased TRA2B concentrations in minigene experiments (Figure 31, Figure 33).

Shugoshin-2 has beed demonstrated to protect the integrity of the cohesin-complexes which regulate sister chromatid cohesion during meiosis I (Llano et al. 2008; Clift and Marston 2011). In contrast to mitosis, in which sister chromatids are separated to achieve equal distribution to daughter cells, sister chromatids remain connected during anaphase of meiosis I and the homologue chromatid pairs are separated instead to form haploid daughter cells. Sister chromatids separate during meiosis II (Campbell and Reece 2002). Absence of Shugoshin-2 in mice has been shown to cause premature release of meiosis-specific cohesin complexes from anaphase I centromeres, which has been demonstrated to result in complete loss of centromer cohesion during metaphase II. This in turn has been described to be causative for generation of aneuploid gametes, which caused infertility in male and female mice (Llano et al. 2008). The same workgroup has demonstrated that *Sgol2* deficiency has no impact on mitosis, as adult tissues and cultured fibroblasts were not impaired over many passages. This is in contrast to Shugoshin (*SGO1*), which has very recently been demonstrated to regulate centromeric cohesion not only in meiosis but also during mitosis (Yin et al. 2013). Database entries for SGOL2 as expected suggest highest expression in reproductive tissue (primarily testis and ovary) but also in brain, muscle and internal organs. Still, very little is known regarding somatic functions of SGOL2. However, the facts of SGOL2 being expressed in non-reproductive tissue and the involvement of its homolog SGO1 in mitosis preserves the possibility of SGOL2 being involved in the regulation of mitotic chromatid cohesion as well. Disturbed chromosome distribution to daughter cells during mitosis could surely abolish cell proliferation and in turn cause brain developmental defects observed in neuronal-specific *Tra2b* knock-out mice. Ablation of *Tra2b* was demonstrated to result in reduced inclusion of exon 4 of the murine *Sgol2* transcript *in vivo*. Exon 4 has a length of 78 bp and its splicing-related skipping removes 26 amino acids from the protein product but does not perturb the coding frame of the transcript. As no functional domains have yet been identified for SGOL2, the molecular consequences of exon 4 loss cannot be assessed. Despite loss of the molecular function of Sgol2 in mitotic tissue might be sufficient to cause cell death, missplicing of *Sgol2* exon 4 is likely not contributing to aberrant brain development in neuronal-specific knock-out mice. Evidence has been provided that Sgol2 is not involved in the regulation of mitotic chromatid cohesion (Llano et al. 2008) and absence of the exon 4 coding information cannot be further assessed with regard to protein function and stability. Most importantly, changes in exon 4 splicing caused by *Tra2b*-depletion were minor *in vivo.* Even though exon 4 was significantly less included in *Sgol2* transcripts, the full-length isoform was still the prominent transcript – regardless of the presence of Tra2b.

Given the involvement of SGOL2 in gametogenesis and the elevated expression of TRA2B in testis, especially during and after meiosis in spermatocytes and spermatids (Grellscheid et al. 2011a), SGOL2 could be a functional TRA2B target during spermatogenesis. Intriguingly, with *Sgol2* and *Tubd1* (subsequent paragraph) two strongly testis-related splicing targets were identified. Taken together, the strong expression of *Tra2b* in testis, splicing of *tNasp* in testis (Grellscheid et al. 2011a), and the identification of the two novel testis-related splicing targets *Sgol2* and *Tubd1* strongly suggests a central role for Tra2b in testis and likely during spermatogenesis. Heterozygous ablation of *Tra2b* in ubiquitous knock-out mice has not been reported to exert any effect on fertility (Mende et al. 2010). However, the compensatory effect from the auto-regulatory splicing-feedback loop, likely prevents any phenotype in ubiquitous heterozygous KO mice. The impact of Tra2b-related missplicing of *Sgol2* exon 4 and *Tubd1* exon 4 will be further investigated in a collaborative approach in the laboratories of David J. Elliott using testis-specific *Tra2b* knock-out mice. This might allow optimal assessment of the impact of *Tra2b* depletion on *Sgol2* and *Tubd1* splicing in the physiologically relevant tissue.

Tubulin delta chain, exon 4

Tubd1 exon 4 was identified as a Tra2b splicing target *in vivo* and was validated in minigene experiments (Figure 31, Figure 33). Delta tubulin has been reported to be highly expressed in testis and to be associated with the manchette, which is a specialized microtubule system that is involved in the reshaping of the sperm head during spermatogenesis (Smrzka et al. 2000). Nuclear elongation is a hallmark of spermatid reshaping and Tubd1 has been suggested to participate in the formation of the manchette, more specifically of the perinuclear ring that is generated during round-to-elongating transition (Kato et al. 2004). The somatic function of delta tubulin is poorly understood: In mammals it has been found in the cytoplasmic and nuclear fractions of cells, but has not been shown to co-localize with microtubules (Smrzka et al. 2000). The same workgroup has demonstrated co-immunoprecipitation of delta and gamma tubulin and suggested a distinct role for mammalian delta tubulin that is different from other known tubulins. While strongly expressed in testis, expression levels in brain are low. Thus, the physiological relevance to brain development cannot be fully clarified. Skipping of exon 4 from the *Tubd1* transcript causes a frameshift resulting in numerous premature termination codons. Results obtained in this study indicate that skipping of exon 4 results in nonsense-mediated decay (Figure 31), thus likely resulting in a decrease of the Tubd1 protein concentration. As similarly mentioned for *Sgol2*, the physiological effects of *Tubd1* missplicing and consequently downregulation could optimally be investigated in testis-specific *Tra2b*

knock-out mice. Functionally, Tra2b-mediated splicing of *Tubd1* exon 4 might constitute a mechanism to regulate *Tubd1* expression on a post-transcriptional level to restrict its expression to specific stages of spermatid development. During spermatogenesis Tubd1 has been demonstrated to be part of the perinuclear ring which forms during the transition from round to elongating spermatids (Kato et al. 2004). This transition is preceded and partly paralleled by massive Tra2b expression during spermatocyte and round spermatid stages (Grellscheid et al. 2011a). These events may probably be temporally well coordinated. Speculatively, elevated *Tra2b* levels in round spermatids might increase *Tubd1* expression via exon 4 inclusion to facilitate transition into spermatid elongation.

Clathrin light chain B, exon 5

Clathrins are functionally versatile proteins mainly involved in vesicle coating and regulation of endocytosis (reviewed in (Brodsky 2012)). Splicing of exon 5 (also exon EN) of Clathrin light chain b (*Cltb*) has been shown to be indirectly regulated by Tra2b in a negative fashion *in vitro* (Stamm et al. 1999). In this study the Tra2b-dependent regulation of *Cltb* exon 5 was detected *in vivo* (Figure 17), suggesting that this repressive mode of splicing modulation by Tra2b is physiologically relevant in the murine brain. The neuronal *Cltb* isoform II (exon 5 included) comprises regulatory features for the C-terminal calmodulin binding domain. This isoform has been demonstrated to be sensitive to calcium and it has been proposed that skipping of *Cltb* exon 5 (isoform III) might keep vesicles in a calcium-insensitive state in non-neuronal tissue (Stamm et al. 1992; Pley et al. 1995). In neurons this might represent a calcium dependent regulatory feature for vesicle release and recycling (Wu et al. 2009; Haucke et al. 2011; Saheki and De Camilli 2012). Altered splicing of *Cltb* exon 5 might therefore perturb the responsiveness of synaptic vesicles to calcium and cause impact synapse formation and function.

Microtubule associated protein tau, exon 10.

Microtubules determine and regulate cell morphology and help organizing intracellular components. The microtubule associated protein Tau (*Mapt*) belongs to the family of microtubule associated proteins (MAP), which can regulate the microtubule polymerization state and may facilitate interactions to various components of the cytoskeleton (Ramirez et al. 1999). Tau has been described to regulate microtubule stability and assembly and to exert direct impact on neuronal morphology (reviewed in (Wang and Liu 2008)). It is enriched in axons of mature as well as developing neurons (Kempf et al. 1996). Hyperphosphorylated Tau has been reported to be the major constituent of neurofibrillary tangles and has been

associated to several diseases that are commonly termed tauopathies (Andreadis 2005). These include e.g. progressive supranuclear palsy, corticobasal degeneration, Pick's disease and FTDP-17 (Buee et al. 2000; Ingram and Spillantini 2002).

There are numerous isoforms being generated from the *MAPT* gene, as 8 out of 16 exons are alternatively spliced in humans (Andreadis 2005). Especially missplicing of exon 10 which encodes 1 out of 4 microtubule binding domains, thereby controlling microtubule affinity, is involved in numerous neurological disorders and has been studied extensively (Jiang et al. 2003; Kondo et al. 2004; Shi et al. 2008; Zhou et al. 2008). Under normal conditions the ratio between the 3R (exon 10 skipped) and the 4R isoform (exon 10 included) is tightly regulated by TRA2B-dependent splicing processes in humans and both isoforms coexist at similar levels. This defined ratio is a prerequisite for normal microtubule function and stability and has been shown to be shifted by increased exon 10 usage in postmortem Alzheimer's disease brains (Glatz et al. 2006). Nevertheless, also a decrease of exon 10 usage has equally been shown to cause neurodegeneration (Andreadis 2005). In this study, ablation of *Tra2b* in KO mice was shown to reduce exon 10 usage causing changes of the 3R:4R ratio in the developing mouse brain. Notably, exon 10 splicing of the murine *Mapt* gene is different from humans. While humans preserve a tightly regulated balance between 3R and 4R protein species during adulthood (Goedert and Jakes 1990; Gao et al. 2000), adult rodents use exon 10 constitutively and thus produce exclusively the 4R isoform. Co-expression of both isoforms in mouse occurs only during embryonic development and expression of the 3R isoform has been reported to decline at postnatal day 8 (Kosik et al. 1989). Despite expressing solely 4R Tau in adulthood, aged rodents have been reported to fail developing neurofibrillary tangles (NFT) and induction of NFT in transgenic animals has appeared equally difficult (Gotz 2001; Andreadis 2005). Even though adult rodents are obviously and in contrast to humans immune to NFT formation based on increased 4R Tau levels, the tight regulation of 3R to 4R ratios during development suggests distinct functions of both isoforms in the developing brain. Shifted compositions of 3R and 4R Tau species during development can add unique regulatory cues to the microtubule network that might be necessary for axonal outgrowth and guidance – processes that are of central importance during brain development. Still, distinct functions of 3R and 4R Tau isoforms beyond microtubule regulation cannot be excluded.

Ablation of *Tra2b* during murine brain development was shown to directly impact the isoform distribution in KO brains. Due to the fact that Tau isoforms are equally balanced in the developing rodent as compared to humans, it appears likely that a disturbance of this splicing regulation might similarly elicit neurodegenerative processes in the mouse and finally contribute to the observed brain phenotype.

Even though single missplicing events discussed here provide good indication for a causative relationship to failure of neuronal development and neurodegeneration, the whole phenotype is more likely to be a combinatorial consequence of multiple factors being deregulated. Furthermore, the inability of high throughput screening approaches to detect Tra2b splicing targets is not indicative of the lack of such targets. With a very high probability many targets of Tra2b are hitherto not identified and each single one of them might contribute to the global brain phenotype of neuronal-specific *Tra2b* knock-out mice.

6.6. Tra2b depletion transcriptionally activates *p21* and causes apoptosis in neurogenic brain areas

Mouse exon array analysis identified a 1.5-fold transcriptional upregulation of the cell cycle regulator p21 (*Cdkn1a*, cyclin dependent kinase inhibitor 1a) in brains of neuronal-specific *Tra2b* knock-out mice compared to controls. P21 is a well known inhibitor of the cell cycle that works via inhibition of the cyclin-dependent kinases CDK2 and CDK4, which leads to G_1-phase cell cycle arrest (Harper et al. 1993). Inhibition of CDKs prevents phosphorylation of retinoblastoma (Rb) and thus prevents dissociation of the histone deacetylase-Rb repressor complex and activation of transcription factors (e.g. E2Fs). This causes repression of S-phase genes and the cell enters G_1-phase arrest (Nevins 1992; Harper et al. 1993). P21 upregulation that was detected on the array was confirmed by comparing cohorts of CTRL, HET and KO mice via semi-quantitative RT-PCR. Quantification indicated only a moderate (but significant) upregulation of *Cdkn1a* transcripts in whole brain RNA. Coming back to an earlier assessment, the measured upregulation of transcripts may deviate negatively from the real upregulation, as tissue loss in KO brains counteracts the detection of transcriptionally upregulated genes. Indeed, it can be speculated that p21 upregulation was restricted to affected brain areas that were demonstrated to undergo apoptosis and subsequentially lose their proliferative potential. These cell populations contribute a relatively small fraction to the whole brain RNA used for quantification experiments. Thus, p21 was supposedly massively upregulated in the affected brain areas, but not upregulated in unaffected regions. To generate evidence for this hypothesis attempts were made to immunohistochemically detect p21 on paraffin-embedded brain sections. Detection on brain sections was highly unspecific, however, which precluded a direct localization of p21 expression in the developing murine brain. In line with this hypothesis and with aberrant brain development in neuronal-specific *Tra2b* KO mice, p21 has

been shown to restrict neuronal proliferation in the subgranular zone of the dentate gyrus of the hippocampus (Pechnick et al. 2008) giving an example of how *p21* expression can impact CNS development.

To validate *in vivo* findings obtained in this study and to assess the extent of p21 upregulation in a homogeneous cellular system, *Tra2b* knock-down experiments were performed in the neuronal precursor cell line NSC34. Transfection of cells with siRNAs specific to *Tra2b* caused *tNasp* depletion due to altered splicing of the *Nasp* transcript which functionally indicated the loss of Tra2b. In contrast to the rather weak upregulation of p21 observed while using whole brain RNA, material from NSC34 cells showed a massive and highly significant upregulation of p21 on transcript and on protein level (Figure 36, Figure 37). Strikingly, the majority of Tra2b-depleted NSC34 cells had died 96 h after Tra2b depletion. Tra2b depletion was already detectable after 24 h, but strongly increased p21 levels were not detectable before 72 h of treatment had passed. This suggests that upregulation of p21 is closely related to cell death in NSC34 cells, which very likely reconstitutes the mechanisms causing apoptosis in brains of neuronal-specific knock-out mice. Moreover, susceptibility to *Tra2b*-depletion which results in cell death as a late consequence does not appear to be a general feature. Murine fibroblasts were depleted of *Tra2b* but were growing at high proliferation rates over many passages (also shown by (Mende et al. 2010)). This indicates that some cell types like fibroblasts probably tolerate absence of Tra2b, while others like neuronal-like cells do not. These findings strongly suggest that p21 upregulation that is caused by Tra2b-depletion promotes apoptosis and loss of the proliferative potential and might strongly be accounting for malformations of the brain observed in neuronal-specific *Tra2b knock*-out mice.

In search of a possible link between depletion of *Tra2b* and the transcriptional increase of *p21*, the splicing of tNasp achieved special attention. The nuclear autoantigenic sperm protein (*Nasp*) is a histone chaperone that is expressed as two distinct isoforms: The *sNasp* (somatic) isoform and the *tNasp* (testis or embryonic) isoform that was demonstrated to be expressed in embryonic brain and has previously been described to be strongly expressed in testis (Grellscheid et al. 2011a). Splicing regulation of the *tNasp*-specific cassette-exon is facilitated by numerous Tra2b binding sites which are evolutionarily highly conserved. The tNasp protein itself has previously been demonstrated to affect progression of the cell cycle, as its absence has been shown to cause failure in passing the G_1/S-phase border (Richardson et al. 2006). Furthermore, direct ablation of *tNasp* via siRNAs has been demonstrated to cause apoptosis of prostate cancer cells via an upregulation of p21 (Alekseev et al. 2011). As *Tra2b*-depletion as a primary effect causes splicing-related downregulation of *tNasp in vivo* and in

NSC34 cells, the question arises whether p21 upregulation is directly based on the loss of *tNasp*.

Mice ubiquitously depleted of *Tra2b* die at early stages of gestation (Mende et al. 2010), which precludes detailed analyses of a systemic upregulation of p21 during embryogenesis. A strong reduction in proliferative potential could well explain developmental arrest. Interestingly, mice ubiquitously depleted of *Nasp* survived until the blastocyst stage of development (approx. 3.5 dpc) (Richardson et al. 2006). The same workgroup proposed that survival until this stage was based on maternal tNasp protein present in the oocyte. Also depletion of tNasp in cells has been demonstrated to cause cell cycle arrest with some delay. Slow degradation of previously produced tNasp protein could well explain this delay and is in line with the delayed cell death observed in NSC34 cells as well.

tNasp has a well-established role as a histone chaperone and transporter. It has been demonstrated that tNasp directly binds to linker histones H1 and transports them into the nucleus in an energy and NLS-dependent fashion (Alekseev et al. 2005). Linker histones regulate nucleosome spacing and chromatin remodeling and thus have impact on gene transcription (Cheung et al. 2002; Horn et al. 2002; Fan et al. 2003). Within the nucleus, linker histones are constantly exchanged (Lever et al. 2000) and histone H1 molecules not bound to DNA have been suggested to be bound to Nasp (Alekseev et al. 2003; Alekseev et al. 2005). Thereby, Nasp has been demonstrated to be non-selective for any Histone H1 subtype (Richardson et al. 2000). These results imply Nasp in having a role in the regulation of linker histone homeostasis. Linking this to the putative role of tNasp in the transcriptional upregulation of p21 it appears intriguing that sole ablation of Histone H1 (independently of Nasp) causes cell cycle arrest and altered expression of cell cycle-related genes including p21 (Sancho et al. 2008). Based on this the following mechanistic model is proposed: *Tra2b*-dependent missplicing of *Nasp* results in deficient histone import to the nucleus and causes failure to pass the G_1/S-phase checkpoint via p21-mediated cell cycle inhibition.

