Biologie

Transgenic mouse models for studying BC1 and BC200 RNA transport and purification of their RNPs, and SINEs as tools for reconstructing rodent phylogeny

Inaugural-Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften im Fachbereich Biologie
der Mathematisch-Naturwissenschaftlichen Fakultät
der Westfälischen Wilhelms-Universität Münster

vorgelegt von

Manoj Kumar Sadasivuni

aus Visakhapatnam, Indien
im Juni 2009

Dekan : Prof. Dr. Christian Klämbt

Erster Gutacher : Prof. Dr. Jürgen Brosius

Zweiter Gutacher : Prof. Dr. Volker Gerke

Tag der mündliche Prüfung: ……29/05/2009……
(wird nach der Prüfung handschriftlich eingestzt)

Tag der Promotion: ……..…09/06/2009……
(wird nach der Prüfung handschriftlich eingestzt)

III Results

Establishing and characterization of the tagged BC RNAs in transgenic mice for localization and purification studies

Hu proteins are the potential proteins partners of BC RNP complexes

Application of SINEs in rodent phylogeny

IV Discussion

Abbreviations:

ABC complex	Avidin : Biotinylated enzyme Complex
BC1 RNA	Brain specific Cytoplasmic RNA 1
BC200 RNA	Brain specific Cytoplasmic RNA of 200 nt long
BC RNAs	BC1 RNA and BC200 RNA collectively termed as BC RNAs
BGH	bovine growth hormone
BSA	bovine serum albumin
CA	cornu ammonis (Amun's horn)
CAM	calmodulin sepharose
CBB	Coomassie Brilliant Blue R-250 and G-250 Dyes
CBP	Calmodulin Binding Peptide
CMV	Human cytomegalovirus
CPAL	Conserved Presence/Absence Loci
DG	dentate gyrus
DAB	3,3'-diamino-benzidine
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
EGTA	ethylene glycol tetraacetic acid
eIF4	eukaryotic initiation factor 4
FAM	fossil Alu monomer
FLAM	free left Alu monomer
FRAM	free right Alu monomer
FMRP	fragile X mental retardation protein
GAP-43	growth associated protein of 43 kDa
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
ID element	a SINE family found in rodents, termed Identifier-elements
IgG	Immunoglobulin domain
IPTG	Isopropyl-β-D-thiogalactopyranoside
IR	immunoreactivity
LB	Luria-Bertani broth
NaF	sodium fluoride
npcRNAs	non-protein coding RNAs
NGS	normal goat serum
PABP	Poly(A)-binding protein
PBS	phosphate-buffered saline
PFA	paraformaldehyde
PMSF	phenylmethanesulfonylfluoride
PNK	polynucleotide kinase
PP7	*Pseudomonas aeruginosa* phage 7
PVDF	polyvinylidene difluoride
rRNA	ribosomal RNA
SRP RNA	signal recognition particle RNA
SINEs	Short INterspersed Elements
$tRNA^{Ala}$	Alanine transfer RNA
TEV	tobacco etch virus
$^{Tg/Tg}$	Transgene homozygous
$^{Tg/+}$	Transgene heterozygous
$^{+/+}$	wild type homozygous
$^{+/-}$	BC1 KO heterozygous
$^{-/-}$	BC1 KO homozygous
TB	terrific broth

Acknowledgements

I wish to express my deep sense of gratitude to my supervisor, Prof. Dr. Jürgen Brosius, Director of the Institute of Experimental Pathology, for introducing me into the field of BC1 and BC200 RNAs and his valuable guidance and constant personal support throughout the course of my work. Besides his substantial scientific knowledge and experience, I highly appreciate all those stimulating discussions with him gave me strength to keep on solving problems faced during my doctoral studies and learned to be more critical towards the work.

I would like to thank Dr. Boris V. Skryabin for generating transgenic mice and helpful suggestions on BC RNAs work. My thanks goes to his group members as well, Sergej Handel, Jens, Thomas Robeck, Leonid V Gubar (for few Southern blots) and Birte Seeger.

I would like to thank Dr. Juergen Schmitz for introducing me into the field of rodent phylogeny and his personal interest towards the work and his constant support in solving the problems faced during the work. I also like to thank Gennady Churakov regarding the rodent phylogeny work and for his friendly help for many things.

I would like to thank Tasneem Khanam, Muddashetty R. S., for helping me at the initial stages of my work. I would like to thank Timofey S Rozhdestvensky, Carsten, A Rabbe and Martin Kiefmann for their helpful suggestions. I take this opportunity to thank all my colleagues from Institute of Experimental Pathology. I would like to thank Marsha Bundman very much for her editorial assistance regarding abstract and summary.

I would like to thank my friends, Aswani Kumar, Phaneeswara Rao, Suresh C. V., who encouraged me throught out my stay in Germany.

I would also like to thank my research advisory committee members Prof. Dr. Volker Gerke and Prof. Dr. Martin Bähler, for their willingness in mentoring me.

Last, but the most important, I wish to thank my parents Mohan Rao, Anasuya, my brothers Kishore Kumar, Rajesh Kumar and my wife Swathy who have been my emotional strength all along.

Abstract

Rodent-specific BC1 RNA and primate-specific BC200 RNA, here referred to collectively as "BC RNAs", are small non-protein coding RNAs transcribed by RNA polymerase III and generated by retroposition of $tRNA^{Ala}$ and 7SL RNA, respectively. Despite their different evolutionary origins, both RNAs share many features, thus they are considered to be functional analogues. Both RNAs are preferentially expressed in neurons and are transported efficiently to the dendrites. They are also involved in translational regulation. However, their precise cellular functions have yet to be discovered. In the cell they are bound to proteins and form ribonucleoprotein (RNP) complexes. To date, over a dozen proteins are reported to bind to these RNAs, but most of them are highly debated and controversial. To determine the precise functions of the BC RNAs, it is essential to purify and characterize their RNPs. Many biochemical purification methods were previously used in an attempt to purify the BC RNP complexes, but most of them were not successful in isolating pure or entire complexes. Thus, we developed an RNA tag-based affinity purification procedure, whereby both BC RNAs were tagged with MS2 RNA motifs in combination with the MS2 coat protein fused to the tandem affinity purification (TAP) tag to purify the endogenously assembled BC RNP complexes using an affinity matrix. The TAP tag is a two-step affinity purification method, whereby contaminants can be eliminated during the purification, and provides a higher purification compared to other conventional biochemical purifications. Preliminary *ex vivo* experiments were carried out in NIH 3T3 cells; however, as these are not neuro-specific cells, such studies might have missed the neurospecific protein partners. In addition to purification, the MS2 RNA tag can also be used to study transport. Thus, I established two different transgenic mouse models, BC1MS2B+MS2TAP and BC200MS2+MS2TAP, to study the purification and transport of both BC RNAs from mouse brain. Initial purification and transport studies provided a hope that both objectives may be achieved in the *in vivo* situation.

The second part of my thesis took advantage of the fact that both BC1 RNA and BC200 RNA are related to Short INterspersed Elements (SINEs). BC1 RNA is a master gene for the rodent-specific ID repetitive elements yielding SINEs. SINEs are abundant in all rodent genomes and have been used as powerful tools for examining rodent evolution. In this study, SINE elements were used as effective markers for reconstructing the phylogeny of rodents. Computerized databases from *Mus musculus* and *Rattus norvegicus*, published human genomic sequences, and preliminary trace sequences from *Dipodomys ordii*, *Cavia porcellus,* and *Spermophilus tridecemlineatus* were used to screen for phylogenetically informative SINE insertions [human and lagomorph (*Oryctolagus cunlculus*) sequences were used as outgroups]. Using different computational strategies for selecting phylogenetically informative loci, we were able to resolve some of the unsettled questions in the rodent phylogenetic tree. Twenty-two independent diagnostic SINE markers were found in 17 different loci. In addition to those already known, we found five new SINE markers that confirm the monophyly of rodents. The most important findings were two diagnostic SINE

markers that group Sciuridae and Aplodontidae in the Squirrel-related clade, two markers that group Caviomorpha and Phiomorpha species together within the Ctenohystrica clade and four new markers that group members of the superfamily Geomyoidea. The remaining nine markers confirmed previously known groupings within the three main rodent clades.