To validate this hypothesis, attempts were made to rescue p21 upregulation and cell death in NSC34 cells by overexpressing tNasp from a vector. First experiments showed high cellular toxicity of tNasp overexpression in NSC34 cells and clear data could not be obtained. Similar to tNasp depletion, also tNasp overexpression has been demonstrated to affect the progression of the cell cycle (Alekseev et al. 2003), but dosage compensation of the expression vector did not reduce toxicity. Very likely, cellular toxicity is mediated by the pEGFP-C2 vector backbone itself, as transfection of an empty vector at equal amounts also caused the cells to die within 2 days after transfection. Thus, usage of another expression system is advisable.

6.7. Concluding remarks: Implications for treatment strategies involving histone deacetylase inhibitors

The present study was aiming at the identification of Tra2b splicing targets *in vivo* using neuronal-specific *Tra2b* knock-out mice. Identification of transcripts processed by Tra2b was in so far important as *Tra2b* is transcriptionally upregulated as a response to treatment with valproic acid and other HDAC inhibitors (Brichta et al. 2003; Riessland et al. 2006). HDAC inhibitor treatment has been demonstrated to be a promising approach as a therapy for SMA (1.6.1). However, HDAC inhibitors and especially valproic acid are also used or clinically tested for the treatment of epilepsy (Davis et al. 1994), bipolar disorder (Cipriani et al. 2013) and for the prophylaxis of episodic migraine (Linde et al. 2013). Furthermore, VPA has been attributed to have anti-cancer properties, as it abolished growth of HeLa cervical cancer cells and induced apoptosis in thyroid carcinoma cells (Cha et al. 2013; Han et al. 2013). Given the broad use of VPA in various conditions and its upregulating effect on the expression of Tra2b (Brichta et al. 2003), the identification of Tra2b-dependent splicing processes is of very high relevance. Splicing of exons that are regulated by Tra2b is likely to be altered by increased Tra2b levels and thus HDAC inhibitor - and especially VPA treatment - could cause splicing-related side effects.

VPA treatment increases the expression of *Tra2b*. In this study attempts were made to identify Tra2b targets in knock-out mice. Splicing related effects caused by Tra2b upregulation and knock-out are supposedly different, depending on the type of exon (6.4). Ablation of Tra2b exerts the biggest effects on exon inclusion, if the exon is generally weakly defined and its inclusion strongly relies on Tra2b. Overexpression of Tra2b has a different effect and exons that under physiological conditions are not or only weakly affected by Tra2b, might be promoted at strongly increased Tra2b concentrations. Notably, expressional increase of Tra2b induced by VPA is not equal to an overexpression but rather remains within a range similar to physiological levels. Such - compared to overexpression experiments - moderate increase of Tra2b expression is very likely to exert impact on the fine tuning of exons. Affected exons are probably characterized by generally weak splice sites and their inclusion to a great extent relies on splice factors other than the spliceosome. This makes putatively affected exons responsive to even moderately increased Tra2b concentrations. That effect of course applies for the weakly defined *SMN2* exon 7, but might also apply for *Mapt* exon 10. The *Mapt* exon 10 constitutes a very good example representing a tighly regulated and fine-tuned exon. *In vivo* depletion of Tra2b caused a moderate but significant decrease of exon 10 inclusion (Figure 17). Moreover, (physiologically) increased levels of *Tra2b* have been associated to enhanced exon

10 inclusion levels and to imbalances in the 3R:4R isoform ratio of Tau in post mortem Alzheimer's disease brains (Glatz et al. 2006). Slight disturbances in the ratio of Tau isoforms can disturb neuronal function, which makes Tau a critical candidate as a putative off target in VPA treatment strategies. For that reason, investigations on *Mapt* splicing during VPA treatment should be conducted to assess the direct impact of VPA treatment on the Tau isoform ratios.

This study was not able to largely expand the list of known Tra2b-depending splicing targets. This was mainly due to failure of high throughput screening approaches based on the strong heterogeneity between CTRL and KO brains used for comparison. Still, the identification of *Sgol2* and *Tubd1* as novel targets of Tra2b and the *in vivo* identification of bona fide Tra2b-related splicing processes, yielded new insights into the different modi of alternative splicing control mediated by Tra2b (6.4). Even though differentially strong impact on splicing of various exons can probably be explained by the number of Tra2b binding sites and the predicted splice site strength, this data and the current understanding of the interplay within the whole splicing machinery is insufficient to predict the impact of Tra2b on exon splicing without experimental testing. Therefore, further experimental studies are necessary to identify targets affected by HDAC inhibitor treatment. This study underlined that a subset of splicing targets is responsive to minor changes of Tra2b concentrations and some exons underlie a fine-tuned regulation of inclusion that is a prerequisite for normal physiological function. Also Tra2b was demonstrated to drive tissue-specific splicing processes that may convey tissue-specific functions to gene products. Thus, HDAC inhibitor-induced expressional changes of Tra2b definitively have the power to alter splicing processes that may cause aberrant function in multiple systems.

Besides Tra2b-mediated splicing aberrations, HDAC inhibitors are very broadly acting compounds that affect transcription of genes very unspecifically and alter the expression of a multitude of genes simultaneously. Even though this work focused on splicing-related aberrations, also deregulation of transcription is a source for side effects in HDAC inhibitor treatment as well and needs to be considered. A very common side effect of VPA treatment in various conditions is a massive gain of weight and the underlying functional mechanism is subject of controversial discussion (Verrotti et al. 2011). Typical features of long-term valproic acid treatment include metabolic disturbances like hyperinsulinemia and insulin resistance and VPA-related weight gain is associated with long-term vascular complications (Belcastro et al. 2013). Recently, metabolites of VPA have been identified to have an activating effect on the AMP-activated protein kinase, which is a key regulator of cellular metabolism (Avery and Bumpus 2013). Involvement of Tra2b in obesity and lipogenesis has been reported via splicing

of exon 6 of the *Lpin1* gene (Pihlajamaki et al. 2011). Inclusion of exon 6 has been demonstrated to promote expression of the lipogenic isoform of *Lpin1* and Tra2b has been proposed to inhibit exon 6 inclusion via competition with U1 snRNA. Increased levels of Tra2b have been shown to inhibit production of the lipogenic isoform. This mechanism of regulation is converse to weight gain during VPA treatment, which with virtual certainty excludes Tra2b to be involved in VPA-mediated weight gain via splicing of *Lpin1*. Of note, *Lpin1* exon 6 was not detected as an alternatively regulated candidate exon on mouse exon arrays, despite expression databases suggest *Lpin1* expression in the embryonic and adult brain.

7. Summary

TRA2B is a serine-arginine-rich splicing factor that contributes to the alternative splicing of exons and depletion of *Tra2b* in the mouse causes early embryonic lethality. It modulates splice site selection in a concentration dependent fashion and associates to target exons either directly via GAA binding motifs or indirectly via interactions with other splice factors. *TRA2B* is highest expressed in neuronal tissue and testis and its expression and activation are controlled via an autoregulatory self-splicing mechanism and by phosphorylation. Depending on interactors and on the position of binding, TRA2B can promote either inclusion or skipping of exons. *TRA2B* has been associated with splicing processes involved in development, vascularization, spermatogenesis and neuronal function and deregulation of splicing processes has been linked to conditions like Alzheimer's disease, dementia and Parkinson's disease. Importantly, TRA2B is involved in the splicing of the *SMN* transcript and has been demonstrated to promote inclusion of exon 7. The functional loss of the *SMN1* gene causes spinal muscular atrophy, whereas the copy number of *SMN2,* which is alternatively spliced to mainly block exon 7 inclusion, is the major determinant of the disease severity. Correction of the *SMN2* splicing pattern to increase the number of functional full-length transcripts is a promising avenue for SMA therapy. The use of HDAC inhibitors like VPA has been demonstrated to correct the *SMN2* splicing pattern to a great extent by transcriptional upregulation of *TRA2B*. Despite promising trial of VPA in SMA therapy and successful application in other conditions like epilepsy and bipolar disorder, it can be assumed that the splicing of other transcripts targeted by TRA2B besides *SMN2* will be affected by its upregulation.

In the present work a neuronal-specific *Tra2b* knock-out mouse was generated to identify transcripts targeted by Tra2b in the central nervous system. Mice homozygously depleted of *Tra2b* in the central nervous system died shortly after birth and presented severe abnormalities in cortical development. These included a loss of cortical patterning, a reduction of the total cortical width and ventriculomegaly. By immunohistochemical analyses of brain sections at developmental stages reaching from 14.5 dpc to birth, massive apoptosis was detected in the ventricular and subventricular layers of the lateral ventricles and in the thalamic region. Apoptosis at 14.5 dpc was followed by a loss of the proliferative potential in a timely fashion, which caused a progressive loss of cortical material and ventricular dilation that reached its end stage at 16-17 dpc. Based on a mosaic knock-out of *Tra2b* in the murine brain, mRNA and protein levels of Tra2b were drastically reduced, but not fully absent as demonstrated by quantitative PCR and semi-quantitative Western Blot. Heterozygous knock-

out animals were fully viable, showed no apoptosis, normal brain development and Tra2b protein levels comparable to control mice. Analysis of *Tra2b* isoforms revealed that the autoregulatory splicing feedback-loop of *Tra2b* is functional *in vivo* in the mouse brain and compensates for the loss of a single gene copy by upregulation of functional *Tra2b* transcripts due to skipping of exon 2.

By the use of quantitative PCR, previously and in minigene approaches identified Tra2b-dependent splicing processes of *Mapt*, *Cltb*, *Tra2a* and *Nasp* were indentified *in vivo* providing proof of concept for the identification of splicing processes in the *Tra2b*-deficient murine brain. For the *in vivo* identification of novel Tra2b target exons, whole transcriptome sequencing and mouse exon array analysis were performed using whole brain RNA from neuronal-specific knock-out mice. Based on dramatic differences between control and knock-out brains and drastically altered cellular compositions, the reliable identification of targeted exons proved to be difficult. Exons of the Shugoshin-like 2 and Tubulin delta chain were identified via exon array analysis to be promoted by Tra2b *in vivo*. These splicing processes were confirmed by quantitative PCR and both identified exons were shown to be responsive to increased Tra2b concentrations in a minigene splicing assay. Some of the *in vivo* identified Tra2b-dependent splicing processes have possible implications in neuronal development and function, and might therefore contribute to the observed aberrations in brain development.

Moreover, the cell cycle regulating gene *Cdkn1a* (encoding p21) was found upregulated in brains of neuronal-specific knock-out mice using exon arrays, and p21 upregulation was validated by quantitative PCR. P21 is a well-known inhibitor of cell cycle progression that causes cells to enter cell cycle arrest at the G_1-S-phase transition. RNAi-mediated depletion of Tra2b in NSC34 neuronal precursor cells caused a drastic increase of p21 expression on RNA and on protein level. Strikingly, NSC34 cells died shortly after p21 expression had increased, indicating an apoptotic effect. This suggests that an increased p21 expression is likely the reason for apoptosis and the loss of the proliferative potential in neurogenic areas of the developing *Tra2b*-depleted brains. Previous studies have linked deficiency of the histone chaperone Nasp to upregulation of p21 and to apoptosis. Further investigations need to clarify, whether Tra2b-related missplicing of Nasp (tNasp depletion) is the underlying reason for p21 upregulation and apoptosis.

8. Zusammenfassung

TRA2B ist ein SR-*related* Spleißfaktor, der das alternative Spleißen von Exonen reguliert und die Abwesenheit von *TRA2B* führt zu früher embryonaler Lethalität in Mäusen. TRA2B moduliert die Auswahl von Spleißstellen in Abhängigkeit von seiner Konzentration. Dabei assoziiert TRA2B entweder direkt über GAA Bindestellen mit seinem Zielexon oder es bindet indirekt durch Interaktion mit anderen Speißfaktoren. *TRA2B* wird am stärksten in neuronalem Gewebe und in Testis exprimiert, während die Expression selbst sowie der Aktivierungszustand über einen sich selbst regulierenden auto-Speißmechanismus und durch Phosphorylierung gesteuert werden. In Abhängigkeit von seinen Interaktionspartnern und der Bindestelle innerhalb eines Transkripts kann TRA2B entweder den Ein- oder Auschluss eines Exons fördern. *TRA2B* wurde in vorhergegangenen Studien mit Spleißprozessen in Verbindung gebracht, welche für die Entwicklung, Vaskularisierung, Spermatogenese und neuronale Funktion von Bedeutung sind. Störungen in TRA2B-abhängigen Spleißprozessen wurden in Krankheiten wie Alzheimer, Demenz und Parkinson nachgewiesen. Von hoher Bedeutung ist der Einfluss von TRA2B auf das Spleißen von *SMN* Exon 7. Der funktionelle Verlust des *SMN1* Gens verursacht spinale Muskelatrophie, wobei die Kopienanzahl des nahezu identischen *SMN2* Gens den Schweregrad der Erkrankung bestimmt. Durch alternatives Spleißen kommt es zum Ausschluss von Exon 7 des *SMN2* Gens in 90 % aller Transkripte. Die Wiederherstellung des *SMN2* Spleißmusters und die damit verbundene Vervielfachung von *SMN2* Volllängentranskripten bietet deshalb einen vielversprechenden Therapieansatz für die SMA. Vorhergegangene Studien haben gezeigt, dass HDAC Inhibitoren wie VPA das *SMN2* Spleißmuster über eine transkriptionelle Heraufregulation von *TRA2B* zu revertieren vermögen. VPA wurde in klinischen Studien für eine mögliche SMA Therapie getestet und VPA wird erfolgreich bei Epilepsie und bipolaren Störungen eingesetzt. Es kann jedoch davon ausgegangen werden, dass die transkriptionelle Verstärkung der *TRA2B* Expression neben *SMN2* das Spleißen von anderen Transkripten beeinflusst.

Im Rahmen dieser Studie wurden neuronal-spezifische *Tra2b* Knock-out Mäuse generiert, um von Tra2b regulierte Transkripte im zentralen Nervensystem zu identifizieren. Mäuse mit homozygoter *Tra2b* Deletion im zentralen Nervensystem starben kurz nach der Geburt und zeigten schwerwiegende Fehlbildungen des Gehirns. Diese äußerten sich insbesondere durch den Verlust der Organisation der cortikalen Laminierung, einer Reduzierung der cortikalen Dicke und Ventrikulomegalie. Durch immunhistochemische Analysen von Dünnschnitten des Gehirns beginnend bei 14.5 dpc bis zur Geburt, konnte in den ventrikulären und subventrikulären Schichten der lateralen Ventrikel sowie in der

Thalamusregion Apoptose nachgewiesen werden. Apoptose wurde erstmals bei 14.5 dpc beobachtet und wurde gefolgt von ausbleibender Proliferation, was einen fortschreitenden Verlust von cortikalem Material und die Erweiterung der lateralen und des dritten Ventrikels zur Folge hatte. Diese erreichten bei 16-17 dpc ihr Endstadium. Basierend auf einem mosaischen Knock-out von *Tra2b* im murinen Gehirn, waren mRNA- und Proteinspiegel von Tra2b zwar drastisch reduziert, aber noch immer nachweisbar. Heterozygot deletierte Mäuse waren uneingeschränkt lebensfähig und zeigten eine normale Gehirnentwicklung, die von Kontrolltieren nicht unterscheidbar war. Durch qPCR Analyse der *Tra2b* isoformen konnte der zuvor beschriebene Selbstspleißmechanismus von *Tra2b* funktionell *in vivo* nachgewiesen werden. Durch einen verstärkten Ausschluss von *Tra2b* Exon 2 wurde der Verlust einer *Tra2b* Genkopie nahezu vollständig kompensiert.

Durch isoformspezifische qPCRs konnten zuvor identifizierte Tra2b-abhängige Spleißprozesse von *Mapt*, *Cltb*, *Tra2a* und *Nasp in vivo* nachgewiesen werden. So konnte im Allgemeinen die Möglichkeit zur Identifizierung von Tra2b-abhängigen Spleißprozessen in konditionalen Knock-out Mäusen bewiesen werden. Zur Identifizierung unbekannter Tra2b-abhängiger Spleißprozesse wurden *whole transcriptome sequencing* und *Exon Array* Analysen mit Gehirn-RNA durchgeführt. Aufgrund der dramatischen Störungen während der Gehirnentwicklung, war die zelluläre Komposition der Gehirne insoweit verändert, sodass eine zuverlässige Identifizierung alternativ gespleißter Exone sich als schwierig erwies. Dennoch konnten Exone der Gene Shugoshin-like 2 und Tubulin delta chain *in vivo* identifiziert werden. Die entsprechenden Spleißprozesse wurden durch qPCR Analysen bestätigt und Minigenkonstrukte beider Exone zeigten gesteigerte Einschlussraten bei erhöhten Tra2b Konzentrationen. Einige der *in vivo* identifizierten Tra2b-abhängigen Spleißvorgänge sind in Prozessen der neuronalen Entwicklung involviert und könnten entscheidend zu den beobachteten Fehlbildungen des Gehirns beigetragen haben.

Des Weiteren wurde durch *Exon Array* Analysen eine Heraufregulation des Zellzyklusregulators *Cdkn1a* (p21) in *Tra2b* Knock-out Gehirnen nachgewiesen und durch qPCR validiert. P21 ist ein etablierter Zellzyklusinhibitor, der eine Unterbrechung des Zellzyklus im Übergang von der G_1- zur S-Phase auslösen kann. RNAi-vermittelte Herunterregulation von Tra2b in NSC34 neuronalen Vorläuferzellen verursachte eine drastische Erhöhung der p21 Expression auf mRNA und auf Proteinebene. Verblüffenderweise starben die meisten NSC34 Zellen kurz nach der Heraufregulation von p21, was auf einen apoptotischen Effekt schließen lässt. Das legt die Vermutung nahe, dass Apoptose und der Verlust proliferativer Aktivität, wie es in neuronal-spezifischen Knock-out Mäusen auftritt, in der *in vivo* erhöhten p21 Expression begründet ist. In vorhergegangenen Studien war die Defizienz des Histonchaperons Nasp mit

der Heraufregulation von p21 und Apoptose in Verbindung gebracht worden. Es sollte daher in zukünftigen Studien geklärt werden, ob Tra2b-vermitteltes Fehlspleißen von *Nasp* (Verminderung von *tNasp*) die Ursache für eine Erhöhung der p21 Expression ist.

9. Publications and poster contributions

Original publications:

*Grellscheid S, *Dalgliesh C, ***Storbeck M**, Best A, Liu Y, Jakubik M, Mende Y, Ehrmann I, Curk T, Rossbach K, Bourgeois CF, Stévenin J, Grellscheid D, Jackson MS, Wirth B, Elliott DJ
Identification of evolutionarily conserved exons as regulated targets for the splicing activator Tra2β in development. **2011, PLoS Genet.** *authors contributed equally to this work

Storbeck M, Hupperich K, Gaspar JA, Meganathan K, Martínez Carrera L, Wirth R, Sachinidis A, Wirth B
Neuronal-specific deficiency of the splicing factor Tra2b causes apoptosis in neurogenic areas of the developing mouse brain. **2014, PLoS One, accepted.**

Poster presentations:

13.05.2011 Annual Human Genetics meeting, Universität zu Köln, Cologne, Germany
M. Storbeck, M. Jakubik, K. Roßbach, L. Martínez, Y. Mende, P. Frommolt and B. Wirth
In vivo identification of the Sfrs10 spliceome in the central nervous system of conditional Sfrs10 knock-out mice

20.-24.05.2012 25th Course in Medical Genetics, Bologna, Italy
M. Storbeck, S. Grellscheid, C. Dalgliesh, M. Jakubik, Y. Mende, K. Roßbach, L. Martínez, A. Best, Y. Liu, I. Ehrmann, T. Curk, C. F. Bourgeois, J. Stévenin, D. Grellscheid, M. S. Jackson, D. J.Elliott and B. Wirth
Identification of evolutionarily conserved exons as regulated targets for the splicing activator Tra2b in conditional Sfrs10 knock-out mice

29.11.2012 Annual Human Genetics meeting, Universität zu Köln, Cologne, Germany
M. Storbeck, S. Grellscheid, C. Dalgliesh, M. Jakubik, Y. Mende, K. Roßbach, L. Martínez, A. Best, Y. Liu, I. Ehrmann, T. Curk, C. F. Bourgeois, J. Stévenin, D. Grellscheid, M. S. Jackson, D. J.Elliott and B. Wirth
The splicing activator Tra2b targets evolutionary conserved exons involved in meiosis and spermatogenesis

10. References

Adams, R. D., C. M. Fisher, S. Hakim, R. G. Ojemann and W. H. Sweet (1965). "Symptomatic Occult Hydrocephalus with "Normal" Cerebrospinal-Fluid Pressure. A Treatable Syndrome." N Engl J Med. **273**: 117-126.

Alekseev, O. M., D. C. Bencic, R. T. Richardson, E. E. Widgren and M. G. O'Rand (2003). "Overexpression of the Linker histone-binding protein tNASP affects progression through the cell cycle." J Biol Chem **278**(10): 8846-8852.

Alekseev, O. M., R. T. Richardson, J. K. Tsuruta and M. G. O'Rand (2011). "Depletion of the histone chaperone tNASP inhibits proliferation and induces apoptosis in prostate cancer PC-3 cells." Reprod Biol Endocrinol **9**: 50.

Alekseev, O. M., E. E. Widgren, R. T. Richardson and M. G. O'Rand (2005). "Association of NASP with HSP90 in mouse spermatogenic cells: stimulation of ATPase activity and transport of linker histones into nuclei." J Biol Chem **280**(4): 2904-2911.

Andersson, R., S. Enroth, A. Rada-Iglesias, C. Wadelius and J. Komorowski (2009). "Nucleosomes are well positioned in exons and carry characteristic histone modifications." Genome Res **19**(10): 1732-1741.

Andreadis, A. (2005). "Tau gene alternative splicing: expression patterns, regulation and modulation of function in normal brain and neurodegenerative diseases." Biochim.Biophys.Acta **1739**(2-3): 91-103.

Avery, L. B. and N. N. Bumpus (2013). "Valproic Acid is a Novel Activator of AMP-Activated Protein Kinase and Decreases Liver Mass, Hepatic Fat Accumulation, and Serum Glucose in Obese Mice." Mol Pharmacol.

Baj, G., D. Del Turco, J. Schlaudraff, L. Torelli, T. Deller and E. Tongiorgi (2013). "Regulation of the spatial code for BDNF mRNA isoforms in the rat hippocampus following pilocarpine-treatment: a systematic analysis using laser microdissection and quantitative real-time PCR." Hippocampus **23**(5): 413-423.

Baker, B. S. (1989). "Sex in flies: the splice of life." Nature **340**(6234): 521-524.

Baughan, T., M. Shababi, T. H. Coady, A. M. Dickson, G. E. Tullis and C. L. Lorson (2006). "Stimulating full-length SMN2 expression by delivering bifunctional RNAs via a viral vector." Mol Ther **14**(1): 54-62.

Baughan, T. D., A. Dickson, E. Y. Osman and C. L. Lorson (2009). "Delivery of bifunctional RNAs that target an intronic repressor and increase SMN levels in an animal model of spinal muscular atrophy." Hum Mol Genet **18**(9): 1600-1611.

Belcastro, V., C. D'Egidio, P. Striano and A. Verrotti (2013). "Metabolic and endocrine effects of valproic acid chronic treatment." Epilepsy Res.

Belote, J. M. and B. S. Baker (1982). "Sex determination in Drosophila melanogaster: analysis of transformer-2, a sex-transforming locus." Proc Natl Acad Sci U S A **79**(5): 1568-1572.

Belov, A. A. and M. Mohammadi (2013). "Molecular mechanisms of fibroblast growth factor signaling in physiology and pathology." Cold Spring Harb Perspect Biol **5**(6).

Berget, S. M. (1995). "Exon recognition in vertebrate splicing." J Biol Chem **270**(6): 2411-2414.

Blencowe, B. J., J. A. Bowman, S. McCracken and E. Rosonina (1999). "SR-related proteins and the processing of messenger RNA precursors." Biochem Cell Biol **77**(4): 277-291.

Bolstad, B. M., R. A. Irizarry, M. Astrand and T. P. Speed (2003). "A comparison of normalization methods for high density oligonucleotide array data based on variance and bias." Bioinformatics **19**(2): 185-193.

Bose, J. K., I. F. Wang, L. Hung, W. Y. Tarn and C. K. Shen (2008). "TDP-43 overexpression enhances exon 7 inclusion during the survival of motor neuron pre-mRNA splicing." J Biol Chem **283**(43): 28852-28859.

Brichta, L., Y. Hofmann, E. Hahnen, F. A. Siebzehnrubl, H. Raschke, I. Blumcke, I. Y. Eyupoglu and B. Wirth (2003). "Valproic acid increases the SMN2 protein level: a well-known drug as a potential therapy for spinal muscular atrophy." Hum.Mol.Genet. **12**(19): 2481-2489.

Brodsky, F. M. (2012). "Diversity of clathrin function: new tricks for an old protein." Annu Rev Cell Dev Biol **28**: 309-336.

Buckler, A. J., D. D. Chang, S. L. Graw, J. D. Brook, D. A. Haber, P. A. Sharp and D. E. Housman (1991). "Exon amplification: a strategy to isolate mammalian genes based on RNA splicing." Proc Natl Acad Sci U S A **88**(9): 4005-4009.

Buee, L., T. Bussiere, V. Buee-Scherrer, A. Delacourte and P. R. Hof (2000). "Tau protein isoforms, phosphorylation and role in neurodegenerative disorders." Brain Res Brain Res Rev **33**(1): 95-130.

Burglen, L., S. Lefebvre, O. Clermont, P. Burlet, L. Viollet, C. Cruaud, A. Munnich and J. Melki (1996). "Structure and organization of the human survival motor neurone (SMN) gene." Genomics **32**(3): 479-482.

Burnett, B. G., E. Munoz, A. Tandon, D. Y. Kwon, C. J. Sumner and K. H. Fischbeck (2009). "Regulation of SMN protein stability." Mol Cell Biol **29**(5): 1107-1115.

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

Butler, B., V. Pirrotta, I. Irminger-Finger and R. Nothiger (1986). "The sex-determining gene tra of Drosophila: molecular cloning and transformation studies." EMBO J **5**(13): 3607-3613.

Campbell, N. A. and J. B. Reece (2002). Biology. San Francisco, Pearson Education Inc.

Capecchi, M. R. (1989). "The new mouse genetics: altering the genome by gene targeting." Trends Genet **5**(3): 70-76.

Cartegni, L., M. L. Hastings, J. A. Calarco, E. de Stanchina and A. R. Krainer (2006). "Determinants of exon 7 splicing in the spinal muscular atrophy genes, SMN1 and SMN2." Am J Hum Genet **78**(1): 63-77.

Cartegni, L. and A. R. Krainer (2002). "Disruption of an SF2/ASF-dependent exonic splicing enhancer in SMN2 causes spinal muscular atrophy in the absence of SMN1." Nat.Genet. **30**(4): 377-384.

Cashman, N. R., H. D. Durham, J. K. Blusztajn, K. Oda, T. Tabira, I. T. Shaw, S. Dahrouge and J. P. Antel (1992). "Neuroblastoma x spinal cord (NSC) hybrid cell lines resemble developing motor neurons." Dev Dyn **194**(3): 209-221.

Cha, H. Y., B. S. Lee, S. Kang, Y. S. Shin, J. W. Chang, E. S. Sung, Y. S. Kim, J. W. Choi, J. H. Kim and C. H. Kim (2013). "Valproic Acid Sensitizes TRAIL-Resistant Anaplastic Thyroid Carcinoma Cells to Apoptotic Cell Death." Ann Surg Oncol.

Chen, M. and J. L. Manley (2009). "Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches." Nat Rev Mol Cell Biol **10**(11): 741-754.

Chen, X., L. Guo, W. Lin and P. Xu (2003). "Expression of Tra2beta isoforms is developmentally regulated in a tissue- and temporal-specific pattern." Cell Biol Int **27**(6): 491-496.

Chen, X., J. Huang, J. Li, Y. Han, K. Wu and P. Xu (2004). "Tra2betal regulates P19 neuronal differentiation and the splicing of FGF-2R and GluR-B minigenes." Cell Biol.Int. **28**(11): 791-799.

Cheung, E., A. S. Zarifyan and W. L. Kraus (2002). "Histone H1 represses estrogen receptor alpha transcriptional activity by selectively inhibiting receptor-mediated transcription initiation." Mol Cell Biol **22**(8): 2463-2471.

Chodavarapu, R. K., S. Feng, Y. V. Bernatavichute, P. Y. Chen, H. Stroud, Y. Yu, J. A. Hetzel, F. Kuo, J. Kim, S. J. Cokus, D. Casero, M. Bernal, P. Huijser, A. T. Clark, U. Kramer, S. S. Merchant, X. Zhang, S. E. Jacobsen and M. Pellegrini (2010). "Relationship between nucleosome positioning and DNA methylation." Nature **466**(7304): 388-392.

Churko, J. M., G. L. Mantalas, M. P. Snyder and J. C. Wu (2013). "Overview of high throughput sequencing technologies to elucidate molecular pathways in cardiovascular diseases." Circ Res **112**(12): 1613-1623.

Cipriani, A., K. Reid, A. H. Young, K. Macritchie and J. Geddes (2013). "Valproic acid, valproate and divalproex in the maintenance treatment of bipolar disorder." Cochrane Database Syst Rev **10**: CD003196.

Cisbani, G. and F. Cicchetti (2012). "An in vitro perspective on the molecular mechanisms underlying mutant huntingtin protein toxicity." Cell Death Dis **3**: e382.

Clery, A., S. Jayne, N. Benderska, C. Dominguez, S. Stamm and F. H. Allain (2011). "Molecular basis of purine-rich RNA recognition by the human SR-like protein Tra2-beta1." Nat Struct Mol Biol **18**(4): 443-450.

Clift, D. and A. L. Marston (2011). "The role of shugoshin in meiotic chromosome segregation." Cytogenet Genome Res **133**(2-4): 234-242.

Collins, C. A. and C. Guthrie (2000). "The question remains: is the spliceosome a ribozyme?" Nat Struct Biol **7**(10): 850-854.

Compton, S. J. and C. G. Jones (1985). "Mechanism of dye response and interference in the Bradford protein assay." Anal Biochem **151**(2): 369-374.

Cooper, T. A., L. Wan and G. Dreyfuss (2009). "RNA and disease." Cell **136**(4): 777-793.

Costa, V., M. Aprile, R. Esposito and A. Ciccodicola (2013). "RNA-Seq and human complex diseases: recent accomplishments and future perspectives." Eur J Hum Genet **21**(2): 134-142.

Cowper, A. E., J. F. Caceres, A. Mayeda and G. R. Screaton (2001). "Serine-arginine (SR) protein-like factors that antagonize authentic SR proteins and regulate alternative splicing." J Biol Chem **276**(52): 48908-48914.

Crawford, T. O. and C. A. Pardo (1996). "The neurobiology of childhood spinal muscular atrophy." Neurobiol Dis **3**(2): 97-110.

Dahlstrand, J., M. Lardelli and U. Lendahl (1995). "Nestin mRNA expression correlates with the central nervous system progenitor cell state in many, but not all, regions of developing central nervous system." Brain Res Dev Brain Res **84**(1): 109-129.

Daoud, R., B. M. Da Penha, F. Siedler, M. Hubener and S. Stamm (1999). "Activity-dependent regulation of alternative splicing patterns in the rat brain." Eur.J.Neurosci. **11**(3): 788-802.

Dauwalder, B. and W. Mattox (1998). "Analysis of the functional specificity of RS domains in vivo." EMBO J **17**(20): 6049-6060.

Davis, R., D. H. Peters and D. McTavish (1994). "Valproic acid. A reappraisal of its pharmacological properties and clinical efficacy in epilepsy." Drugs **47**(2): 332-372.

Denk, F. and R. Wade-Martins (2009). "Knock-out and transgenic mouse models of tauopathies." Neurobiol.Aging **30**(1): 1-13.

Dreumont, N., C. F. Bourgeois, F. Lejeune, Y. Liu, I. E. Ehrmann, D. J. Elliott and J. Stevenin (2010). "Human RBMY regulates germline-specific splicing events by modulating the function of the serine/arginine-rich proteins 9G8 and Tra2-{beta}." J Cell Sci **123**(Pt 1): 40-50.

Erkelenz, S., W. F. Mueller, M. S. Evans, A. Busch, K. Schoneweis, K. J. Hertel and H. Schaal (2013). "Position-dependent splicing activation and repression by SR and hnRNP proteins rely on common mechanisms." RNA **19**(1): 96-102.

Eswarakumar, V. P., I. Lax and J. Schlessinger (2005). "Cellular signaling by fibroblast growth factor receptors." Cytokine Growth Factor Rev. **16**(2): 139-149.

Fan, Y., T. Nikitina, E. M. Morin-Kensicki, J. Zhao, T. R. Magnuson, C. L. Woodcock and A. I. Skoultchi (2003). "H1 linker histones are essential for mouse development and affect nucleosome spacing in vivo." Mol Cell Biol **23**(13): 4559-4572.

Faustino, N. A. and T. A. Cooper (2003). "Pre-mRNA splicing and human disease." Genes Dev **17**(4): 419-437.

Ford-Perriss, M., H. Abud and M. Murphy (2001). "Fibroblast growth factors in the developing central nervous system." Clin Exp Pharmacol Physiol **28**(7): 493-503.