INTRODUCTION

The mouse genome sequence revealed that less than 2% of genome codes for proteins, while 37.5% of the genome is composed of repetitive elements (Waterston et al., 2002, Venter et al., 2001). Among them B1, B2, B4 and ID elements are the most abundant repetitive elements, which amount to more than 1 million copies representing 7.8% of the genome. Retroposition is the mechanism responsible for amplification of these elements. Retroposed elements belong to the Short INterspersed Elements (SINEs) family. SINEs are derived from small RNA templates with an internal RNA polymerase III promoter, especially 7SL RNA or tRNAs (Okada, 1991, Ohshima and Okada, 1994). SINE repeats, representing the major class of non-autonomous retroposons, are usually about 75-600 nucleotides long. SINEs are probably best defined as short DNA sequences arising from reverse-transcribed RNA molecules, originally transcribed mostly by RNA polymerase III (Singer, 1990). Most of the SINEs are transcriptionally not active, but some SINEs are actively transcribed and participate in retroposition are termed master genes (Deininger et al., 1992). Alu and B1 elements are derived from 7SL RNA [a component of signal recognition particle, SRP RNA], while ID and B2 elements are derived from tRNA genes. SINEs can be used to resolve the phylogenetic relationships among rodents, this will be discussed in the second part of my thesis.

BC1 RNA is a master gene for generating SINEs namely ID elements. BC200 is a transcribed monomeric Alu element derived from 7SL RNA. Both are small non-protein coding RNAs (npcRNAs) (Brosius, 1999, Wang et al., 2002) transcribed by RNA polymerase III, (Martignetti and Brosius, 1993a) complexed with proteins that form ribonucleoprotein particles [RNPs] in the cell (Chen et al., 1997a, Kremerskothen et al., 1998b, Muddashetty et al., 2002). Rodent-specific BC1 RNA/RNP and primate-specific BC200 RNA/ RNP, here referred to collectively as “BC RNAs/RNPs”. Unlike the majority of small stable RNAs, these cytoplasmic RNAs are almost exclusively expressed in neurons of brain and developmentally regulated. In neurons, both the RNAs/ RNPs are efficiently transported into distal parts of dendrites (Tiedge et al., 1993, Tiedge et al., 1991) and their expression during development coincides with synapse formation and is involved in synaptic activity. The unique expression pattern, both temporally and spatially and their ability to inhibit translation *in vitro* and *ex vivo* [transfected cells] makes them very interesting molecules to study (Kondrashov et al., 2005, Wang et al., 2005, Wang et al., 2002). Both RNAs share similar domain organization, the tRNA-related ID region for BC1 RNA, and the 5' repetitive domain, representing the Alu domain in BC200 RNA, a central A-rich region and a unique region at the 3' end (Martignetti and Brosius, 1993a). In order to fully understand and experimentally test the role of both BC RNAs in neuronal dendrites, it is important to identify all the proteins that are associated with their respective RNPs. Studies of RNP components [RNA-protein interactions] are far from technically simple. Many proteins bind to cellular RNAs, usually in a dynamic way that reflects changes in cellular metabolism and regulation or function. These proteins bind and remain bound with varying affinities that can be influenced by other constituents of the RNP complex. To date, there are several *in vitro* and *in vivo* methods available to

study RNA-protein interactions. *In vitro* techniques have been instrumental for the precise characterization of RNA and protein interaction sites. Early studies on RNA-protein interactions employed filter binding assays or electrophoretic mobility shift assays, in which a known RNA-binding protein was incubated under equilibrium conditions with the radio labeled RNA harboring a potential binding sequence. Methods such as UV cross-linking (Dreyfuss et al., 1984), have been effective in identifying the interacting components as well as in precisely pinpointing the site of RNA-protein interactions *in vitro*. Recently RNA affinity tags (Srisawat and Engelke, 2001) were found to be useful for studying RNPs. Although *in vitro* methods are valuable in understanding the structure/function relationship of RNP complexes, they are prone to artifacts and may fail to reflect physiological interactions, which can be addressed only by *in vivo* systems.

To date, over a dozen proteins are reported to be the components of BC1 or BC200 RNP (Kondrashov et al., 2005). Most of them are *in vitro* or interpretational artifacts except poly (A)-binding protein (PABP) for both BC RNAs (Muddashetty et al., 2002, West et al., 2002) and SRP9/14 for BC200 RNA (Skryabin et al., 1998, Bovia et al., 1997). Thus, we have developed an RNA tag-based affinity purification procedure, to our knowledge, this is an advanced affinity purification strategy of endogenously assembled either BC1 or BC200 RNPs from the mouse brain extract. We believe this method may be useful to overcome the aforementioned general biochemical purification problems, where intact and complete purification of BC RNPs were not possible.

In our method bacteriophage MS2 translational repressor or RNA motif, (Horn et al., 2004, Johansson et al., 1998, van den Worm et al., 1998) were tagged to the BC RNAs to study the RNA-protein interactions and the transport of these BC RNAs. The 19 nt stem loop MS2 motif (stem-loop operator) has a high *in vitro* affinity to MS2 coat protein. The MS2 coat protein had been fused to the Tandem Affinity Purification [TAP] tag (Rigaut et al., 1999), where the BC MS2/MS2TAP RNP complexes will be immobilized on an affinity matrix for a two step affinity purification. In this MS2/MS2 coat protein (MS2 RNA motif and its MS2 coat protein) complex have been conceived as a handle for the purification of BC RNP complexes and the whole complex represents (reflects) an RNA tag based affinity purification (Hogg and Collins, 2007).

1.0 Origin and evolutionary significance of BC RNAs

In general, retroposition is an ancient process that has been instrumental in the transformation of an exclusively RNA/RNP world into the modern world with three major macromolecules DNA-, RNA-, Protein (Brosius and Tiedge, 1995). The process of retroposition is mediated by the enzyme reverse transcriptase, which converts RNA to DNA. The event is rather random whereby any cellular RNAs can serve as a template for reverse transcription and the resulting cDNA copies are integrated into the genome more or less at any position.

Neuronal BC1 RNA has been generated by retroposition of $tRNA^{Ala}$ (DeChiara and Brosius, 1987). The cDNA copy of $tRNA^{Ala}$ integrated fortuitously into a locus [chromosome 7 between Fgf3 and Itgal (Oraov) in a common ancestor of all rodents] (Taylor et al., 1997) where flanking promoter elements were available in conjuction with internal promoter elements (box A+B'). (Martignetti and Brosius, 1993b) The 5' end of BC1 RNA has ~80% sequence similarity with $tRNA^{Ala}$, however this domain does not fold like a cloverleaf structure of tRNA but rather forms a stable stem/loop **(Fig 1.0.)** (Rozhdestvensky et al., 2001). The remainder of BC1 RNA consists of the central A-rich region and the unique region.The BC1 gene is present in all rodent species examined so far. BC1 RNA possesses very specific, specialized expression patterns in the brain. Furthermore, the expression of BC1 RNA is restricted to the mammalian order Rodentia. Its sequence in distantly related rodent genomes is highly conserved compared to its flanking regions and it is the first known RNA polymerase III transcript expressed almost exclusively in nerve cells, where it is efficiently transported up to the distal parts of dendrites (Tiedge et al., 1991). Another important finding is that BC1 RNA also serves as master gene for ID1 SINE elements (Kim et al., 1994). Other ID elements, including ID2-4 presumably are retroposed from a few transcribed ID-related RNAs (Kim et al., 1994, Shen et al., 1997).

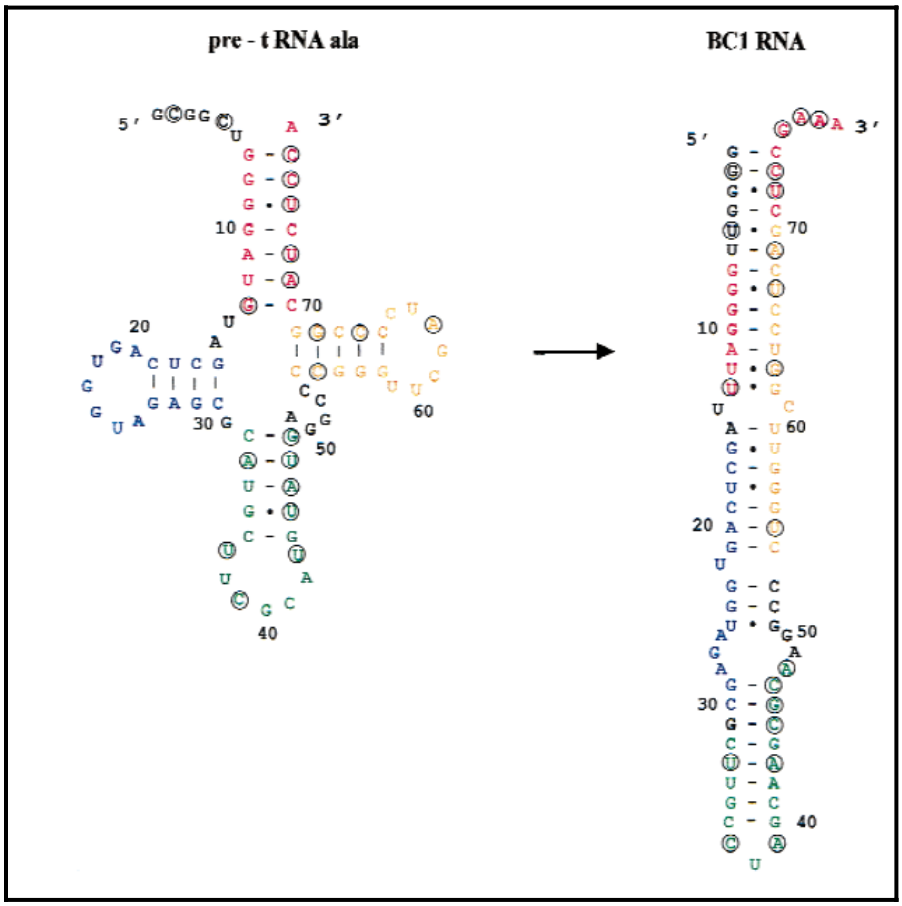


Fig 1.0 Origin of BC1 RNA gene the pre-$tRNA^{Ala}$ is the precursor of BC1 RNA gene. Base changes leading from its putative $tRNA^{Ala}$ precursor are indicated by circles to facilitate comparison. BC1 RNA gene serves as master gene for ID repetitive elements. The figure was adapted from (Rozhdestvensky et al., 2001).

BC200 RNA, which is present in anthropoid primates, located on human chromosome 2p16 (Basile et al., 1998), but the human genome sequence [build 36.3] places it on 2p21 between TACSTD1 (tumor-associated calcium signal transducer 1)/MIRN559 and CALM2 [calmodulin 2]/LSL [leptin, serum level] has been exapted from a monomeric Alu element. The progenitor of this Alu family of SINEs began with a 7SL RNA which is derived from a deletion of 141 nucleotides in the 7SL RNA gene starting from position 98, followed by a few point mutations in the "B-box" of the RNA polymerase III promoter. The resulting master gene given rise to a SINE called the fossil Alu monomer [FAM] (Quentin, 1992). It amplified to a number of copies that are still detectable in the human genome. The FAM most probably given rise to another SINE, the free left Alu monomer [FLAM (PB1)] by the deletion of another 42 base pairs starting from the position 84 (Kriegs et al., 2007). The FLAM probably is the common ancestor for the BC200 RNA gene in primates, via intermediate ancestral elements to the $4.5S_H$ gene as well as to the B1 SINEs in rodents. The FLAM fused to another FAM-derived element, the free right Alu monomer [FRAM] that given rise to the Alu family of SINEs **(Fig 1.1)** .The insertion of the FLAM copy that given rise to BC200 gene must have occurred after the divergence of the anthropoid lineage from the prosimian lineage (Kuryshev et al., 2001). Like its analogue BC1 RNA, BC200 RNA is evolutionarily conserved and thus under selective pressure, characteristic for functionally active genes. The gene is conserved in all anthropoids including Old World monkeys, New World monkeys and Hominoids (Skryabin et al., 1998). However this gene is absent in representatives of the prosimian lineages [galagos, lemurs and tarsiers]. Interestingly, the orthologus loci in G. moholi and E. coronatus harbor a dimeric Alu element nearly exactly at the same position of the BC200 gene. G22 RNA is almost twice the size of BC200 RNA has a similar sub-cellular localization like in BC200 RNA. Further experimental evidence revealed that G22 RNA also forms RNP complex by interacting with proteins, SRP9/14 and PABP (Khanam et al., 2007b) like BC200 RNA. BC200 RNA and G22 RNA represents the rare offsprings of repetitive DNA families that is transcriptionally and retropositionally active (Khanam et al., 2007b, Kuryshev et al., 2001, Ludwig et al., 2005, Skryabin et al., 1998, Watson and Sutcliffe, 1987) and may be associated with higher cognitive functions in the nervous system. The BC200 RNA gene itself given rise to several [~200] retropseudogenes, but also pseudogenes generated by segmental duplication, examples are BC200β and BC200γ (Martignetti and Brosius, 1993a). BC200 RNA gene is relatively young, the estimated age is about 35-55 million years.

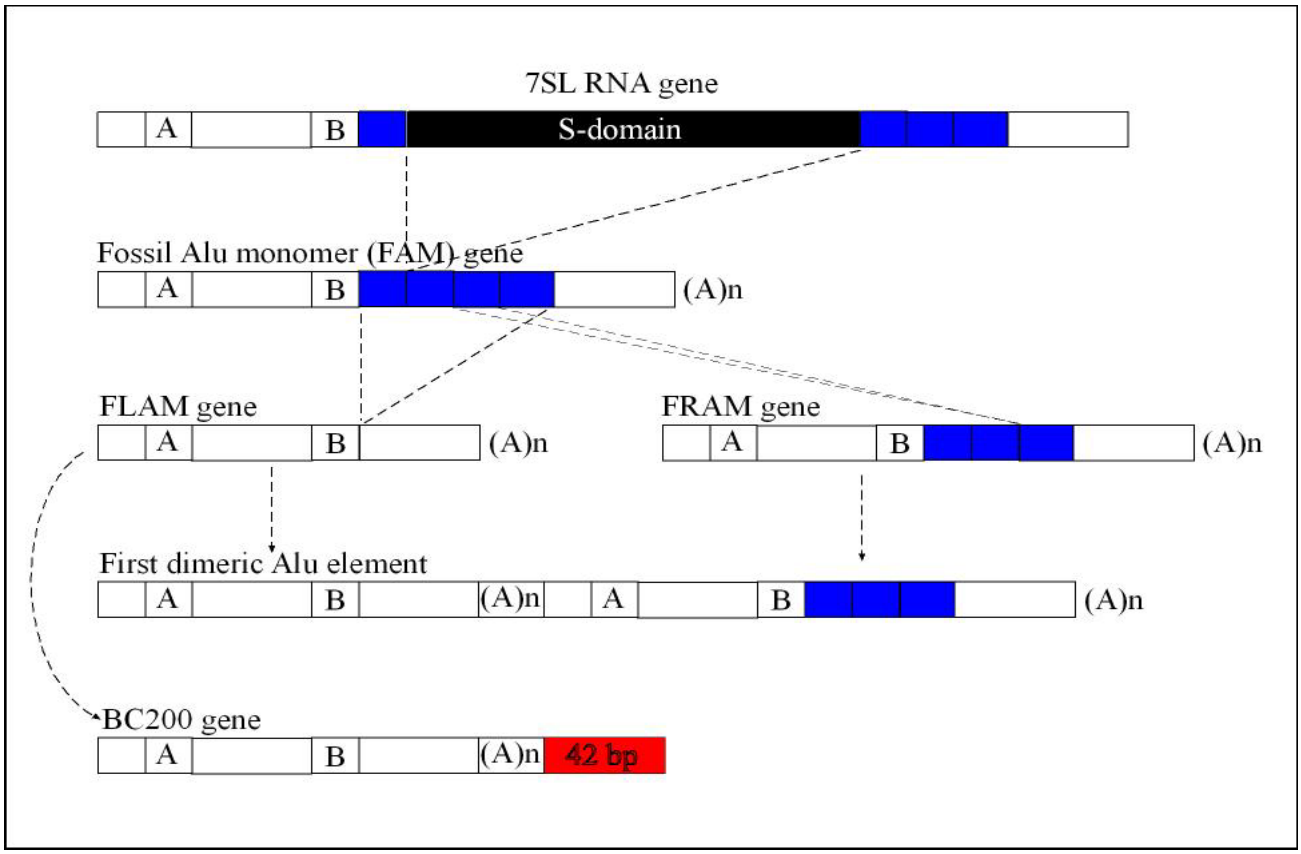


Fig 1.1 Origin of BC200 RNA gene The 7SL RNA is a progenitor for the origin of proto-Alu gene and the neuronal BC200 RNA gene was exapted from a monomeric Alu element. The central S-domain (black) is deleted from the 7SL RNA gave rise to a fossil Alu monomer [FAM]. FAM given rise to FRAM part of Alu SINEs (blue). The red 42 bp long portion of the BC200 gene is a unique sequence derived from the locus of integration and not from the 7SL RNA gene. Both BC200 gene and the first dimeric Alu element arose through independent retroposition.