Forni, P. E., C. Scuoppo, I. Imayoshi, R. Taulli, W. Dastru, V. Sala, U. A. Betz, P. Muzzi, D. Martinuzzi, A. E. Vercelli, R. Kageyama and C. Ponzetto (2006). "High levels of Cre expression in neuronal progenitors cause defects in brain development leading to microencephaly and hydrocephaly." J.Neurosci. **26**(37): 9593-9602.

Fox-Walsh, K. L., Y. Dou, B. J. Lam, S. P. Hung, P. F. Baldi and K. J. Hertel (2005). "The architecture of pre-mRNAs affects mechanisms of splice-site pairing." Proc Natl Acad Sci U S A **102**(45): 16176-16181.

Fu, X. D. (1993). "Specific commitment of different pre-mRNAs to splicing by single SR proteins." Nature **365**(6441): 82-85.

Furman, E. and D. G. Glitz (1995). "Purification of the spliceosome A-complex and its visualization by electron microscopy." J Biol Chem **270**(26): 15515-15522.

Gao, Q. S., J. Memmott, R. Lafyatis, S. Stamm, G. Screaton and A. Andreadis (2000). "Complex regulation of tau exon 10, whose missplicing causes frontotemporal dementia." J Neurochem **74**(2): 490-500.

Garbes, L., M. Riessland, I. Holker, R. Heller, J. Hauke, C. Trankle, R. Coras, I. Blumcke, E. Hahnen and B. Wirth (2009). "LBH589 induces up to 10-fold SMN protein levels by several independent mechanisms and is effective even in cells from SMA patients non-responsive to valproate." Hum Mol Genet **18**(19): 3645-3658.

Glatz, D. (2005). Alternatives Fehlspleißen von Tau Exon 10, Htra2-a und Clk2 in post mortem Hirngewebe von Alzheimer-Patienten.

Glatz, D. C., D. Rujescu, Y. Tang, F. J. Berendt, A. M. Hartmann, F. Faltraco, C. Rosenberg, C. Hulette, K. Jellinger, H. Hampel, P. Riederer, H. J. Moller, A. Andreadis, K. Henkel and S. Stamm (2006). "The alternative splicing of tau exon 10 and its regulatory proteins CLK2 and TRA2-BETA1 changes in sporadic Alzheimer's disease." J.Neurochem. **96**(3): 635-644.

Goedert, M. and R. Jakes (1990). "Expression of separate isoforms of human tau protein: correlation with the tau pattern in brain and effects on tubulin polymerization." EMBO J. **9**(13): 4225-4230.

Gotz, J. (2001). "Tau and transgenic animal models." Brain Res Brain Res Rev **35**(3): 266-286.

Graham, F. L., J. Smiley, W. C. Russell and R. Nairn (1977). "Characteristics of a human cell line transformed by DNA from human adenovirus type 5." J Gen Virol **36**(1): 59-74.

Graveley, B. R. (2000). "Sorting out the complexity of SR protein functions." RNA **6**(9): 1197-1211.

Grellscheid, S., C. Dalgliesh, M. Storbeck, A. Best, Y. Liu, M. Jakubik, Y. Mende, I. Ehrmann, T. Curk, K. Rossbach, C. F. Bourgeois, J. Stevenin, D. Grellscheid, M. S. Jackson, B. Wirth and D. J. Elliott (2011a). "Identification of evolutionarily conserved exons as regulated targets for the splicing activator tra2beta in development." PLoS Genet **7**(12): e1002390.

Grellscheid, S. N., C. Dalgliesh, A. Rozanska, D. Grellscheid, C. F. Bourgeois, J. Stevenin and D. J. Elliott (2011b). "Molecular design of a splicing switch responsive to the RNA binding protein Tra2beta." Nucleic Acids Res **39**(18): 8092-8104.

Gu, H., J. D. Marth, P. C. Orban, H. Mossmann and K. Rajewsky (1994). "Deletion of a DNA polymerase beta gene segment in T cells using cell type-specific gene targeting." Science **265**(5168): 103-106.

Gunderson, F. Q. and T. L. Johnson (2009). "Acetylation by the transcriptional coactivator Gcn5 plays a novel role in co-transcriptional spliceosome assembly." PLoS Genet **5**(10): e1000682.

Gunter, C. and R. Dhand (2002). "Human biology by proxy." Nature **420**(6915): 509-509.

Hahn, W. C., S. K. Dessain, M. W. Brooks, J. E. King, B. Elenbaas, D. M. Sabatini, J. A. DeCaprio and R. A. Weinberg (2002a). "Enumeration of the simian virus 40 early region elements necessary for human cell transformation." Mol.Cell Biol. **22**(7): 2111-2123.

Hahn, W. C., S. K. Dessain, M. W. Brooks, J. E. King, B. Elenbaas, D. M. Sabatini, J. A. DeCaprio and R. A. Weinberg (2002b). "Enumeration of the simian virus 40 early region elements necessary for human cell transformation." Mol Cell Biol **22**(7): 2111-2123.

Hamilton, D. L. and K. Abremski (1984). "Site-specific recombination by the bacteriophage P1 lox-Cre system. Cre-mediated synapsis of two lox sites." J Mol Biol **178**(2): 481-486.

Han, B. R., B. R. You and W. H. Park (2013). "Valproic acid inhibits the growth of HeLa cervical cancer cells via caspase-dependent apoptosis." Oncol Rep **30**(6): 2999-3005.

Harper, J. W., G. R. Adami, N. Wei, K. Keyomarsi and S. J. Elledge (1993). "The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases." Cell **75**(4): 805-816.

Hattori, D., S. S. Millard, W. M. Wojtowicz and S. L. Zipursky (2008). "Dscam-mediated cell recognition regulates neural circuit formation." Annu Rev Cell Dev Biol **24**: 597-620.

Haucke, V., E. Neher and S. J. Sigrist (2011). "Protein scaffolds in the coupling of synaptic exocytosis and endocytosis." Nat Rev Neurosci **12**(3): 127-138.

Hauke, J., M. Riessland, S. Lunke, I. Y. Eyupoglu, I. Blumcke, A. El-Osta, B. Wirth and E. Hahnen (2009). "Survival motor neuron gene 2 silencing by DNA methylation correlates with spinal muscular atrophy disease severity and can be bypassed by histone deacetylase inhibition." Hum.Mol.Genet. **18**(2): 304-317.

Hertel, K. J. (2008). "Combinatorial control of exon recognition." J Biol Chem **283**(3): 1211-1215.

Hofmann, Y., C. L. Lorson, S. Stamm, E. J. Androphy and B. Wirth (2000). "Htra2-beta 1 stimulates an exonic splicing enhancer and can restore full-length SMN expression to survival motor neuron 2 (SMN2)." Proc.Natl.Acad.Sci.U.S.A **97**(17): 9618-9623.

Hofmann, Y. and B. Wirth (2002). "hnRNP-G promotes exon 7 inclusion of survival motor neuron (SMN) via direct interaction with Htra2-beta1." Hum.Mol.Genet. **11**(17): 2037-2049.

Horn, P. J., L. M. Carruthers, C. Logie, D. A. Hill, M. J. Solomon, P. A. Wade, A. N. Imbalzano, J. C. Hansen and C. L. Peterson (2002). "Phosphorylation of linker histones regulates ATP-dependent chromatin remodeling enzymes." Nat Struct Biol **9**(4): 263-267.

Hua, Y., T. A. Vickers, H. L. Okunola, C. F. Bennett and A. R. Krainer (2008). "Antisense masking of an hnRNP A1/A2 intronic splicing silencer corrects SMN2 splicing in transgenic mice." Am J Hum Genet **82**(4): 834-848.

Ikonomidou, C., F. Bosch, M. Miksa, P. Bittigau, J. Vockler, K. Dikranian, T. I. Tenkova, V. Stefovska, L. Turski and J. W. Olney (1999). "Blockade of NMDA receptors and apoptotic neurodegeneration in the developing brain." Science **283**(5398): 70-74.

Ingram, E. M. and M. G. Spillantini (2002). "Tau gene mutations: dissecting the pathogenesis of FTDP-17." Trends Mol Med **8**(12): 555-562.

Jacob, A. G., D. O'Brien, R. K. Singh, D. F. Comiskey, Jr., R. M. Littleton, F. Mohammad, J. T. Gladman, M. C. Widmann, S. C. Jeyaraj, C. Bolinger, J. R. Anderson, D. A. Barkauskas, K. Boris-Lawrie and D. S. Chandler (2013). "Stress-Induced Isoforms of MDM2 and MDM4 Correlate with High-Grade Disease and an Altered Splicing Network in Pediatric Rhabdomyosarcoma." Neoplasia **15**(9): 1049-1063.

Jiang, Z., H. Tang, N. Havlioglu, X. Zhang, S. Stamm, R. Yan and J. Y. Wu (2003). "Mutations in tau gene exon 10 associated with FTDP-17 alter the activity of an exonic splicing enhancer to interact with Tra2 beta." J.Biol.Chem. **278**(21): 18997-19007.

Johnson, M. B., Y. I. Kawasawa, C. E. Mason, Z. Krsnik, G. Coppola, D. Bogdanovic, D. H. Geschwind, S. M. Mane, M. W. State and N. Sestan (2009). "Functional and evolutionary insights into human brain development through global transcriptome analysis." Neuron **62**(4): 494-509.

Kalinina, J., K. Dutta, D. Ilghari, A. Beenken, R. Goetz, A. V. Eliseenkova, D. Cowburn and M. Mohammadi (2012). "The alternatively spliced acid box region plays a key role in FGF receptor autoinhibition." Structure **20**(1): 77-88.

Kan, Z., P. W. Garrett-Engele, J. M. Johnson and J. C. Castle (2005). "Evolutionarily conserved and diverged alternative splicing events show different expression and functional profiles." Nucleic Acids Res **33**(17): 5659-5666.

Kashima, T. and J. L. Manley (2003). "A negative element in SMN2 exon 7 inhibits splicing in spinal muscular atrophy." Nat.Genet. **34**(4): 460-463.

Kato, A., Y. Nagata and K. Todokoro (2004). "Delta-tubulin is a component of intercellular bridges and both the early and mature perinuclear rings during spermatogenesis." Dev Biol **269**(1): 196-205.

Kaufman, M. H. and J. B. L. Bard (1999). The anatomical basis of mouse development. San Diego, Academic Press.

Kempf, M., A. Clement, A. Faissner, G. Lee and R. Brandt (1996). "Tau binds to the distal axon early in development of polarity in a microtubule- and microfilament-dependent manner." J Neurosci **16**(18): 5583-5592.

Keren, H., G. Lev-Maor and G. Ast (2010). "Alternative splicing and evolution: diversification, exon definition and function." Nat Rev Genet **11**(5): 345-355.

Kondo, S., N. Yamamoto, T. Murakami, M. Okumura, A. Mayeda and K. Imaizumi (2004). "Tra2 beta, SF2/ASF and SRp30c modulate the function of an exonic splicing enhancer in exon 10 of tau pre-mRNA." Genes Cells **9**(2): 121-130.

Kornblihtt, A. R. (2005). "Promoter usage and alternative splicing." Curr Opin Cell Biol **17**(3): 262-268.

Kosik, K. S., L. D. Orecchio, S. Bakalis and R. L. Neve (1989). "Developmentally regulated expression of specific tau sequences." Neuron **2**(4): 1389-1397.

Krubitzer, L. and K. J. Huffman (2000). "Arealization of the neocortex in mammals: genetic and epigenetic contributions to the phenotype." Brain Behav Evol **55**(6): 322-335.

Laemmli, U. K. (1970). "Cleavage of structural proteins during the assembly of the head of bacteriophage T4." Nature **227**(5259): 680-685.

Lareau, L. F., M. Inada, R. E. Green, J. C. Wengrod and S. E. Brenner (2007). "Unproductive splicing of SR genes associated with highly conserved and ultraconserved DNA elements." Nature **446**(7138): 926-929.

Lefebvre, S., L. Burglen, S. Reboullet, O. Clermont, P. Burlet, L. Viollet, B. Benichou, C. Cruaud, P. Millasseau and M. Zeviani (1995). "Identification and characterization of a spinal muscular atrophy-determining gene." Cell **80**(1): 155-165.

Lever, M. A., J. P. Th'ng, X. Sun and M. J. Hendzel (2000). "Rapid exchange of histone H1.1 on chromatin in living human cells." Nature **408**(6814): 873-876.

Lewandoski, M. (2001). "Conditional control of gene expression in the mouse." Nat Rev Genet **2**(10): 743-755.

Li, S. J., Y. Qi, J. J. Zhao, Y. Li, X. Y. Liu, X. H. Chen and P. Xu (2013). "Characterization of nuclear localization signals (NLSs) and function of NLSs and phosphorylation of serine residues in subcellular and subnuclear localization of transformer-2beta (Tra2beta)." J Biol Chem **288**(13): 8898-8909.

Licatalosi, D. D. and R. B. Darnell (2006). "Splicing regulation in neurologic disease." Neuron **52**(1): 93-101.

Lim, S. R. and K. J. Hertel (2001). "Modulation of survival motor neuron pre-mRNA splicing by inhibition of alternative 3' splice site pairing." J Biol Chem **276**(48): 45476-45483.

Linde, M., W. M. Mulleners, E. P. Chronicle and D. C. McCrory (2013). "Valproate (valproic acid or sodium valproate or a combination of the two) for the prophylaxis of episodic migraine in adults." Cochrane Database Syst Rev **6**: CD010611.

Liu, F. and C. X. Gong (2008). "Tau exon 10 alternative splicing and tauopathies." Mol.Neurodegener. **3**: 8.

Llano, E., R. Gomez, C. Gutierrez-Caballero, Y. Herran, M. Sanchez-Martin, L. Vazquez-Quinones, T. Hernandez, E. de Alava, A. Cuadrado, J. L. Barbero, J. A. Suja and A. M. Pendas (2008). "Shugoshin-2 is essential for the completion of meiosis but not for mitotic cell division in mice." Genes Dev **22**(17): 2400-2413.

Lopez-Bendito, G. and Z. Molnar (2003). "Thalamocortical development: how are we going to get there?" Nat Rev Neurosci **4**(4): 276-289.

Lorson, C. L., E. Hahnen, E. J. Androphy and B. Wirth (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**(11): 6307-6311.

Lorson, C. L., J. Strasswimmer, J. M. Yao, J. D. Baleja, E. Hahnen, B. Wirth, T. Le, A. H. Burghes and E. J. Androphy (1998). "SMN oligomerization defect correlates with spinal muscular atrophy severity." Nat Genet **19**(1): 63-66.

Luco, R. F., M. Allo, I. E. Schor, A. R. Kornblihtt and T. Misteli (2011). "Epigenetics in alternative pre-mRNA splicing." Cell **144**(1): 16-26.

Manley, J. L. and A. R. Krainer (2010). "A rational nomenclature for serine/arginine-rich protein splicing factors (SR proteins)." Genes Dev **24**(11): 1073-1074.

Mansour, S. L., K. R. Thomas, C. X. Deng and M. R. Capecchi (1990). "Introduction of a lacZ reporter gene into the mouse int-2 locus by homologous recombination." Proc Natl Acad Sci U S A **87**(19): 7688-7692.

Markowitz, J. A., M. B. Tinkle and K. H. Fischbeck (2004). "Spinal muscular atrophy in the neonate." J Obstet Gynecol Neonatal Nurs **33**(1): 12-20.

Martinez, E., V. B. Palhan, A. Tjernberg, E. S. Lymar, A. M. Gamper, T. K. Kundu, B. T. Chait and R. G. Roeder (2001). "Human STAGA complex is a chromatin-acetylating transcription coactivator that interacts with pre-mRNA splicing and DNA damage-binding factors in vivo." Mol Cell Biol **21**(20): 6782-6795.

Mattox, W., M. J. Palmer and B. S. Baker (1990). "Alternative splicing of the sex determination gene transformer-2 is sex-specific in the germ line but not in the soma." Genes Dev **4**(5): 789-805.

McGlincy, N. J. and C. W. Smith (2008). "Alternative splicing resulting in nonsense-mediated mRNA decay: what is the meaning of nonsense?" Trends Biochem Sci **33**(8): 385-393.

Mende, Y. (2008). "Generierung und Charakterisierung von konditionalen knock-outs des Spleissfaktors Sfrs10 im Mausmodell." Dissertation.

Mende, Y., M. Jakubik, M. Riessland, F. Schoenen, K. Rossbach, A. Kleinridders, C. Kohler, T. Buch and B. Wirth (2010). "Deficiency of the splicing factor Sfrs10 results in early embryonic lethality in mice and has no impact on full-length SMN/Smn splicing." Hum Mol Genet **19**(11): 2154-2167.

Michlewski, G., J. R. Sanford and J. F. Caceres (2008). "The splicing factor SF2/ASF regulates translation initiation by enhancing phosphorylation of 4E-BP1." Mol Cell **30**(2): 179-189.

Mignone, J. L., V. Kukekov, A. S. Chiang, D. Steindler and G. Enikolopov (2004). "Neural stem and progenitor cells in nestin-GFP transgenic mice." J.Comp Neurol. **469**(3): 311-324.

Monani, U. R., C. L. Lorson, D. W. Parsons, T. W. Prior, E. J. Androphy, A. H. Burghes and J. D. McPherson (1999). "A single nucleotide difference that alters splicing patterns distinguishes the SMA gene SMN1 from the copy gene SMN2." Hum Mol Genet **8**(7): 1177-1183.

Mullis, K., F. Faloona, S. Scharf, R. Saiki, G. Horn and H. Erlich (1992). "Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. 1986." Biotechnology **24**: 17-27.

Nagoshi, R. N., M. McKeown, K. C. Burtis, J. M. Belote and B. S. Baker (1988). "The control of alternative splicing at genes regulating sexual differentiation in D. melanogaster." Cell **53**(2): 229-236.

Nagy, A. (2000). "Cre recombinase: the universal reagent for genome tailoring." Genesis **26**(2): 99-109.

Nayler, O., C. Cap and S. Stamm (1998). "Human transformer-2-beta gene (SFRS10): complete nucleotide sequence, chromosomal localization, and generation of a tissue-specific isoform." Genomics **53**(2): 191-202.

Neves, G., J. Zucker, M. Daly and A. Chess (2004). "Stochastic yet biased expression of multiple Dscam splice variants by individual cells." Nat Genet **36**(3): 240-246.

Nevins, J. R. (1992). "E2F: a link between the Rb tumor suppressor protein and viral oncoproteins." Science **258**(5081): 424-429.

Ngo, J. C., S. Chakrabarti, J. H. Ding, A. Velazquez-Dones, B. Nolen, B. E. Aubol, J. A. Adams, X. D. Fu and G. Ghosh (2005). "Interplay between SRPK and Clk/Sty kinases in phosphorylation of the splicing factor ASF/SF2 is regulated by a docking motif in ASF/SF2." Mol Cell **20**(1): 77-89.

Nilsen, T. W. and B. R. Graveley (2010). "Expansion of the eukaryotic proteome by alternative splicing." Nature **463**(7280): 457-463.
Noctor, S. C., A. C. Flint, T. A. Weissman, R. S. Dammerman and A. R. Kriegstein (2001). "Neurons derived from radial glial cells establish radial units in neocortex." Nature **409**(6821): 714-720.
Novoyatleva, T., B. Heinrich, Y. Tang, N. Benderska, M. E. Butchbach, C. L. Lorson, M. A. Lorson, C. Ben-Dov, P. Fehlbaum, L. Bracco, A. H. Burghes, M. Bollen and S. Stamm (2008). "Protein phosphatase 1 binds to the RNA recognition motif of several splicing factors and regulates alternative pre-mRNA processing." Hum Mol Genet **17**(1): 52-70.
O'Leary, D. D. and Y. Nakagawa (2002). "Patterning centers, regulatory genes and extrinsic mechanisms controlling arealization of the neocortex." Curr Opin Neurobiol **12**(1): 14-25.
O'Leary, D. D., B. L. Schlaggar and R. Tuttle (1994). "Specification of neocortical areas and thalamocortical connections." Annu Rev Neurosci **17**: 419-439.
Owens, G. K. (1995). "Regulation of differentiation of vascular smooth muscle cells." Physiol Rev **75**(3): 487-517.
Pan, Q., O. Shai, L. J. Lee, B. J. Frey and B. J. Blencowe (2008). "Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing." Nat Genet **40**(12): 1413-1415.
Pechnick, R. N., S. Zonis, K. Wawrowsky, J. Pourmorady and V. Chesnokova (2008). "p21Cip1 restricts neuronal proliferation in the subgranular zone of the dentate gyrus of the hippocampus." Proc Natl Acad Sci U S A **105**(4): 1358-1363.
Peitz, M., K. Pfannkuche, K. Rajewsky and F. Edenhofer (2002). "Ability of the hydrophobic FGF and basic TAT peptides to promote cellular uptake of recombinant Cre recombinase: a tool for efficient genetic engineering of mammalian genomes." Proc Natl Acad Sci U S A **99**(7): 4489-4494.
Pelisch, F., J. Gerez, J. Druker, I. E. Schor, M. J. Munoz, G. Risso, E. Petrillo, B. J. Westman, A. I. Lamond, E. Arzt and A. Srebrow (2010). "The serine/arginine-rich protein SF2/ASF regulates protein sumoylation." Proc Natl Acad Sci U S A **107**(37): 16119-16124.
Pihlajamaki, J., C. Lerin, P. Itkonen, T. Boes, T. Floss, J. Schroeder, F. Dearie, S. Crunkhorn, F. Burak, J. C. Jimenez-Chillaron, T. Kuulasmaa, P. Miettinen, P. J. Park, I. Nasser, Z. Zhao, Z. Zhang, Y. Xu, W. Wurst, H. Ren, A. J. Morris, S. Stamm, A. B. Goldfine, M. Laakso and M. E. Patti (2011). "Expression of the splicing factor gene SFRS10 is reduced in human obesity and contributes to enhanced lipogenesis." Cell Metab **14**(2): 208-218.
Plass, M., E. Agirre, D. Reyes, F. Camara and E. Eyras (2008). "Co-evolution of the branch site and SR proteins in eukaryotes." Trends Genet **24**(12): 590-594.
Pley, U. M., B. L. Hill, C. Alibert, F. M. Brodsky and P. Parham (1995). "The interaction of calmodulin with clathrin-coated vesicles, triskelions, and light chains. Localization of a binding site." J Biol Chem **270**(5): 2395-2402.
Qiu, T., L. Zhou, W. Zhu, T. Wang, J. Wang, Y. Shu and P. Liu (2013). "Effects of treatment with histone deacetylase inhibitors in solid tumors: a review based on 30 clinical trials." Future Oncol **9**(2): 255-269.
Query, C. C., M. J. Moore and P. A. Sharp (1994). "Branch nucleophile selection in pre-mRNA splicing: evidence for the bulged duplex model." Genes Dev **8**(5): 587-597.
Rakic, P. (1988). "Specification of cerebral cortical areas." Science **241**(4862): 170-176.
Ramirez, G., A. Alvarez, J. Garcia-Abreu, F. C. Gomes, V. Moura Neto and R. B. Maccioni (1999). "Regulatory roles of microtubule-associated proteins in neuronal morphogenesis. Involvement of the extracellular matrix." Braz J Med Biol Res **32**(5): 611-618.
Richardson, R. T., O. M. Alekseev, G. Grossman, E. E. Widgren, R. Thresher, E. J. Wagner, K. D. Sullivan, W. F. Marzluff and M. G. O'Rand (2006). "Nuclear autoantigenic sperm protein (NASP), a linker histone chaperone that is required for cell proliferation." J Biol Chem **281**(30): 21526-21534.

Richardson, R. T., I. N. Batova, E. E. Widgren, L. X. Zheng, M. Whitfield, W. F. Marzluff and M. G. O'Rand (2000). "Characterization of the histone H1-binding protein, NASP, as a cell cycle-regulated somatic protein." J Biol Chem **275**(39): 30378-30386.

Riessland, M., B. Ackermann, A. Forster, M. Jakubik, J. Hauke, L. Garbes, I. Fritzsche, Y. Mende, I. Blumcke, E. Hahnen and B. Wirth (2010). "SAHA ameliorates the SMA phenotype in two mouse models for spinal muscular atrophy." Hum Mol Genet **19**(8): 1492-1506.

Riessland, M., L. Brichta, E. Hahnen and B. Wirth (2006). "The benzamide M344, a novel histone deacetylase inhibitor, significantly increases SMN2 RNA/protein levels in spinal muscular atrophy cells." Hum.Genet. **120**(1): 101-110.

Risso, G., F. Pelisch, A. Quaglino, B. Pozzi and A. Srebrow (2012). "Regulating the regulators: serine/arginine-rich proteins under scrutiny." IUBMB Life **64**(10): 809-816.

Roberts, J. M., H. Ennajdaoui, C. Edmondson, B. Wirth, J. Sanford and B. Chen (2013). "Splicing factor TRA2B is required for neural progenitor survival." J Comp Neurol.

Roth, K. A. and C. D'Sa (2001). "Apoptosis and brain development." Ment Retard Dev Disabil Res Rev **7**(4): 261-266.

Ruskin, B., P. D. Zamore and M. R. Green (1988). "A factor, U2AF, is required for U2 snRNP binding and splicing complex assembly." Cell **52**(2): 207-219.

Saheki, Y. and P. De Camilli (2012). "Synaptic vesicle endocytosis." Cold Spring Harb Perspect Biol **4**(9): a005645.

Sancho, M., E. Diani, M. Beato and A. Jordan (2008). "Depletion of human histone H1 variants uncovers specific roles in gene expression and cell growth." PLoS Genet **4**(10): e1000227.

Sanford, J. R., P. Coutinho, J. A. Hackett, X. Wang, W. Ranahan and J. F. Caceres (2008). "Identification of nuclear and cytoplasmic mRNA targets for the shuttling protein SF2/ASF." PLoS.One. **3**(10): e3369.

Sauer, B. and N. Henderson (1989). "Cre-stimulated recombination at loxP-containing DNA sequences placed into the mammalian genome." Nucleic Acids Res **17**(1): 147-161.

Schwartz, S., E. Meshorer and G. Ast (2009). "Chromatin organization marks exon-intron structure." Nat Struct Mol Biol **16**(9): 990-995.

Segade, F., B. Hurle, E. Claudio, S. Ramos and P. S. Lazo (1996). "Molecular cloning of a mouse homologue for the Drosophila splicing regulator Tra2." FEBS Lett **387**(2-3): 152-156.

Seraphin, B. and M. Rosbash (1989). "Identification of functional U1 snRNA-pre-mRNA complexes committed to spliceosome assembly and splicing." Cell **59**(2): 349-358.

Shepard, P. J., E. A. Choi, A. Busch and K. J. Hertel (2011). "Efficient internal exon recognition depends on near equal contributions from the 3' and 5' splice sites." Nucleic Acids Res **39**(20): 8928-8937.

Shi, J., T. Zhang, C. Zhou, M. O. Chohan, X. Gu, J. Wegiel, J. Zhou, Y. W. Hwang, K. Iqbal, I. Grundke-Iqbal, C. X. Gong and F. Liu (2008). "Increased dosage of Dyrk1A alters alternative splicing factor (ASF)-regulated alternative splicing of tau in Down syndrome." J.Biol.Chem. **283**(42): 28660-28669.

Shukla, S. and S. A. Fisher (2008). "Tra2beta as a novel mediator of vascular smooth muscle diversification." Circ.Res. **103**(5): 485-492.

Simeone, A., M. Gulisano, D. Acampora, A. Stornaiuolo, M. Rambaldi and E. Boncinelli (1992). "Two vertebrate homeobox genes related to the Drosophila empty spiracles gene are expressed in the embryonic cerebral cortex." EMBO J **11**(7): 2541-2550.

Singh, N. N., E. J. Androphy and R. N. Singh (2004). "In vivo selection reveals combinatorial controls that define a critical exon in the spinal muscular atrophy genes." RNA **10**(8): 1291-1305.

Skordis, L. A., M. G. Dunckley, B. Yue, I. C. Eperon and F. Muntoni (2003). "Bifunctional antisense oligonucleotides provide a trans-acting splicing enhancer that stimulates SMN2 gene expression in patient fibroblasts." Proc Natl Acad Sci U S A **100**(7): 4114-4119.

Smrzka, O. W., N. Delgehyr and M. Bornens (2000). "Tissue-specific expression and subcellular localisation of mammalian delta-tubulin." Curr Biol **10**(7): 413-416.

Smyth, G. K. (2004). "Linear models and empirical bayes methods for assessing differential expression in microarray experiments." Stat Appl Genet Mol Biol **3**: Article3.
Sobue, K., K. Hayashi and W. Nishida (1999). "Expressional regulation of smooth muscle cell-specific genes in association with phenotypic modulation." Mol Cell Biochem **190**(1-2): 105-118.
Somlyo, A. P. and A. V. Somlyo (1994). "Signal transduction and regulation in smooth muscle." Nature **372**(6503): 231-236.
Staknis, D. and R. Reed (1994). "SR proteins promote the first specific recognition of Pre-mRNA and are present together with the U1 small nuclear ribonucleoprotein particle in a general splicing enhancer complex." Mol Cell Biol **14**(11): 7670-7682.
Stamm, S., D. Casper, J. Dinsmore, C. A. Kaufmann, J. Brosius and D. M. Helfman (1992). "Clathrin light chain B: gene structure and neuron-specific splicing." Nucleic Acids Res. **20**(19): 5097-5103.
Stamm, S., D. Casper, V. Hanson and D. M. Helfman (1999). "Regulation of the neuron-specific exon of clathrin light chain B." Brain Res.Mol.Brain Res. **64**(1): 108-118.
Stoilov, P., R. Daoud, O. Nayler and S. Stamm (2004). "Human tra2-beta1 autoregulates its protein concentration by influencing alternative splicing of its pre-mRNA." Hum.Mol.Genet. **13**(5): 509-524.
Tacke, R., M. Tohyama, S. Ogawa and J. L. Manley (1998). "Human Tra2 proteins are sequence-specific activators of pre-mRNA splicing." Cell **93**(1): 139-148.
Tardiff, D. F., S. A. Lacadie and M. Rosbash (2006). "A genome-wide analysis indicates that yeast pre-mRNA splicing is predominantly posttranscriptional." Mol Cell **24**(6): 917-929.
Tarn, W. Y. and J. A. Steitz (1994). "SR proteins can compensate for the loss of U1 snRNP functions in vitro." Genes Dev **8**(22): 2704-2717.
Tevethia, M. J., H. A. Lacko and A. Conn (1998). "Two regions of simian virus 40 large T-antigen independently extend the life span of primary C57BL/6 mouse embryo fibroblasts and cooperate in immortalization." Virology **243**(2): 303-312.
Toung, J. M., M. Morley, M. Li and V. G. Cheung (2011). "RNA-sequence analysis of human B-cells." Genome Res **21**(6): 991-998.
Towbin, H., T. Staehelin and J. Gordon (1979). "Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications." Proc Natl Acad Sci U S A **76**(9): 4350-4354.
Tronche, F., C. Kellendonk, O. Kretz, P. Gass, K. Anlag, P. C. Orban, R. Bock, R. Klein and G. Schutz (1999). "Disruption of the glucocorticoid receptor gene in the nervous system results in reduced anxiety." Nat Genet **23**(1): 99-103.
Tsuda, K., T. Someya, K. Kuwasako, M. Takahashi, F. He, S. Unzai, M. Inoue, T. Harada, S. Watanabe, T. Terada, N. Kobayashi, M. Shirouzu, T. Kigawa, A. Tanaka, S. Sugano, P. Guntert, S. Yokoyama and Y. Muto (2011). "Structural basis for the dual RNA-recognition modes of human Tra2-beta RRM." Nucleic Acids Res **39**(4): 1538-1553.
Turunen, J. J., E. H. Niemela, B. Verma and M. J. Frilander (2013). "The significant other: splicing by the minor spliceosome." Wiley Interdiscip Rev RNA **4**(1): 61-76.
Umen, J. G. and C. Guthrie (1995a). "A novel role for a U5 snRNP protein in 3' splice site selection." Genes Dev **9**(7): 855-868.
Umen, J. G. and C. Guthrie (1995b). "Prp16p, Slu7p, and Prp8p interact with the 3' splice site in two distinct stages during the second catalytic step of pre-mRNA splicing." RNA **1**(6): 584-597.
Venables, J. P., C. F. Bourgeois, C. Dalgliesh, L. Kister, J. Stevenin and D. J. Elliott (2005). "Up-regulation of the ubiquitous alternative splicing factor Tra2beta causes inclusion of a germ cell-specific exon." Hum Mol Genet **14**(16): 2289-2303.
Venables, J. P., D. J. Elliott, O. V. Makarova, E. M. Makarov, H. J. Cooke and I. C. Eperon (2000). "RBMY, a probable human spermatogenesis factor, and other hnRNP G proteins interact with Tra2beta and affect splicing." Hum.Mol.Genet. **9**(5): 685-694.

Verrotti, A., C. D'Egidio, A. Mohn, G. Coppola and F. Chiarelli (2011). "Weight gain following treatment with valproic acid: pathogenetic mechanisms and clinical implications." Obes Rev **12**(5): e32-43.

Voziyanov, Y., S. Pathania and M. Jayaram (1999). "A general model for site-specific recombination by the integrase family recombinases." Nucleic Acids Res **27**(4): 930-941.

Wang, J. Z. and F. Liu (2008). "Microtubule-associated protein tau in development, degeneration and protection of neurons." Prog.Neurobiol. **85**(2): 148-175.

Wang, Y., J. Wang, L. Gao, R. Lafyatis, S. Stamm and A. Andreadis (2005). "Tau exons 2 and 10, which are misregulated in neurodegenerative diseases, are partly regulated by silencers which bind a SRp30c.SRp55 complex that either recruits or antagonizes htra2beta1." J.Biol.Chem. **280**(14): 14230-14239.

Will, C. L. and R. Luhrmann (2011). "Spliceosome structure and function." Cold Spring Harb Perspect Biol **3**(7).

Wirth, B. (2000). "An update of the mutation spectrum of the survival motor neuron gene (SMN1) in autosomal recessive spinal muscular atrophy (SMA)." Hum.Mutat. **15**(3): 228-237.