1.1 Structural features of BC RNAs

The 152 nt long rat BC1 RNA sequence can be subdivided into three major domains **(Fig 1.2)**. The 5' ID domain of 75 nucleotides [nt] which forms a stable stem-loop, this domain harbors a dendritic targeting, a cis-acting element that is responsible for dendritic transport (Muslimov et al., 2006, Muslimov et al., 2002) and a stretch of 22 nt of adenosine residues followed by a unique 3' terminal stem loop structure of about 55 nt. The chemical and enzymatic probing showed that, this domain underwent a structural transition into an extended rod-like stem-loop structure, which is substantially different from the original tRNA cloverleaf structure (Rozhdestvensky et al., 2001). Nucleotide changes in BC1 RNA that are responsible for this structural transition are predominantly located in the sequences corresponding to the acceptor and anti-codon stems of authentic $tRNA^{Ala}$, which participate now in formation of the 5' stem structure (Rozhdestvensky et al., 2001).

Like rat BC1 RNA, 200 nt long human BC200 RNA can be subdivided into three major domains (Skryabin et al., 1998) **(Fig 1.2)**. The 5' domain consists of ~120 nt similar to monomeric Alu repetitive elements, originated from 7SL RNA, the RNA component of the signal recognition particle (Ullu et al., 1982, Ullu and Tschudi, 1984). The central 41 nt of adenosine [A-rich] domain and a 3' terminal segment of 42 nt is unique for BC200 RNA.

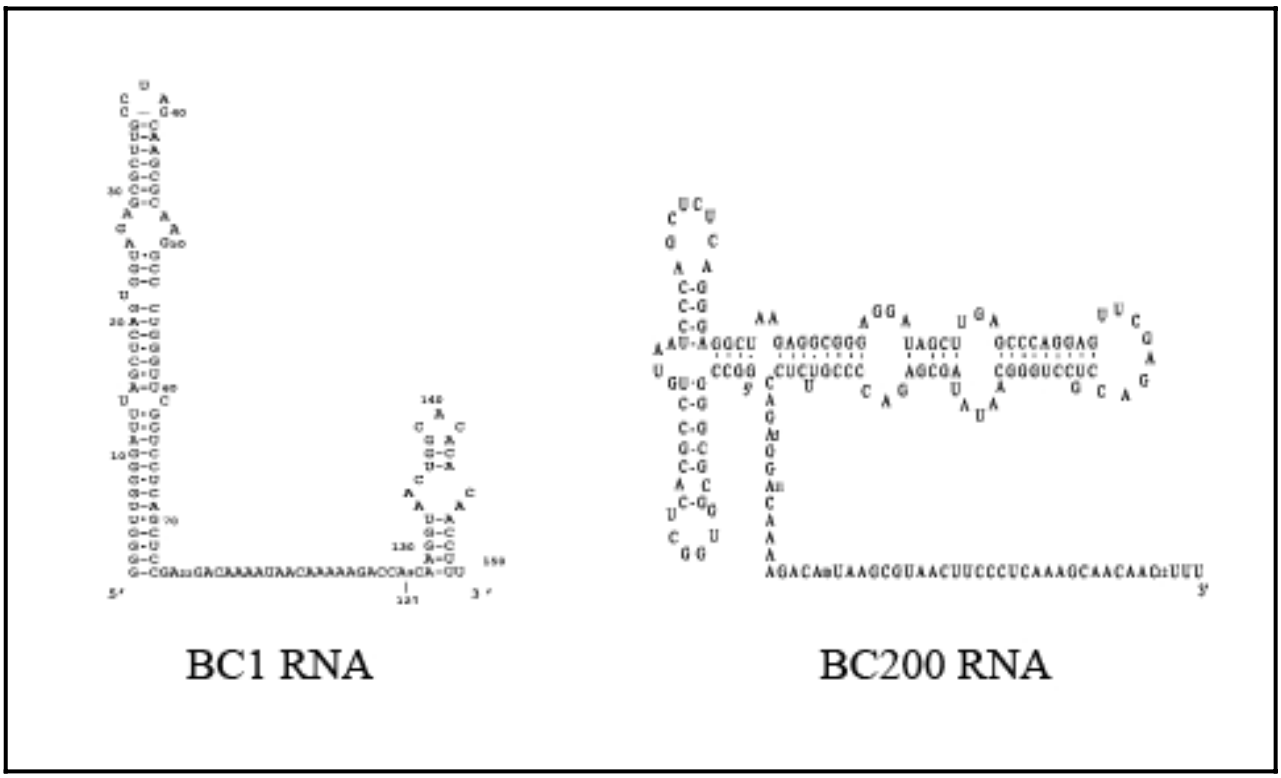


Fig 1.2 Schematic representation of secondary structures of BC1 RNA and BC200 RNA

1.2 Dendritic localization and subcellular distribution of BC RNAs

Despite their different evolutionary origins, both RNAs have many features in common and they are considered to be functional analogues. These RNAs are exclusively expressed in the nervous system and are developmentally regulated. *In situ* hybridization studies on BC RNAs revealed the sub-cellular distribution. Both RNAs are expressed in the retina of ganglion cell layer, innermost part of the inner nuclear layer and inner plexiform layer. Little or no specific labeling was observed over the outer nuclear layer and the outer plexiform layer cells (Tiedge et al., 1991). In the hippocampus CA1, CA2 has moderate signal and CA3 had intense signal for both BC RNAs. In dentate gyrus, no signal was observed for BC1 RNA and a weak signal observed in case of BC200 RNA. In the hippocampus, both BC RNAs are highly expressed in the pyramidal cells in Ammon's horn and in the hilar region.

The olfactory bulb and neo-cortex containing all layers (II – IV except layer I) showed extensive staining**.** In the cerebellum BC1 RNA labeling was low to moderate (Tiedge et al., 1991). White matter areas throughout the brain showed little or no labeling, indicating the absence of BC1 RNA in axons and glial cells.

Two important conclusions were drawn from the studies on BC RNAs [BC1 and BC200] in comparison to GAP-43 mRNA *in situ* hybridization in the adult rat brain and the human brain [respectively]. First the hippocampal neurons that express BC RNAs also contain GAP-43 mRNA indicating that BC RNAs and GAP-43 mRNA are co-expressed by the same type of neuronal cells (Lin et al., 2001, Tiedge et al., 1993). Second BC RNAs are located in postsynaptic domains, while GAP-43 mRNA has been observed in presynaptic domains. It has been shown that BC1 RNA is a prominent component of isolated dendritic spines and synaptodendrosomes (Chicurel et al., 1993, Rao and Steward, 1993). Therefore BC RNAs colocalization in postsynaptic microdomain further supports speculation on postsynaptic RNA metabolism and regulation of mRNA translation.

1.3 Functional Significance and Synaptic Plasticity

BC RNAs localized to dendritic processes of neurons may play an important role in physiological activity and synaptic plasticity responsible for learning and spatial memory. BC RNAs operate as modulators of dendritic RNA functions, like transport of other dendritic components (Tiedge and Brosius, 1996, Tiedge et al., 1993, Tiedge et al., 1991) and dendritic translational regulation (Brosius and Tiedge, 1995, Wang et al., 2005, Wang et al., 2002), Brosius, 1996). Both were speculated earlier – in 2003, the first possibility could be ruled out (Skryabin et al., 2003). Both RNAs are under selective pressure to be expressed in brain and testes of rodents and anthropoid primates respectively. This reflects that these RNAs must have a functional role, but that function must be a subtle rather than essential. Because neither of the RNAs are present in all mammalian orders. This was confirmed, when the gene encoding BC1 RNA was deleted, and no noticeable phenotype was detected and the mice did not reveal any neurological abnormalities. The BC1 deficient mice had shown normal brain morphology and the other dendritic mRNA transport and localization was not changed. Nevertheless, BC1 deficient mice exhibited behavioral changes in terms of reduced exploration and increased anxiety (Skryabin et al., 2003).