Wirth, B., L. Brichta, B. Schrank, H. Lochmuller, S. Blick, A. Baasner and R. Heller (2006). "Mildly affected patients with spinal muscular atrophy are partially protected by an increased SMN2 copy number." Hum.Genet. **119**(4): 422-428.

Wong, A., S. Zhang, D. Mordue, J. M. Wu, Z. Zhang, Z. Darzynkiewicz, E. Y. Lee and M. Y. Lee (2013). "PDIP38 is translocated to the spliceosomes/nuclear speckles in response to UV-induced DNA damage and is required for UV-induced alternative splicing of MDM2." Cell Cycle **12**(19).

Wu, X. S., B. D. McNeil, J. Xu, J. Fan, L. Xue, E. Melicoff, R. Adachi, L. Bai and L. G. Wu (2009). "Ca(2+) and calmodulin initiate all forms of endocytosis during depolarization at a nerve terminal." Nat Neurosci **12**(8): 1003-1010.

Xiao, S. H. and J. L. Manley (1998). "Phosphorylation-dephosphorylation differentially affects activities of splicing factor ASF/SF2." EMBO J **17**(21): 6359-6367.

Yeo, G., D. Holste, G. Kreiman and C. B. Burge (2004). "Variation in alternative splicing across human tissues." Genome Biol **5**(10): R74.

Yin, F. X., G. P. Li, C. L. Bai, Y. Liu, Z. Y. Wei, C. G. Liang, T. D. Bunch and L. S. Zan (2013). "SGO1 maintains bovine meiotic and mitotic centromeric cohesions of sister chromatids and directly affects embryo development." PLoS One **8**(9): e73636.

Zahler, A. M. (1999). "Purification of SR protein splicing factors." Methods Mol Biol **118**: 419-432.

Zerres, K. and S. Rudnik-Schoneborn (1995). "Natural history in proximal spinal muscular atrophy. Clinical analysis of 445 patients and suggestions for a modification of existing classifications." Arch Neurol **52**(5): 518-523.

Zhang, Z., O. Kelemen, M. A. van Santen, S. M. Yelton, A. E. Wendlandt, V. M. Sviripa, M. Bollen, M. Beullens, H. Urlaub, R. Luhrmann, D. S. Watt and S. Stamm (2011). "Synthesis and characterization of pseudocantharidins, novel phosphatase modulators that promote the inclusion of exon 7 into the SMN (survival of motoneuron) pre-mRNA." J Biol Chem **286**(12): 10126-10136.

Zheng, S. and D. L. Black (2013). "Alternative pre-mRNA splicing in neurons: growing up and extending its reach." Trends Genet.

Zhou, J., Q. Yu and T. Zou (2008). "Alternative splicing of exon 10 in the tau gene as a target for treatment of tauopathies." BMC.Neurosci. **9 Suppl 2**: S10.

Zhu, J. Y., M. Abate, P. W. Rice and C. N. Cole (1991). "The ability of simian virus 40 large T antigen to immortalize primary mouse embryo fibroblasts cosegregates with its ability to bind to p53." J.Virol. **65**(12): 6872-6880.

Appendix

I. Top 100 ranking candidates obtained by whole transcriptome sequencing

Candidate hits of the initial whole transcriptome sequencing analysis were sorted by significance level. The 100 top ranking candidates are listed below. The chromosomal coordinates of each exon is given by CHR, EX_START and EX_END. Length gives the exon length in bp.

GENE	CHR	EX_START	EX_END	length bp	x-fold change	p-val
Appbp2	chr11	85027738	85027861	123	-3.24	2.5131E-05
Renbp	chrX	71174455	71174679	224	-4.11	3.9022E-05
Htt	chr5	35242421	35242576	155	1.52	5.5188E-05
Elavl2	chr4	90978264	90978354	90	17.30	6.4309E-05
Rpl34	chr3	130431250	130431769	519	-1.71	6.6274E-05
AI413582	chr17	27700718	27700961	243	1.25	1.0320E-04
Tm7sf2	chr19	6063324	6063446	122	1.28	1.2083E-04
Nasp	chr4	116286935	116287015	80	5.20	1.3034E-04
Zfp507	chr7	36572737	36572851	114	1.47	1.9921E-04
Bahcc1	chr11	120094261	120094421	160	-2.44	2.0975E-04
Slc43a2	chr11	75385817	75385914	97	1.80	2.4567E-04
Tanc2	chr11	105718711	105718942	231	-1.22	2.5942E-04
Zhx1	chr15	57878558	57880340	1782	-1.27	2.8100E-04
Gm6644	chr1	55500577	55500886	309	1.33	2.7701E-04
Gm6644	chr1	55501376	55501908	532	-1.22	2.7707E-04
Galk2	chr2	125713579	125713723	144	-1.67	2.8350E-04
Lin37	chr7	31342284	31342313	29	-3.17	2.8438E-04
Dlgap4	chr2	156526358	156527410	1052	1.19	2.9671E-04
Fbxl6	chr15	76367719	76367850	131	-1.25	3.1398E-04
Galk2	chr2	125713600	125713723	123	-1.75	3.6431E-04
Spnb4	chr7	28157116	28157394	278	1.46	3.7232E-04
Tank	chr2	61431134	61431265	131	2.01	4.7103E-04
Tank	chr2	61431154	61431265	111	2.01	4.7037E-04
AI413582	chr17	27700716	27700961	245	1.25	4.8913E-04
Lrba	chr3	86519435	86519602	167	1.63	4.9624E-04
Ypel1	chr16	17082650	17082787	137	3.86	5.0883E-04
Map4k1	chr7	29767073	29767283	210	-2.24	5.5680E-04
Eya2	chr2	165596534	165596762	228	-1.44	6.3657E-04
Adcy2	chr13	68796255	68796339	84	2.32	6.3663E-04
Robo3	chr9	37224282	37225087	805	-1.41	6.5205E-04
Pias3	chr3	96503916	96504000	84	-1.55	6.8355E-04

GENE	CHR	EX_START	EX_END	length bp	x-fold change	p-val
Aox4	chr1	58314059	58314154	95	1.62	6.8791E-04
Dlgap4	chr2	156526342	156527410	1068	1.19	6.8325E-04
Dlg1	chr16	31871761	31871934	173	-1.91	7.1727E-04
Dlg1	chr16	31871761	31871932	171	-1.91	7.1718E-04
2700029M09Rik	chr8	63378881	63379031	150	-2.71	7.5417E-04
1700021K19Rik	chr16	32844633	32844792	159	1.49	7.7601E-04
Slc2a3	chr6	122686597	122686759	162	1.31	8.0889E-04
Mtx1	chr3	89017735	89017784	49	3.37	8.3369E-04
Sbf2	chr7	117519086	117519342	256	-1.43	8.3037E-04
1700026L06Rik	chr2	28549144	28549346	202	2.19	8.3661E-04
Rasgrp3	chr17	75899399	75899572	173	1.86	8.5611E-04
Scfd2	chr5	74854330	74854667	337	3.52	8.9710E-04
Rims1	chr1	22475030	22475160	130	-1.94	9.0557E-04
Syvn1	chr19	6047085	6047128	43	-3.65	9.2388E-04
Araf	chrX	20429056	20429154	98	1.46	9.3155E-04
Gzf1	chr2	148510348	148510704	356	-1.24	9.4564E-04
Tnrc6c	chr11	117608491	117608610	119	1.42	9.4755E-04
Surf1	chr2	26770378	26770415	37	-7.11	9.8963E-04
Ap1g2	chr14	55717415	55717998	583	-1.83	9.9639E-04
Myo10	chr15	25643626	25643971	345	2.28	1.0292E-03
Plod1	chr4	147292889	147293002	113	1.30	1.0395E-03
Agxt2l2	chr11	51398259	51398381	122	4.06	1.0618E-03
Irf1	chr11	53587564	53587696	132	2.85	1.0620E-03
Adcy6	chr15	98425344	98425473	129	1.26	1.0577E-03
Tm2d1	chr4	98041611	98041704	93	5.76	1.0674E-03
Wbp1	chr6	83069862	83070038	176	1.25	1.1906E-03
Kdm4c	chr4	73944345	73944538	193	1.52	1.2451E-03
Wdfy1	chr1	79730428	79730448	20	1.86	1.2606E-03
Nadsyn1	chr7	150985696	150985827	131	1.63	1.2704E-03
Nrxn3	chr12	90088723	90089293	570	-1.63	1.2916E-03
Ftl1	chr7	52713335	52713653	318	1.24	1.3283E-03
Zfp511	chr7	147225225	147225344	119	-1.76	1.3308E-03
Snrpd2	chr7	19737825	19738084	259	-1.50	1.3403E-03
Eml4	chr17	83872990	83873088	98	1.97	1.3514E-03
Pnmal1	chr7	17547369	17547668	299	1.33	1.3520E-03
Supt6h	chr11	78059050	78059489	439	-1.65	1.3841E-03
BC024978	chr7	27987646	27987716	70	3.69	1.4037E-03
Shisa5	chr9	108957377	108957521	144	-1.69	1.5042E-03
Ppp1r10	chr17	36063439	36063518	79	1.88	1.5377E-03
Trim66	chr7	116604007	116604528	521	-1.82	1.5999E-03
Prosapip1	chr2	130462679	130463155	476	-1.52	1.6142E-03
Acin1	chr14	55263521	55263628	107	-2.01	1.6334E-03
Tgfbr1	chr4	47396621	47396866	245	1.72	1.6417E-03
Dlg1	chr16	31871761	31872347	586	-1.87	1.6731E-03
Pold1	chr7	51788574	51788671	97	-2.70	1.6840E-03

GENE	CHR	EX_START	EX_END	length bp	x-fold change	p-val
Akap9	chr5	4034748	4034867	119	1.82	1.7357E-03
Xpo6	chr7	133267886	133268045	159	1.31	1.7453E-03
Gm15952	chr3	35831967	35832089	122	2.70	1.7552E-03
Utp15	chr13	99032830	99032947	117	1.88	1.7505E-03
Pcbd2	chr13	55876728	55876808	80	1.87	1.7610E-03
Usp19	chr9	108401625	108401841	216	-1.25	1.7871E-03
Cnot6l	chr5	96509091	96509154	63	2.94	1.7875E-03
Kpna2	chr11	106852383	106852646	263	1.94	1.8029E-03
Tubgcp3	chr8	12630410	12632373	1963	-1.44	1.8288E-03
Pank2	chr2	131122009	131122044	35	2.59	1.8084E-03
Ncoa1	chr12	4270747	4270848	101	2.13	1.8306E-03
Nr6a1	chr2	38596063	38596127	64	-4.63	1.8323E-03
Parp8	chr13	117715324	117715394	70	-3.51	1.8357E-03
Pappa	chr4	64954353	64954466	113	3.20	1.8991E-03
Fnip2	chr3	79310129	79310200	71	1.70	1.9032E-03
Srgap3	chr6	112704565	112704667	102	-1.53	1.8775E-03
Rnd2	chr11	101329489	101329753	264	-2.78	1.9459E-03
Col6a1	chr10	76184148	76184276	128	-1.32	1.9570E-03
Cd2bp2	chr7	134337172	134337444	272	-1.47	1.9626E-03
Ppt2	chr17	34764111	34764470	359	-1.78	2.0033E-03
Phyh	chr2	4851759	4851940	181	1.31	1.9936E-03
Kcnab2	chr4	151851510	151851653	143	-3.18	2.0201E-03
Fam19a5	chr15	87511910	87512059	149	-1.15	2.0401E-03
Rictor	chr15	6745896	6746034	138	-2.09	2.0469E-03

II. Differentially expressed transcripts identified on mouse exon arrays

Of all identified differentially expressed transcripts only those with relative changes in expression larger than +1.5-fold or smaller than -1.5-fold and p-values smaller than 0.05 were considered and are listed below. Transcript ID refers to the transcript cluster ID of the NetAffx database for mouse WT 1.0 ST arrays.

Gene symbol	Transcript ID	Refseq	x-fold change	p-value
Defb11	6980933	NM_139221	2.0658	0.048522
Gstp2	6765276	NM_181796	1.8474	0.032396
Ccng1	6787474	NM_009831	1.8472	0.000038
Lilrb4	6767782	NM_013532	1.8272	0.025706
Lmbr1	6936930	NM_020295	1.8137	0.000605
Nfyc	6793170	NM_008692	1.7295	0.029643
Vsnl1	6798951	NM_012038	1.7261	0.028591
Sprr2e	6899429	NM_011471	1.7181	0.020085
Eda2r	7018366	NM_001161432	1.6672	0.000681
Olfr61	6965107	NM_146964	1.6627	0.018089
Sepp1	6828326	NM_009155	1.6604	0.024251
Krtap4-7	6791359	NM_029613	1.6428	0.046919
C3ar1	6957059	NM_009779	1.6267	0.001238
Mab21l2	6906620	NM_011839	1.5870	0.030646
Serpine2	6760009	NM_009255	1.5699	0.004763
D930031A20Rik	6964233	AK086471	1.5693	0.004062
Gpr50	7011861	NM_010340	1.5544	0.014725
Cdkn1a	6849595	NM_007669	1.5447	0.000405
Hmox1	6977260	NM_010442	1.5303	0.001385
Gpnmb	6946055	NM_053110	1.5184	0.011782
Fam96b	6984930	NM_026753	1.5115	0.012756
B230220N19Rik	7007728	AK045674	1.5112	0.026976
Glra1	6788423	NM_020492	1.5043	0.034485
Ctsc	6962376	NM_009982	1.5008	0.007041
Speer4f	6928962	NM_027609	-1.5027	0.037082
Dpy19l1	6994030	NM_172920	-1.5037	0.000029
Mpped1	6832276	NM_172610	-1.5039	0.004595
Sfrp1	6974663	NM_013834	-1.5084	0.022188
Nfix	6983873	NM_001081981	-1.5138	0.014852
Trim59	6905821	NM_025863	-1.5149	0.036341
Casc5	6880492	NM_029617	-1.5182	0.002964
H1f0	6831858	NM_008197	-1.5194	0.010765
Atad2	6835934	NM_027435	-1.5204	0.008777
E130114P18Rik	6923584	NR_015513	-1.5206	0.003672
Cdon	6988112	NM_021339	-1.5206	0.011740
Tmpo	6775984	NM_001080129	-1.5215	0.003863
Abcd2	6837935	NM_011994	-1.5223	0.000455
Igfbpl1	6921331	NM_018741	-1.5234	0.000428
Tcf4	6861850	NM_013685	-1.5247	0.002076
Zeb2	6886356	NM_015753	-1.5326	0.001918
Olfr510	6963311	NM_146311	-1.5354	0.033983
Sox2ot	6896671	NR_015580	-1.5355	0.001574
Cdkn2c	6924466	NM_007671	-1.5362	0.001138
Spc25	6887520	NM_025565	-1.5368	0.005265

Gene symbol	Transcript ID	Refseq	x-fold change	p-value
Zbtb20	6841019	NM_019778	-1.5642	0.011133
Hist1h1c	6805381	NM_015786	-1.5670	0.009790
Fezf2	6822729	NM_080433	-1.5751	0.007063
Bub1	6890715	NM_001113179	-1.5766	0.005811
Kif15	6993098	NM_010620	-1.5882	0.000537
D030056L22Rik	6868519	BC020125	-1.5906	0.008919
Pax6	6879741	NM_013627	-1.5982	0.000013
Smc4	6898241	NM_133786	-1.6021	0.013139
Nusap1	6880529	NM_133851	-1.6152	0.005198
Hist1h2ba	6811715	NM_175663	-1.6162	0.034395
Kif11	6869503	NM_010615	-1.6222	0.007155
Usp1	6915734	NM_146144	-1.6303	0.020022
Rlbp1	6968704	NM_020599	-1.6363	0.000674
Mir9-2	6808621	NR_029545	-1.6379	0.000071
Neurod1	6888114	NM_010894	-1.6418	0.000130
C79407	6801335	NM_172578	-1.6463	0.001367
Nfib	6922846	NM_001113209	-1.6486	0.007677
Cdk6	6928457	NM_009873	-1.6576	0.001684
Olfr1325	7012030	NM_146398	-1.6761	0.023110
Epha3	6846670	NM_010140	-1.6916	0.001027
Nr2e1	6773556	NM_152229	-1.7001	0.000058
G2e3	6795024	NM_001167963	-1.7030	0.000861
Usp-ps	6969968	NR_033576	-1.7222	0.009437
Tiam2	6848368	NM_011878	-1.7223	0.011583
Cdca7	6878053	NM_025866	-1.7261	0.001688
Ect2	6903753	NM_007900	-1.7433	0.000693
Tra2b	6844553	NM_009186	-1.7445	0.010315
Zfp238	6755494	NM_001012330	-1.7494	0.002712
Satb2	6758862	NM_139146	-1.7605	0.007654
Olfr433	6755294	NM_146717	-1.7657	0.017448
Ccnd2	6957263	NM_009829	-1.7997	0.000541
Top2a	6791298	NM_011623	-1.8076	0.009440
Eomes	6992849	NM_010136	-1.8263	0.000286
Foxg1	6794939	NM_001160112	-1.8445	0.000809
Hist1h1b	6811530	NM_020034	-1.8480	0.006732
Dlx1	6878021	NM_010053	-1.8513	0.000522
2810417H13Rik	6989974	NM_026515	-1.8574	0.018898
Nfib	6922857	NM_001113209	-1.8625	0.001273
Arhgap11a	6889894	NM_181416	-2.0788	0.000112
Neurod6	6953866	NM_009717	-2.0829	0.010316
2610017I09Rik	6758223	NR_027826	-2.1843	0.000693
Hist1h2bb	6805380	NM_175664	-2.5570	0.020253

III. Differentially spliced exons identified on exon arrays

Of all identified exons only those with relative changes in inclusion larger than +1.5-fold or smaller than -1.5-fold and p-values smaller than 0.05 were considered. Filtering strategies explained in chapter 5.6.2 and in Figure 27 were applied. Exons fulfilling the filtering conditions are listed below. Probeset ID refers to identification number of a group of probes representing an exon as it is listed in the NetAffx database for mouse WT 1.0 ST arrays.

Gene symbol	Probeset ID	exon length [bp]	AGAA freq.	x-fold change	p-value
SLU7	**4714947**	**114**	**13.47**	**-1.9191**	**0.02477**
KIF20B	4410288	156	13.13	-1.5488	0.00222
KIF2C	4603783	101	12.67	-1.8807	0.01001
MTHFD1L	4548959	42	12.19	1.5596	0.01384
SORBS3	4394107	87	11.77	1.6978	0.02896
PSMC3IP	4585989	146	10.52	-1.7234	0.00084
CENPK	4424232	151	10.17	-1.7252	0.00686
NRG1	**4394148**	**51**	**10.04**	**-1.7979**	**0.00613**
CCNB2	4583775	129	9.92	-1.6695	0.01335
ANKRD27	4374584	155	9.91	-1.5020	0.01201
TRANK1	4556735	104	9.85	1.5552	0.00786
SGOL2	**4596578**	**78**	**9.85**	**-2.6673**	**0.04683**
MLF1IP	4579279	79	9.72	-2.0832	0.01112
CENPE	4448452	159	9.66	-2.0459	0.04910
SLC26A4	4513897	55	9.31	-1.5611	0.02733
XPC	4540836	85	9.04	1.8936	0.02175
INTS6	4661791	143	8.95	-1.5575	0.02020
OFD1	4518495	174	8.83	-1.5665	0.00514
MEIS2	4356182	59	8.68	-1.5856	0.00315
SIAH1B	4542386	61	8.39	-1.6958	0.04213
OFD1	4349753	153	8.37	-1.6282	0.00951
CCDC73	4461583	184	8.35	-1.5059	0.03378
SGOL2	4709146	154	8.31	-1.7447	0.00706
DLX5	4410923	185	8.30	-1.7827	0.01353
KLHDC4	4507108	188	8.17	-1.9123	0.01836
MKI67	4310542	417	7.98	-1.5090	0.04443
PRR11	4320972	99	7.76	-1.5154	0.00922
CDCA5	4546320	66	7.76	-1.8696	0.01511
BRCA1	4554019	99	7.76	-1.8131	0.02303
HEG1	4578225	104	7.38	-2.1142	0.01104
ENOX2	4530030	108	7.11	-1.7849	0.00098
ARHGAP26	4327278	72	7.11	1.7719	0.01261
PTPN14	4746205	219	7.01	1.5709	0.03546
FAM107B	4727147	184	6.96	-1.6696	0.00154
SNX10	4597042	221	6.95	-1.7356	0.00079
MLH1	4544833	113	6.80	-1.5836	0.01376

Gene symbol	Probeset ID	exon length [bp]	AGAA freq.	x-fold change	p-value
SYTL4	4694836	76	6.74	2.3830	0.02549
GTF2F1	4486003	154	6.65	-1.7839	0.00200
SYNE2	4611825	156	6.56	-1.5839	0.02015
GRAMD1C	4489059	118	6.51	-2.0551	0.03330
D930015E06RIK	4532105	79	6.48	-1.7058	0.00727
APBA1	4609243	40	6.40	1.7743	0.01322
TTK	4707365	81	6.32	-2.0089	0.00007
CACNA2D3	4313575	81	6.32	1.5554	0.00522
ZBTB40	4670785	163	6.28	-1.6228	0.02769
OVCH2	4329814	204	6.27	-1.5507	0.03872
RFTN1	4652925	82	6.24	-2.0106	0.01390
KIF18A	4462409	412	6.21	-1.7284	0.03035
ELMO3	4712245	83	6.17	1.6138	0.04499
ASPM	4513114	167	6.13	-1.8145	0.00335
GRIK1	4641203	251	6.12	1.6661	0.01374
SPTY2D1	4327656	169	6.06	-1.5769	0.04401
ASPM	4726405	170	6.02	-1.8942	0.02406
RNASEH2B	4635673	43	5.95	-1.6809	0.00022
SYNE2	4360206	129	5.95	-1.7968	0.00087
TAF5	4530560	130	5.91	-1.7218	0.01403
GCA	**4430879**	**44**	**5.82**	**-2.5955**	**0.00346**
4930447C04RIK	4442257	176	5.82	1.6903	0.02751
LGSN	4338761	133	5.77	-1.5240	0.04271
KDM5A	4721789	182	5.63	-1.5596	0.00644
GPR64	4738054	46	5.57	1.6862	0.03385
PLA2R1	4349740	416	5.54	-1.6316	0.02991
SSH3	4647689	47	5.45	1.6535	0.00593
PLCH1	4393121	95	5.39	-1.5513	0.01493
HNF4A	4534336	95	5.39	1.6081	0.03012
SUV420H1	4495303	239	5.36	-1.5129	0.00050
NIN	4425864	191	5.36	-1.5214	0.04612
CNTNAP2	4361055	144	5.33	1.5486	0.00680
ECE1	4628587	96	5.33	2.7274	0.01753
ARHGAP19	4562132	192	5.33	-1.5435	0.03178
ZFP105	4626084	289	5.31	-1.5801	0.00169
EZR	4627884	97	5.28	-1.9445	0.00498
CEP55	4334451	195	5.25	-1.8542	0.01566
TFDP2	4509008	147	5.22	-1.7324	0.00205
RPRD1A	4466185	98	5.22	-2.1037	0.03573
ECE2	4421701	99	5.17	1.6365	0.00414
CDH17	4555128	99	5.17	1.5607	0.01284
ATP8B3	4579465	100	5.12	1.6908	0.00031
SF3B2	4555575	100	5.12	-1.5298	0.01763
HOXC8	4691504	1705	5.10	1.6038	0.04103
BCAP29	4720125	101	5.07	-1.5115	0.02851

Gene symbol	Probeset ID	exon length [bp]	AGAA freq.	x-fold change	p-value
FOXM1	4305232	152	5.05	-2.2088	0.00038
POLR2A	4458434	204	5.02	-1.5385	0.00020
RAPGEF4	4677994	51	5.02	2.1572	0.00355
SYNE1	4529049	153	5.02	1.6361	0.00800
SIDT1	4555772	102	5.02	1.5260	0.03208
ALLC	4462024	51	5.02	1.5400	0.04320
PRPF38B	4623817	2096	5.01	-1.5071	0.00399
GNAO1	4489479	154	4.99	-1.6702	0.00030
CPA3	4448551	103	4.97	-1.5479	0.02166
CCNA2	4670846	208	4.92	-1.7401	0.01545
ST18	**4501014**	**104**	**4.92**	**-2.1608**	**0.01835**
VEPH1	4493598	210	4.88	-2.1549	0.00510
DGKK	4405005	105	4.88	3.0980	0.04066
ALDOB	4374933	894	4.87	1.8427	0.01753
KIF18A	4403782	370	4.84	-1.6174	0.00304
KIF18A	4706655	370	4.84	-1.8986	0.02277
RCOR2	**4517066**	**53**	**4.83**	**-2.6669**	**0.00325**
DMC1	4547125	53	4.83	1.5978	0.01469
OFD1	4327484	107	4.79	-1.6517	0.00438
PLXNA2	4628626	269	4.76	-1.6269	0.00386
KCNIP4	4650677	108	4.74	1.5880	0.04691
PCLO	4319628	109	4.70	1.7506	0.01280
MBIP	4664789	219	4.68	-1.5382	0.03569
NES	4343695	5031	4.58	-1.7326	0.02448
NES	4434708	5031	4.58	-1.7789	0.04031
HCLS1	4603834	56	4.57	1.6110	0.00187
NTNG2	4678504	448	4.57	-1.5195	0.02573
MBOAT1	4523473	56	4.57	-1.6739	0.04129
RSBN1	4612612	674	4.56	-1.5093	0.02094
TPX2	4389535	169	4.54	-1.5292	0.00635
NUP107	4661967	113	4.53	-1.5011	0.01902
SPATA19	4390533	57	4.49	-1.5098	0.01334
NPR1	4573063	114	4.49	1.6498	0.03552
EFNB2	4406005	114	4.49	-1.5061	0.04358
1700025F22RIK	4728486	114	4.49	-1.6757	0.04824
TMEM176B	4395179	57	4.49	1.5149	0.04914
ASF1A	4345722	116	4.41	-1.6885	0.00595
TTC7	4445556	118	4.34	1.7040	0.01565
ARHGAP9	4457012	59	4.34	-1.5093	0.02528
MAP3K5	4572388	178	4.31	1.9921	0.04547
CENPJ	4563789	243	4.21	-1.8277	0.02338
CKAP2L	4655699	61	4.20	-1.9860	0.00757
LARS	4705431	369	4.16	-1.5188	0.00615
ANP32E	4479636	123	4.16	-1.6111	0.00770
SH3TC1	4498854	123	4.16	1.5202	0.02225

Gene symbol	Probeset ID	exon length [bp]	AGAA freq.	x-fold change	p-value
MID1	4452815	308	4.16	-2.1009	0.04247
TGM3	4653958	246	4.16	1.5098	0.04855
FRMD4A	4437044	248	4.13	-1.6302	0.00541
PLA2G7	4404775	124	4.13	1.6547	0.01115
AOX4	4498242	62	4.13	-1.9342	0.01310
CPNE8	4467664	62	4.13	1.5033	0.01780
DDX21	4350871	186	4.13	-1.5083	0.03626
EPHA4	4665920	125	4.10	-1.6541	0.00200
CP	4745390	376	4.09	1.9315	0.04285
SAP30	4647459	126	4.06	-1.5454	0.00052
PRKG1	4517208	63	4.06	1.5457	0.01848
DUB2A	4430289	1959	4.05	-1.7222	0.00944
DTNA	4512612	127	4.03	1.8568	0.01410
TRHDE	4403978	127	4.03	1.5086	0.01631
D2ERTD750E	4616609	127	4.03	-1.8003	0.01678
TNFRSF22	4437715	129	3.97	1.7165	0.01008
ZFP445	4554588	2075	3.95	-1.5427	0.01132
LRRIQ3	4734429	324	3.95	1.6544	0.02787
INPP4B	4592600	65	3.94	1.7812	0.00958
EPHA7	4519544	65	3.94	-3.0103	0.01593
SLC27A2	4491865	131	3.91	1.6788	0.01049
AIFM1	4696782	131	3.91	-1.6289	0.01333
ANO7	4550734	131	3.91	1.5821	0.04959
ZFYVE26	4575982	137	3.74	-1.5468	0.00893
GRM8	4744663	137	3.74	1.6249	0.02911
ZSCAN2	4339247	1581	3.72	-1.6928	0.02355
DDX27	4743790	277	3.70	1.5777	0.01701
RAD17	4393293	70	3.66	-1.6108	0.00622
BC088983	4696394	211	3.64	-1.5485	0.03231
D930015E06RIK	4371766	142	3.61	-1.6043	0.00239
STK30	4377299	71	3.61	-1.9443	0.03151
MCF2	4492513	143	3.58	-1.7141	0.00434
MYBL1	4392383	72	3.56	-1.5096	0.00721
APLF	4443727	72	3.56	-1.5235	0.01628
C3	4738818	72	3.56	1.5342	0.02861
TUBD1	**4335622**	**217**	**3.54**	**-1.6770**	**0.00326**
POLE	4598107	145	3.53	-1.5859	0.00031
NASP	4616314	652	3.53	-1.7145	0.02272
SPATA20	4662583	363	3.53	-1.5004	0.03368
DSN1	4321696	73	3.51	-1.6469	0.02201
ORC5	4657817	221	3.48	-1.7598	0.00165
ADAMTS12	4477489	368	3.48	1.5242	0.03622
PLA2G7	4346269	149	3.44	1.6709	0.00637
LIMA1	4368416	149	3.44	-1.5060	0.01213
9130014G24RIK	4392481	823	3.42	-1.6688	0.02775

Gene symbol	Probeset ID	exon length [bp]	AGAA freq.	x-fold change	p-value
IFITM3	4471047	374	3.42	1.6120	0.03210
CSTF3	4419823	75	3.41	-1.6309	0.00785
MTHFD1	4582739	75	3.41	-1.6544	0.02141
LRRC1	4596354	75	3.41	-1.6541	0.02152
RASGEF1B	4433006	152	3.37	-1.8033	0.04420
EME1	4653832	610	3.36	-1.6390	0.01435
SPAG5	4363783	1457	3.34	-1.5176	0.00669
RHPN2	4477249	77	3.32	1.5636	0.01444
CNST	4422191	629	3.26	3.1702	0.01696
SPOCK3	4515551	1892	3.25	1.7465	0.01574
SLC4A10	4657136	158	3.24	1.8556	0.01678
KDM2B	4328332	79	3.24	-1.5005	0.03221
H47	4448059	79	3.24	-1.5361	0.03398
ENAH	4681038	79	3.24	-1.5019	0.03725
TAF7L	4620888	79	3.24	2.9805	0.04087
PRSS16	4556206	635	3.23	1.5390	0.00734
SACS	4560545	480	3.20	1.6823	0.01206
HTR3A	4399364	161	3.18	-1.5998	0.02738
BC003331	4416326	647	3.17	-1.6843	0.03530
RAD9B	4377905	242	3.17	-1.5470	0.04848
TRA2A	4500206	811	3.16	1.5258	0.00208
SLCO1A5	4455480	655	3.13	-1.5728	0.00898
ST18	4492882	246	3.12	-1.6802	0.00237
PKHD1L1	4506571	166	3.08	-1.5203	0.00193
MPDZ	4511539	166	3.08	-1.5949	0.01005
PHF20L1	4706505	250	3.07	-1.6025	0.00013
	4483198	250	3.07	1.7295	0.02964
PHF20L1	4677219	250	3.07	-1.6314	0.03172
LCP2	4475995	84	3.05	1.5375	0.00598
ACE	4629811	168	3.05	1.5395	0.03204
4930504O13RIK	4516818	253	3.04	1.6303	0.03007
LHX3	4607122	169	3.03	1.7543	0.03022
MYOCD	4466868	2456	3.02	1.9269	0.03129
F730047E07RIK	4740068	170	3.01	-1.7486	0.00042
HTR3A	4595344	170	3.01	-1.6996	0.00124
SEMA5B	4544194	428	2.99	-1.7438	0.01615
PKHD1L1	4436724	172	2.98	-1.8760	0.00066
CBFA2T2	4370307	86	2.98	-1.5000	0.00329
THNSL2	4488806	172	2.98	1.5663	0.00453
PRKD3	**4344822**	**86**	**2.98**	**-2.8835**	**0.01817**
LRP1B	4393083	172	2.98	1.5616	0.02738
PAM	4549185	86	2.98	-1.5351	0.04407
NUT	4406073	2232	2.98	1.6016	0.04663
PABPC2	4515030	2586	2.97	1.7722	0.01478
GAS6	4650063	173	2.96	1.6455	0.00096