From the studies of BC1 deficient mice, we learned dendritic transport of other RNAs, such as mRNAs is not affected. The other possible speculation about these RNAs is that they must involve in translational regulation. This is supported by finding poly(A)-binding protein [PABP], the protein partner of both BC RNAs. This might indicate that BC1 RNA inhibits translation *in vitro* also in *ex vivo* [transfected cells]. These important findings came from our laboratory (Kondrashov et al., 2005, Muddashetty et al., 2002). Wang et al have confirmed, and further demonstrated that BC1 RNA inhibited translational initiation by blocking the formation of 48S initiation complexes (Wang et al., 2005, Wang et al., 2002). They also

demonstrate that BC1 RNA-mediated inhibition occurred by a competition between mRNA and BC1 RNA for binding of eukaryotic initiation factor 4 [eIF4] and PABP (Wang et al., 2002).

1.4 BC RNAs – Expression and deregulation in tumours

In general, BC RNAs are selectively expressed in neurons of brain, also in germ cells such as spermatogonia, oocytes and in cultured immortal cell lines of non-neural origins (Muslimov et al., 2002, Tiedge et al., 1993). But there were reports that BC1 RNA and BC200 RNA apparently were expressed in murine and human tumours respectively (Chen et al., 1997a, Chen et al., 1997b). Distinctively, BC200 RNA was expressed in certain cancerous cells such as breast, cervix, oesophagus, lung, ovary and tongue but not in the corresponding normal tissues (Iacoangeli et al., 2004). The detection of BC200 RNA expression in human tumours demonstrated that BC200 RNA is deregulated in those tumours. Recently Tiedge's laboratory reported, that BC200 RNA is differentially regulated in aging and in Alzheimer's disease (Mus et al., 2007).

1.5 Preceding work on BC RNPs

Role of BC RNAs and associated proteins

The size of the BC1 RNP complex is about 8.8S [naked RNA 6.1S] (Cheng et al., 1996) and the size of BC200 RNP complex 11.4S [naked RNA 7.6S] (Cheng et al., 1997) shows that BC RNAs are bound to proteins. In any RNA-protein complex, not all interactions are identical, some of the proteins bind to the RNA directly, where as other proteins bind to the RNA indirectly via other proteins or transiently. It is very important to establish the composition of these RNP complexes, to identify the proteins bind to these RNAs to determine their fate and functional role.

To date, over a dozen proteins have been reported to be part of BC1 RNP complex. Nevertheless, most of them must be *in vitro* or interpretational artificats. The best example, La (Kremerskothen et al., 1998a) has been published from our laboratory and later we found it is not a part of BC1 or BC200 RNP. Another prominent example is FMRP, which is highly debated and controversial (Napoli et al., 2008, Zalfa et al., 2003), and has been disproven in five laboratories independently (Iacoangeli et al., 2008b, Iacoangeli et al., 2008a).

Poly(A)-binding protein [PABP]

Out of so many candidate proteins of BC1 RNP, poly(A)-binding protein [PABP] is the only *bona fide* protein for both BC RNPs. PABP is a unique translation initiation factor that binds to the mRNA 3' poly(A)-tail and involved in regulation of translation. Because BC RNAs have a stretch of poly (A)-rich sequence 54 nt and 41 nt respectively for BC1 and BC200 RNA, we assumed that PABP can be part of these RNP complexes. Several lines of evidences implicates this protein as binding partner for both BC RNAs, like

immunoprecipitation, gel shift assays, localization etc (Khanam et al., 2006, Muddashetty et al., 2002, Mullin et al., 2004, West et al., 2002).

SRP 9/14 (signal recognition particle)

Many *Alu* containing transcripts such as SRP RNA, BC200 RNA, and B1 transcript share a similar secondary structure as well as conserved sequences, which can potentially interact with the SRP9/14 kDa proteins (Bovia and Strub, 1996, Bui and Strub, 1999, Hasler and Strub, 2006, Kremerskothen et al., 1998b, Tiedge et al., 1993). SRP9/14 proteins form a stable heterodimer in the signal recognition particle and binds with high affinities to 7SL RNA and BC200 RNA. These proteins are bound to the Alu domain of BC200 RNA. This has been confirmed by different lines of evidences like filter binding assays, immunoprecipitation etc. In addition to the SRP9/14 proteins bound to the BC200 RNA, PABP is another protein that binds to the BC200 RNA (Khanam et al., 2006, Muddashetty et al., 2002, Mullin et al., 2004, West et al., 2002).

1.6 Isolation of BC RNPs

The general biochemical methods including gel filtration, sucrose gradients, purifications like DEAE, Mono-Q, superdex-G200, heparin resins [matrix] could not purify BC1 or BC200 RNP complexes to near homogeneity (Muddashetty, R., thesis). Therefore, we sought to develop RNA tag based affinity purification, where BC RNAs are tagged to the MS2 stem loop motif, and the corresponding MS2 coat protein fused to TAP tag co-expressed for purification. The TAP **(Fig 1.3.)** is a two-step affinity purification method, developed by *Bertrand Seraphin* (Rigaut et al., 1999). TAP tag consists of two IgG binding [ProtA] domains, and a calmodulin-binding peptide [CBP] separated by a tobacco etch virus [TEV] cleavage site. The CBP tag allows for efficient selection and specific release from the affinity resin under mild conditions. In contrast, ProtA can only be released from the matrix-bound IgG under denaturing conditions at low pH. Therefore, TEV protease recognition sequence has been inserted between CBP tag and ProtA tag. The TEV protease enzyme specifically cleaves the ProtA tag from CBP tag under mild conditions, so the complex can be eluted under native conditions.

MS2 tag was chosen for two reasons. First, it could be used for transport studies. The second main purpose is to purify the BC RNP complexes to homogeneity and to identify the proteins bound to these complexes. MS2 tag was preferred in comparison to other tags for two reasons. First, MS2 stem [MS2 RNA motif] structure is quite stable and the corresponding MS2 coat protein protects the MS2 stem from processing compared to other tags, which do not have a protein to protect the tagged RNA from processing. Another reason is MS2 RNA motif, a naturally available tag and this is the most selective and efficient in binding to MS2 coat protein.

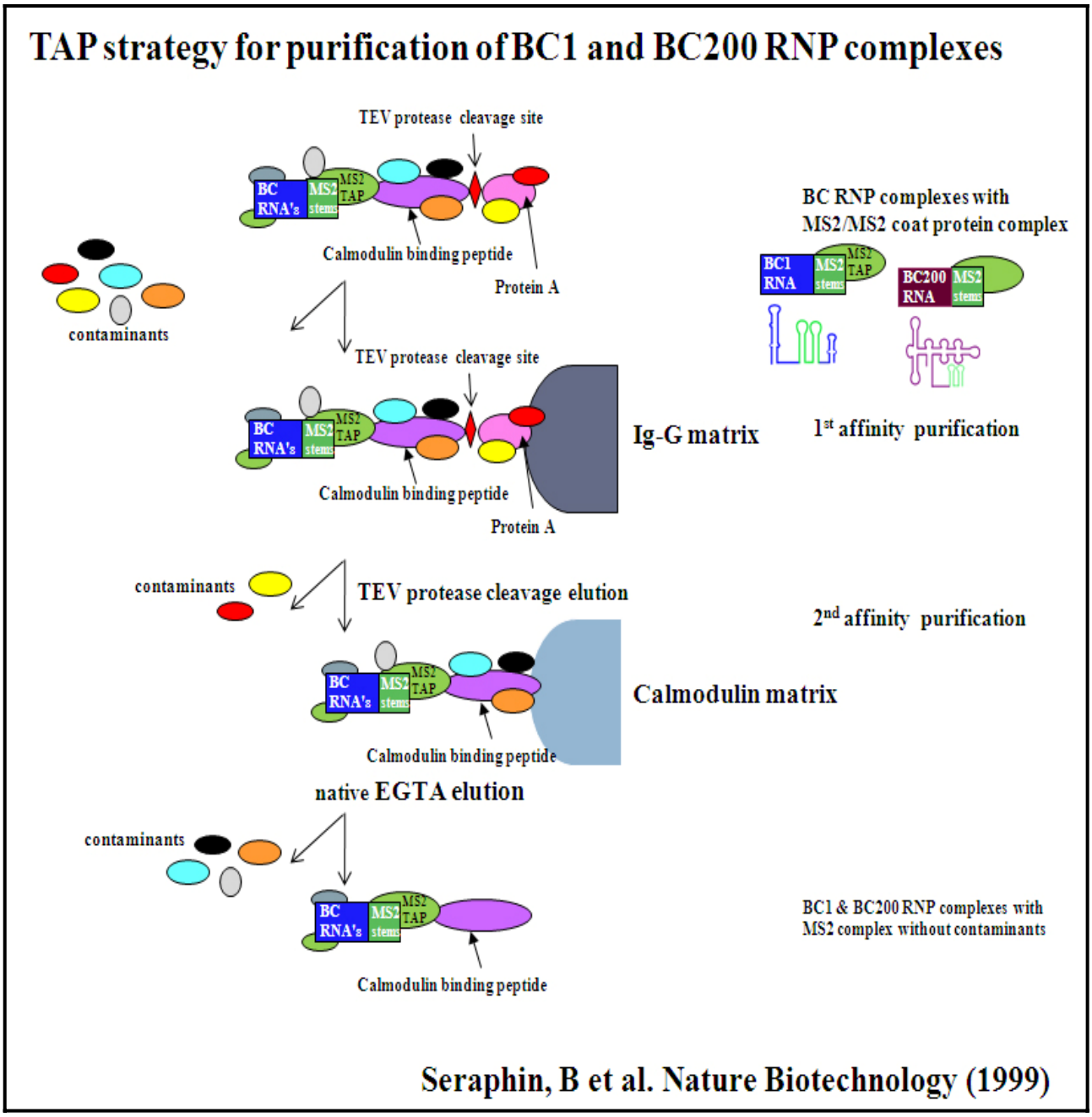


Fig 1.3 Over view of RNA-tag based tandem affinity purification

BC RNAs [BC1 RNA-blue rectangle and BC200 RNA-purple rectangle] tagged with the MS2 RNA structure [green rectangle] bound to the MS2 coat protein [green oval] complex [other domains shown in purple and magenta]. In the first affinity purification the protein A [ProtA - magenta] portion of the TAP fusion polypeptide is bound to the IgG matrix. After the first wash, the bound complex is treated with TEV [tobacco etch virus - red] protease for cleaving the ProtA portion from the remainder of the TAP polypeptide. In the second affinity purification step, the TEV cleaved complex is bound to calmodulin matrix via the CBP tag [calmodulin binding peptide - purple] of the TAP polypeptide. The complex bound to the CBP matrix is washed with buffer to remove additional contaminants. Finally, the BC RNP complexes are eluted from the matrix.

1.7 SINEs (Short INterspersed Elements) as tools in Rodent phylogeny

Up to now, I have given a description of two interesting neuronal dendritic npcRNAs, BC1 RNA and BC200 RNAs, both of which exhibit relationships to SINE elements. Now I will describe the application of SINEs in rodent phylogeny. SINEs are usually very short elements such as ID elements in rodents of about 75 nt long, while some are rather long such as DAS-III SINEs (about 600 nt) (Churakov and Novikov, 2000). SINEs may be amplified to a copy number well beyond $>10^5$ total copies per genome. The presence/absence patterns of SINE elements have proven as invaluable tools for reconstructing virtually homoplasy-free phylogenetic trees (Schmitz et al., 2001, Schmitz et al., 2005, Shedlock and Okada, 2000, Shedlock et al., 2004). The presence of a retroposed element at an orthologous position in related taxon indicates that they are inherited from a common ancestor, while its absence in a taxon indicates they are distant. In general retroposed elements possess several features, they are very unlikely to occur twice independently at orthologous genomic positions. To take advantage of orthologous SINEs as phylogenetic markers in a given group, there must be some SINEs that were actively transposed before the divergence points of speciation. Mouse and rat genome sequence revealed the activity of SINEs during their evolution, B4 probably extinct before the mouse-rat speciation, while B1, B2 and ID are still actively transposed (Gibbs et al., 2004). However, a large number of retroposed SINEs in rodents accomplish in selecting SINEs as reliable markers for inferring evolutionary history of rodents.

The order Rodentia represents almost half of all living mammalian species [2277 species], classified into 30-33 families (Hartenberger, 1998, Reyes et al., 2000) is the most diverse order among placental mammals. The relationships among rodents have been highly debated. Based on morphological and paleontological data, rodents have a single origin and were traditionally divided into either Hystricognathi and Sciurognathi (Tullberg, 1899) or Myomorpha, Sciuromorpha and Hystricomorpha (Brandt, 1855) suborders. Early molecular studies on rodent evolution suggested a paraphyletic grouping of rodents (D'Erchia et al., 1996, Graur et al., 1991, Reyes et al., 1998). Rodent paraphyly has been disproven (Cao et al., 1994, Farwick et al., 2006, Huchon et al., 2007, Martignetti and Brosius, 1993b). Recent molecular analyses suggest that Rodentia are monophyletic and divided into seven, well supported groups: (1) Myodonta [rats, mice, jerboas]; (2) Anomaluromorpha [scaly-tailed flying squirrels, springhares]; (3) Castoridae [beavers]; (4) Geomyoidea [pocket gophers, pocket mice]; (5) Ctenohystrica [gundi, porcupines, guinea-pigs]; (6) Sciuroidea [mountain beavers, squirrels, woodchucks]; and (7) Gliridae [dormice]. But these seven groups are divided into three major clades, a “mouse-related clade” (1 to 4), “ctenohystrica” (5), and a “squirrel-related clade” [6, 7] (Adkins et al., 2001, DeBry, 2003, Huchon et al., 2002, Murphy et al., 2001). Molecular analysis using the nuclear-encoded gene IRBP (DeBry and Sagel, 2001) supports the classical hypothesis that groups Castoridae with Sciuroidea and not the grouping of Castoridae with Geomyoidea, Anomaluroidea and Myodonta, as suggested by other molecular data. Recent studies using fast-

evolving nucleotides in mitochondrial, exon and intron DNA fragments supports the grouping of Anomaluromorpha, Myomorpha as sister clade to Castorimorpha and other grouping comprises Sciuromorpha and Hystricomorpha together (Montgelard et al., 2008). Even molecular analyses were often contradictory, because of different rates of molecular clocks in different taxa/lineages or different selection pressure on the same gene or protein in different taxa/lineages leads to ambiguous conclusions.

To resolve the ambiguous relationships within the rodent phylogeny, we have selected retroposed elements (Schmitz et al., 2001, Shedlock and Okada, 2000). Even though LINE & LTR presence/absence pattern provided a homoplasy-free phylogeny reconstruction of the placental mammalian tree (Kriegs et al., 2006), still two possible sources of homoplasy have been reported in SINE presence/absence evaluation (I) exact excision (II) independent orthologous insertion (van de Lagemaat et al., 2005). The probability of exact excision of SINE element is very low and it would be insignificant as compared to incomplete lineage sorting, an obvious problem for any marker systems. Another possible homoplasy situation is independent insertion at the same genomic loci observed in gene conversion, a rare event. Salem et al., (Salem et al., 2003) found only three cases of gene conversion and only one of them involved full gene conversion, among 133 hominoid Alu-insertions. Analysis of SINE retroposed elements promises a nearly homoplasy-free reconstruction of the rodent phylogenetic tree.