Gene symbol	Probeset ID	exon length [bp]	AGAA freq.	x-fold change	p-value
IQUB	4375567	173	2.96	1.9279	0.01136
POLE2	4625040	173	2.96	-1.6420	0.04054
LRP1B	4697372	2254	2.95	1.7630	0.04585
MANBA	4544437	87	2.94	1.8919	0.02124
HGF	4727408	175	2.93	-1.5910	0.04558
SPAG5	4490876	527	2.91	-1.7068	0.00012
VRK1	4651024	88	2.91	-1.9744	0.00096
ASPM	4629397	1408	2.91	-1.5253	0.01388
SLA	4388873	1937	2.91	-1.6603	0.03951
ASPM	4733897	1408	2.91	-1.5789	0.04961
EDN3	4465373	177	2.89	1.5972	0.00593
EZH2	**4681112**	**354**	**2.89**	**-1.6003**	**0.01694**
SLC7A3	4507707	178	2.88	1.9742	0.00225
AURKA	4612817	623	2.88	-1.5211	0.01405
CDK5RAP2	4705444	179	2.86	-1.7833	0.04189
RCBTB2	4429704	897	2.85	-1.5164	0.00013
SOX6	**4595680**	**90**	**2.84**	**-2.2940**	**0.00049**
HACE1	4562739	90	2.84	-1.5657	0.00115
MSH6	4329278	90	2.84	-1.6149	0.00803
KIF20A	4412443	90	2.84	-2.2795	0.01341
RACGAP1	4337642	90	2.84	-1.5121	0.02875
DHFR	4532630	4329	2.84	-1.6291	0.03442
ANGEL2	4487000	992	2.84	-1.5319	0.03675
4930583H14RIK	4349596	909	2.82	1.5726	0.00360
USP3	4688768	455	2.81	-1.5697	0.00535
MAP4K1	4584957	91	2.81	1.5749	0.01142
USP3	4689496	455	2.81	-1.7695	0.01173
PCNX	4570893	91	2.81	-1.9095	0.02433
ATP8A1	4565843	183	2.80	2.0702	0.04385
GPR158	4313474	4312	2.79	1.5638	0.00874
PTPRN	4744856	275	2.79	1.5346	0.01022
BRCC3	4538606	3318	2.78	-1.7085	0.00528
MBL2	4520356	184	2.78	-2.1122	0.00931
SRSF15	4382671	1288	2.78	-1.5258	0.01029
POLE	4488858	277	2.77	-1.7014	0.00163
PIH1D2	4617724	278	2.76	1.7806	0.02267
LIMA1	4531012	2596	2.76	-1.5865	0.02588
POLA1	4306555	93	2.75	-1.5848	0.02509
	4373942	3456	2.74	1.6091	0.00478
NDP	4581991	374	2.74	-2.0329	0.01218
FBXO38	4663537	94	2.72	-1.5067	0.00691
PKD1L2	4365239	94	2.72	-1.5440	0.01219
BOC	4735672	755	2.71	-1.5133	0.01100
ADK	4466769	755	2.71	-1.5576	0.04199
PRKX	4329571	96	2.67	-1.5539	0.00619

Gene symbol	Probeset ID	exon length [bp]	AGAA freq.	x-fold change	p-value
LCP1	4595737	96	2.67	1.5881	0.01335
DOK2	4503584	288	2.67	1.5278	0.02367
ELF2	4659413	96	2.67	-1.6847	0.02379
MCM2	4560441	192	2.67	-1.5402	0.04644
MS4A6D	4464067	578	2.66	1.5040	0.04085
EIF2AK4	4388293	386	2.65	-1.5193	0.00818
CTSG	4605174	193	2.65	2.0718	0.00880
KIF20A	4487970	965	2.65	-1.8862	0.03413
SYT9	4327356	2137	2.64	1.7922	0.01575
ADAM22	4442753	195	2.63	1.5670	0.03879
TMED5	4584874	3024	2.62	-1.7698	0.03078
TMTC4	4496065	1173	2.62	1.5565	0.04164
FLT1	4500095	98	2.61	1.7689	0.02734
SGOL1	4726910	792	2.59	-1.6013	0.00376
DMGDH	4415535	199	2.57	1.8070	0.02996
LCTL	4378199	200	2.56	1.6867	0.01358
ACVR1	4422845	100	2.56	-1.5270	0.01416
SORCS2	4423836	100	2.56	1.5393	0.02582
NLK	4505594	1808	2.55	-1.6462	0.00128
ZFP191	4354070	501	2.55	-1.5352	0.04879
	4488454	4240	2.54	-1.8165	0.00005
DDX11	4681249	101	2.53	-1.5091	0.02505
POLR2A	4484954	101	2.53	-1.5278	0.03929
MAD2L1	4363305	1221	2.52	-1.5584	0.00296
CCNG2	4718143	2750	2.51	-1.6380	0.00185
DMD	4641080	102	2.51	1.6966	0.00276
SLC2A3	4479978	102	2.51	1.5092	0.01243
TXNDC8	4367418	204	2.51	1.5540	0.03470
AKT3	4479233	206	2.49	-1.5602	0.01513
	4641319	2796	2.47	-1.7974	0.00247
FMO3	4489116	725	2.47	-2.0509	0.01331
	4486460	2796	2.47	-2.4199	0.02907
SPATS2L	4310354	207	2.47	1.5334	0.04660
RWDD3	4556722	1249	2.46	-2.1795	0.01520
GAB3	4515765	208	2.46	1.8418	0.03085
EPHX1	4361792	209	2.45	1.6537	0.03559
MSRA	4551929	105	2.44	-1.5724	0.01922
A930024E05RIK	4586822	738	2.43	-1.5868	0.00134
KIF4	4650460	106	2.42	-1.6806	0.00492
KLHL38	4559615	106	2.42	1.6933	0.03897
FOXRED1	4743516	213	2.40	-1.5259	0.02774
RNASEL	4685349	1607	2.39	-1.6046	0.01455
TLR7	4732269	3557	2.38	1.7712	0.01063
CRY1	4621702	108	2.37	1.5289	0.01750
STAG3	4659038	108	2.37	1.6741	0.02355

Gene symbol	Probeset ID	exon length [bp]	AGAA freq.	x-fold change	p-value
ZRSR2	4595108	325	2.36	-1.5142	0.01557
TRANK1	4476297	3050	2.35	1.7118	0.00523
PSMC6	4404885	109	2.35	-1.7306	0.04386
DLC1	4588622	219	2.34	-1.5196	0.02058
KLRA7	4519746	219	2.34	-2.2906	0.03110
LIPA	4351453	110	2.33	1.6418	0.01401
BMPER	4355300	110	2.33	-1.5710	0.01413
CACNB2	4415831	110	2.33	1.6237	0.01448
TCFAP2B	4539678	220	2.33	1.5376	0.01920
SLA	4547683	4066	2.33	-1.8236	0.02292
PHLDA3	4435898	659	2.33	2.1969	0.02718
ADCYAP1	4737637	1548	2.32	1.9273	0.00312
B230354K17RIK	4467887	1107	2.31	-1.6395	0.00303
CREM	4650584	1774	2.31	-1.5764	0.02471
C330027C09RIK	4360097	1335	2.30	-1.6455	0.00133
FAM5C	4518059	223	2.30	1.5546	0.02862
EPHA5	4434153	670	2.29	-1.5359	0.00949
LHX6	4593283	2126	2.29	-1.7716	0.01260
SDC2	4308175	2354	2.28	1.5059	0.01286
KLHL1	4524502	225	2.28	1.6689	0.02347
MCM2	4598865	113	2.27	-1.7781	0.00542
D630003M21RIK	4358288	903	2.27	1.9894	0.01714
D630003M21RIK	4486570	903	2.27	1.5524	0.01882
CALN1	4500272	113	2.27	1.5991	0.02083
THY1	4462459	1241	2.27	1.8176	0.02306
FUNDC1	4737894	339	2.27	-1.7487	0.02882
ASAP3	4715296	226	2.27	-1.6745	0.02993
WWC2	4556420	2702	2.27	-1.6076	0.03983
EPHA5	4733959	1019	2.26	-1.7866	0.01325
RWDD2A	4527541	227	2.26	1.8085	0.03397
ST8SIA6	4330824	6271	2.25	1.9448	0.03751
NFKBID	4443319	229	2.24	1.5006	0.01512
DHFR	4503067	116	2.21	-1.9853	0.00541
FAM20A	4566179	116	2.21	1.5554	0.00662
SLC16A9	4694299	233	2.20	1.5558	0.01308
GRM7	4706559	936	2.19	1.5800	0.02217
RFC4	4651099	353	2.18	-2.1348	0.01560
LRRTM4	4695212	1176	2.18	1.6368	0.03003
LETM1	4694406	2940	2.18	-1.8168	0.03278
SLC5A4B	4448388	1406	2.18	1.5124	0.04781
SLC15A2	4538095	1777	2.16	-1.7086	0.01396
CEP97	4369410	710	2.16	-1.5453	0.01415
PRPF3	4402326	119	2.15	-1.5628	0.03461
ADAMTS1	4380069	2274	2.14	1.6430	0.01942
A730037C10RIK	4311846	1671	2.14	1.5700	0.03940

Gene symbol	Probeset ID	exon length [bp]	AGAA freq.	x-fold change	p-value
NUAK1	4463507	120	2.13	-1.7956	0.00035
GRID2	4394768	2998	2.13	1.5192	0.00334
SLC24A2	4584831	1081	2.13	1.5315	0.01175
CENPE	4487174	120	2.13	-2.2500	0.03693
ZFP143	4472442	120	2.13	-1.5769	0.03861
	4632901	3139	2.12	-1.5679	0.00056
	4676952	3139	2.12	-1.7990	0.00857
SOX6	4669046	241	2.12	-1.5687	0.02169
AU021034	4348091	1936	2.12	1.6732	0.03054
ANKRD34B	4596043	3420	2.10	1.6112	0.00553
KCNE3	4590228	732	2.10	1.6163	0.03732
BMP10	4525025	365	2.10	1.6265	0.04137
ENPP5	4738159	864	2.07	1.5101	0.00739
TBC1D1	4683923	871	2.06	1.5337	0.00900
TMIE	4312239	1247	2.05	1.6778	0.00452
AMY1	4373243	126	2.03	1.6669	0.01001
SWAP70	4652190	2273	2.03	1.7312	0.01313
CDC25A	4340311	126	2.03	-1.5322	0.02588
N4BP3	4421228	255	2.01	1.5978	0.00580
PPRC1	4377908	255	2.01	1.9666	0.02180
RFX5	4367286	3568	2.01	-1.6865	0.02182
PMFBP1	4324734	255	2.01	1.7253	0.04228
POU3F2	4603318	6272	2.00	-1.5123	0.01185
CLSPN	4428061	128	2.00	-1.5184	0.03616
PDE1C	4388113	129	1.98	-1.6272	0.01852
CD1D1	4508745	517	1.98	-1.7724	0.02527
SUPT16H	4650061	130	1.97	-1.5031	0.01206
ZFHX3	4509945	6255	1.96	1.5191	0.00720
KIF18B	4595407	526	1.95	-1.6458	0.00976
GDA	4350979	3952	1.94	1.6029	0.02740
MRVI1	4348027	132	1.94	1.5131	0.03021
RET	4499628	264	1.94	1.5090	0.04439
	4318939	132	1.94	2.0020	0.04653
LRP1B	4663695	265	1.93	1.7006	0.02002
ADD3	4712355	2402	1.92	1.7947	0.00902
CASP1	**4305036**	**267**	**1.92**	**1.5430**	**0.02294**
KIF23	4592278	400	1.92	-1.5841	0.04619
ZBTB8A	4436521	803	1.91	-1.8701	0.02436
OLFR791	4502247	939	1.91	1.5085	0.04465
MCM10	4461183	269	1.90	-1.8054	0.00128
NAA16	4521127	135	1.90	-1.6204	0.01197
DNAHC5	4585508	135	1.90	1.5461	0.03136
PPP2R1B	4378289	135	1.90	-1.6646	0.03331
SP1	4623057	5680	1.89	-2.0571	0.00242
HSD17B3	4338295	406	1.89	-1.5467	0.01858

Gene symbol	Probeset ID	exon length [bp]	AGAA freq.	x-fold change	p-value
TMEM74	4401375	1353	1.89	-1.6337	0.02176
OLFR1046	4373246	951	1.88	-2.2074	0.01460
GRTP1	4408948	274	1.87	-1.6022	0.00733
PPARGC1B	4391082	1102	1.86	1.5293	0.02159
FOXC1	4514970	3990	1.86	1.5662	0.04646
SYT2	4353516	5675	1.85	1.8027	0.04042
YPEL1	4374592	2221	1.84	-1.6274	0.00694
GADD45G	4348997	418	1.84	-1.6410	0.01916
SULT1A1	4675456	140	1.83	1.6315	0.00056
OPRM1	4379014	420	1.83	-1.5537	0.00223
CP	4623278	140	1.83	1.5105	0.02035
ABCA4	4624016	140	1.83	1.7457	0.02959
SLC11A1	4734800	280	1.83	1.5680	0.04254
HEPACAM	4536924	282	1.82	1.7162	0.02700
HSPA12A	4344248	4248	1.81	1.5990	0.01971
IYD	4419840	848	1.81	1.5495	0.03836
ASPM	4344173	142	1.80	-1.9044	0.01217
PEO1	4632961	142	1.80	-1.5126	0.04437
HBP1	4539199	142	1.80	-1.5597	0.04723
IL1F9	4664033	143	1.79	-1.8881	0.03463
CCL24	4739850	144	1.78	-1.5646	0.00250
MOS	4530521	1449	1.77	1.6046	0.01875
2010300C02RIK	4508931	145	1.77	1.7034	0.03742
GTDC1	4429845	2188	1.76	-1.5118	0.00441
DLL1	4705503	727	1.76	-1.5596	0.01742
SLC6A13	4725872	440	1.75	1.6944	0.00621
COL4A1	4597318	1460	1.75	1.5878	0.02461
FARS2	4354425	292	1.75	-1.5153	0.04131
C330027C09RIK	4553525	148	1.73	-1.7838	0.00516
FANCI	4374601	443	1.73	-1.5347	0.02256
TRAF2	4544512	149	1.72	-1.5788	0.02430
MAP1LC3A	4439349	599	1.71	3.6272	0.00415
DLGAP5	4700686	1945	1.71	-1.6636	0.01610
DCT	4525957	300	1.71	-1.5898	0.03031
DLGAP5	4668991	1945	1.71	-1.9304	0.03748
CNTN6	4552356	151	1.70	1.8251	0.03341
TRANK1	4653312	1517	1.69	1.6748	0.00202
USP3	4499487	3942	1.69	-1.8275	0.01850
CGA	4337438	305	1.68	1.8465	0.01683
RAE1	4380864	457	1.68	-1.8731	0.03900
E330009J07RIK	4521937	154	1.66	-1.5049	0.00230
ITIH2	4569424	309	1.66	1.6252	0.00294
D930015E06RIK	4620859	154	1.66	-1.6091	0.01076
NEUROD2	4741421	2927	1.66	-1.9637	0.01450
SYNJ2	4389452	2007	1.66	1.7214	0.04156

Gene symbol	Probeset ID	exon length [bp]	AGAA freq.	x-fold change	p-value
MYRIP	4643639	626	1.64	1.5798	0.00228
KCNC4	4508651	937	1.64	1.6273	0.00580
HNF4A	4631354	156	1.64	1.7118	0.00793
EGR1	4553204	2520	1.63	1.7389	0.00034
EGR1	4472382	2520	1.63	1.5447	0.02558
LIMA1	4546203	2821	1.63	-1.7192	0.02617
ADAMTS10	4744306	157	1.63	-1.9474	0.03812
CENPA	4515076	790	1.62	-1.5357	0.01323
CNTN2	4378107	6213	1.61	-1.5483	0.00218
DNAJC11	4583902	1588	1.61	-1.8878	0.01322
CNTN2	4373088	6213	1.61	-1.6561	0.01422
TBR1	4497886	159	1.61	-1.5634	0.03035
SC4MOL	4654742	1273	1.61	-1.5144	0.03061
PDE4B	4539972	479	1.60	1.5419	0.03749
RACGAP1	4513235	969	1.59	-1.5917	0.00378
PLXDC2	4671369	161	1.59	1.5397	0.00624
FRMD4B	4589852	1768	1.59	-1.7008	0.00797
ZFP784	4503608	1129	1.59	-1.5264	0.01001
BLM	4586409	161	1.59	-1.9113	0.01449
RCOR2	4355418	162	1.58	-1.8296	0.00145
CDC45	4355843	162	1.58	-1.5256	0.01734
PRKD1	4549513	162	1.58	-1.5065	0.02775
PLEKHG2	4531289	162	1.58	-1.7737	0.04283
ALG13	4598957	163	1.57	-1.5787	0.00378
MS4A6D	4410250	163	1.57	1.6310	0.02211
CHAF1B	4625459	163	1.57	-1.9919	0.02430
INPP5B	4455005	163	1.57	1.6600	0.04838
GM5607	4582719	987	1.56	-1.7349	0.00017
CELSR1	4438558	164	1.56	-2.1430	0.00204
FOXM1	4339271	3117	1.56	-1.7902	0.00257
SPAG5	4731535	330	1.55	-1.6863	0.00747
	4683721	2659	1.54	1.6527	0.00118
SLC43A1	4414094	832	1.54	-1.5599	0.01719
ITGA2	4698122	166	1.54	1.5499	0.02519
RNF144A	4569727	3993	1.54	-1.5186	0.03155
KIF20A	4527819	166	1.54	-1.5349	0.04300
HSD3B3	4650071	1172	1.53	1.5317	0.00859
2900005J15RIK	4737400	2025	1.52	1.5068	0.01412
IL1R1	4545538	168	1.52	1.5133	0.03741
RCAN2	4372917	2544	1.51	1.6379	0.02219

Erklärung

Ich versichere, dass ich die von mir vorgelegte Dissertation selbstständig angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen -, die anderen Werken im Wortlaut oder im Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; dass diese Dissertation noch an keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie – abgesehen von unten angegebenen Teilpublikationen – noch nicht veröffentlicht worden ist sowie, dass ich eine solche Veröffentlichung vor Abschluss des Promotionsverfahrens nicht vornehmen werde.
Die Bestimmungen dieser Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Prof. Dr. rer. nat. Brunhilde Wirth und Prof. Dr. Elena Rugarli betreut und in der Arbeitsgruppe von Prof. Dr. rer. nat. Brunhilde Wirth durchgeführt worden.

Teilpublikationen sind in Kapitel 9 angegeben.

_____ _____

Ort, Datum Markus Storbeck

Lebenslauf