1.8 SINEs in Rodents

Rodents have experienced intense mobile element activity of 7SL RNA, tRNA and 5S rRNA derived SINEs. SINE elements are usually about 75-600 nt long and are defined as short DNA sequences amplified in the genome via an RNA intermediate, transcribed mostly by RNA polymerase III (Singer, 1982). SINEs are widely distributed in the mammalian genome amount to 33% - 45% (Lander et al., 2001, Waterston et al., 2002). They are thought to be dependent on LINE elements and were classified as nonautonomous. Most of the SINEs except human Alu, rodent B1-related and Ped-1 elements are derived from tRNA genes (Daniels and Deininger, 1985, Ohshima and Okada, 2005, Gogolevsky et al., 2008). Rarely SINE RNAs are exapted to serve cellular functions such as Alu derived BC200 RNA and tRNA derived BC1 RNA (Martignetti and Brosius, 1993b). In mammals more than 30 different SINE families have been identified so far, and each of them categorized within distinct taxonomic superorders, orders, families (Kramerov and Vassetzky, 2005). Approximately 33% of the mouse genome consists of rodent-specific repeats (Waterston et al., 2002). The sequencing and comparative analysis of mouse genome revealed 564,000 B1 elements; 348,000 B2 elements; 391,000 B4 elements and 79,000 ID-related elements (Waterston et al., 2002). Moreover, different SINE families were distributed in various rodent lineages at different periods. For example, B1-related SINEs are common to all 30 rodent families (Vassetzky et al., 2003). The first described tRNA derived SINE is B2, about 190 nt in length (Krayev et al., 1982) and, present in four rodent families: Muridae [mice and rats],

Cricitidae [hamsters, voles, and gerbils], Spalacidae [mole rats], and Rhizomyidae [bamboo rats] (Serdobova and Kramerov, 1998). ID SINEs are presumably present in all rodents. ID elements originated from tRNA-like regions. A monomeric 200 nt SINE referred to as DIP element was found in two rodent families, jerboas [Dipodidae] and birch mice [Zapodidae] (Serdobova and Kramerov, 1998). $4.5S_H$ RNA is present in six rodent families: Muridae, Cricetidae, Spalacidae, Rhizomyidae, Dipodidae and Zapodidae (Gogolevskaya et al., 2005). In addition, dimeric SINEs composed of ID and B1 moiety can be very abundant, for instance, B4 in rat and mouse (Lee et al., 1998, Waterston et al., 2002) or B1-dID in squirrels [Sciuridae] and dormice [Gliridae] (Borodulina and Kramerov, 2001, Kramerov, 1999, Kramerov and Vassetzky, 2001, Rogozin et al., 2000). The aforementioned composite SINEs have a different order of the monomers, ID-B1 in B4 and vice versa in B1-dID. Based on SINE information mentioned above, we have used the presence/absence patterns of SINE elements as phylogenetic markers in Rodentia.

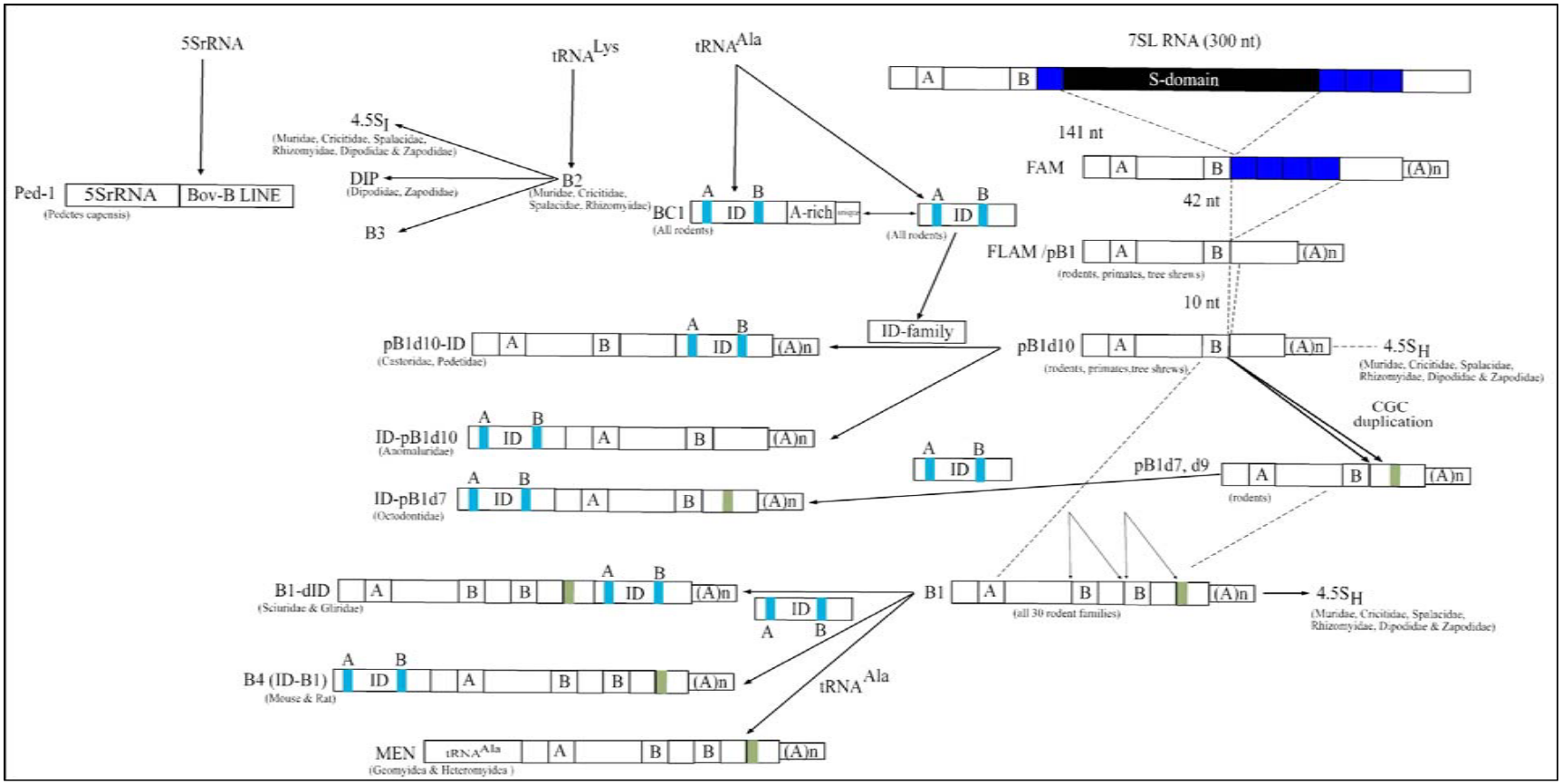


Fig 1.4 Rodent SINEs structures and evolution. At the top right 7SL derived SINE sequence region (A-box and B-box labeled inside the element). Evolutionary events follow one another from top to bottom and represent the most likely events that took place. The central S-domain (black) deleted from the 7SL RNA given rise to fossil Alu monomer [FAM]. FAM given rise to FRAM part of Alu SINEs (blue). FAM retroposon evolved from a 7SL RNA sequence and given rise to FLAM (PB1), a large central deletion believed to given rise FLAM (free left Alu monomer). By diagnostic 10 bp deletion in the FLAM given rise to pB1d10 and the subsequent CGC duplication (green) given rise to pB1d7 and d9. $4.5S_H$ RNA gene assumed to be given rise to B1 and probably from pB1d10. Probably pB1d7 or pB1d10 (dotted line) given rise to a highly efficient B1 SINE which is present in all rodent families. Besides from monomeric SINEs, dimeric SINEs found in rodents. tRNA-derived ID element (A-box and B-box [turquoise]) combined with B1 SINEs given rise to B1-ID or ID-B1(B4) of different order. In addition to the pB1D10 and ID given rise to pB1d10-ID or ID-pB1D10. pB1d7 and ID gave pB1d7-ID. The MEN element has an unusual dimeric structure composed of left $tRNA^{Ala}$ (B2) sequence and the right monomer is B1-related sequences. B2, DIP, MEN and $4.5S_I$ believed to arise from tRNA-derived SINE. The distribution of different SINEs present in different rodent families shown in parentheses.

1.9 Previous work on Rodents

In our lab Farwick et al., has developed an automatic computational strategy to efficiently identify retroposed SINE presence/absence loci with highly conserved flanking regions to resolve the phylogenetic relationships among rodents. From that study, they found 16 informative loci with 31 diagnostic SINE markers (Farwick et al., 2006, Huchon et al., 2007). Some of the deepest branches were not fully assigned, further information is needed to resolve the rodent phylogenic clades, which include “Mouse-related clade”, “Ctenohystrica”, and “Squirrel-related clade”. Furthermore, we expanded our species sampling compared to the previous studies. We further developed new SINE searching strategies, to find more informative loci to cover all three major clades. Available trace database on NCBI GenBank from the representative species in each clade [mouse, rat, kangaroo rat, guinea pig, spermophilus and rabbit as out-group] encouraged us to develop alternative strategies to solve the rodent phylogeny.

Objectives of the study

1 The main objective of this study is to generate a transgenic mouse system to purify the BC1 and BC200 RNP complexes in brain and to identify the proteins bound to the complexes using RNA affinity tags. To achieve this goal, two different transgenic mouse lines should be established, which harbors the BC1MS2 RNA expressing gene in combination with the MS2TAP protein-expressing gene for affinity purification of the BC1 RNP complex and the BC200MS2 RNA expressing gene in combination with the MS2TAP protein-expressing gene.

The established transgenic mouse line should be homozygous with respect to both transgenes [BC1MS2$^{Tg/Tg}$+MS2TAP$^{Tg/Tg}$_BC1 RNA$^{-/-}$] in the endogenous BC1 RNA-depleted homozygous transgenic mouse. In addition BC200MS2$^{Tg/Tg}$+MS2TAP$^{Tg/Tg}$ transgenic mouse line should be homozygous to with respect to both transgenes.

2 A second important objective of the study is to examine the dendritic transport of the MS2 tagged BC RNAs *in vivo*.

3 The third main objective of the study is to solve rodent phylogenetic relationships applying retroposed SINE elements including BC1 RNA related ID repetitive elements.

MATERIALS AND METHODS

2.0 Generation of transgenic mice for RNP purification and to examine dendritic transport

2.1 Construction and generation of BC1MS2 and BC200MS2 (MS2 protein binding sites at 3' end) transgenic mice

For construction of the vector for expressing rat BC1MS2 and human BC200MS2 (MS2 located at the very 3' end of the RNA but proximal to the RNA polymerase III terminator), fragments harboring the RNA genes were cloned into pCR TOPOII cloning vector [Invitrogen]. The two MS2 stems were introduced at the very 3' end of the both RNA genes by PCR mutagenesis as outlined below. PCR based mutagenesis was performed using the following oligonucleotides for the BC1MS2 RNA gene.

[Lowercase letters indicate MS2 stems and italicized lower case letters indicate restriction endonuclease recognition sites, uppercase letters correspond to the oligonucleotides used to amplify the BC1MS2 and BC200MS2 RNA gene fragments.] A spacer (underlined) with the sequence GCAATACCACTATAGAA was added between the two stems to avoid possible sterical hinderance of protein binding to BC1MS2 and BC200MS2.

BC1 forv (5'-AGTCATC*gagctc*TCATTTTCAAAGAGAC-3')

BC1 unirev (5'-GTCTTTTTTAAGGTTGTGTGTGCCAGTT-3')

BC1uniMS2(5'-AACCTTAAAAAAgacatgaggatcacccatgtctGCAATACCACTATAGAAgacatgaggatcaccatgtctTTT-3')

BC1 rev (5'-GTCGGACAGGCCCCTTT*cctagg*ATGTTCT-3').

The rat BC1 RNA gene was used as a template for generating BC1MS2 RNA gene. In the first step, PCR amplification was performed using BC1 forv - BC1 unirev oligonucleotide pair, amplified the upstream of BC1 RNA locus until the very 3' end of the BC1 RNA gene (i.e. before UUU [3U's]). In another amplification step, using BC1 uniMS2-BC1 rev oligonucleotides, the downstream part of the BC1 RNA locus was amplified including the two MS2 stems at the 3' end of the BC1 RNA gene in place of UUU, UUU and the following were inserted after the two MS2 stems for termination of the BC1MS2 transcript followed by wild type locus sequence (TCATTTTCAAA). In the second PCR amplification step, both the above mentioned PCR fragments were mixed, and the entire 1.5 kb fragment including the two MS2 stems was amplified using the BC1 forv and BC1 rev oligonucleotide pair **(Fig 2.0 A)**. Thus, two MS2 RNA motifs were introduced at the very 3' end of BC1 RNA (proximal to the unique region).

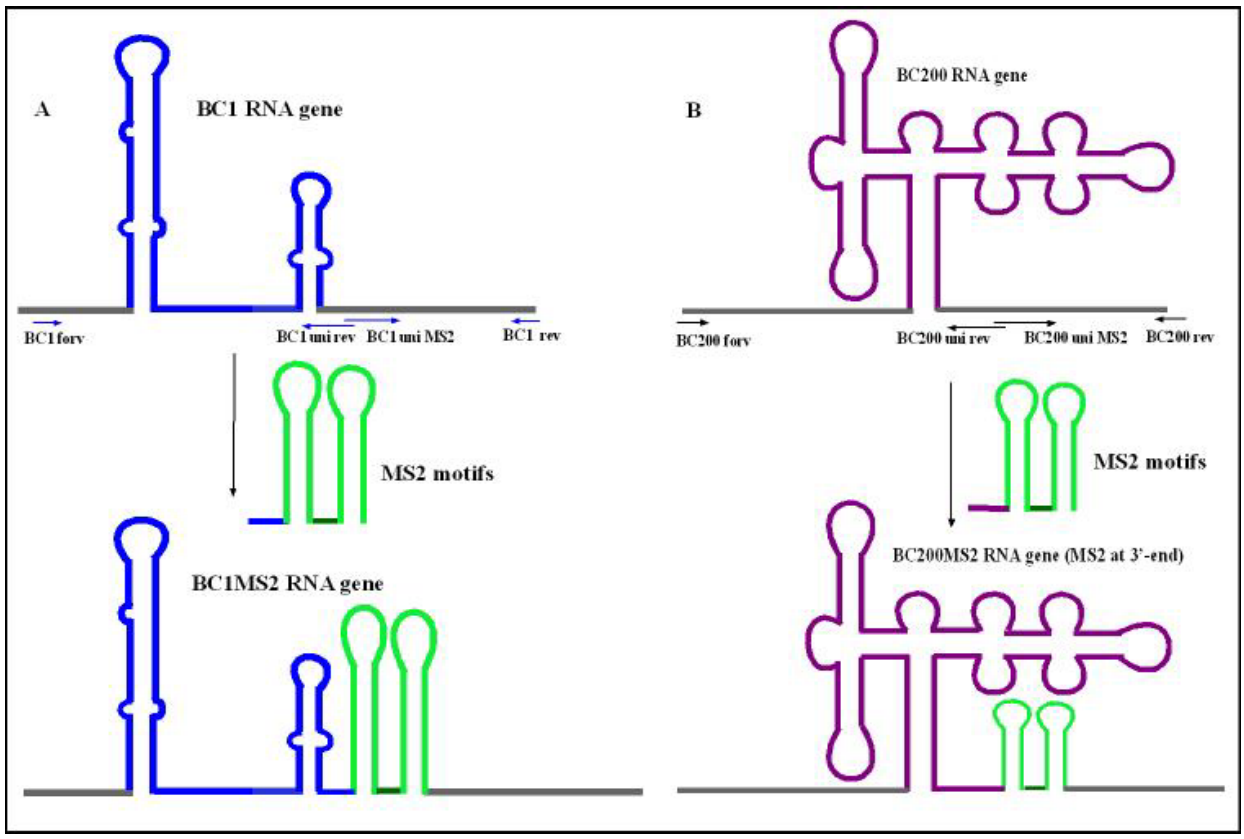


Figure 2.0 Scheme of the cloning strategy for fusing two MS2 stems [green color, with a short spacer between the two stems in dark green] at the very 3' end of the BC1 RNA gene [blue color] and BC200 RNA gene [purple color]. MS2 stems were placed closer to the ID domain (BC1MS2A) and closer to the 3' stem (BC1MS2B) followed by RNA polymerase III transcription termination signals as occurring in the wild type loci. The flanking regions that include control regions for transcription are shown in gray. Arrows denote PCR amplification schemes and primers (for details, see text).

To amplify the BC200MS2 fragment following oligonucleotides were used. BC200 forv (5'-ACGATG*gaattc*TAAGACAGTGTCATAGA-3') and
BC200 uniMS2 (5'-CCCTTAAAAAAgacatgaggatcacccatgtctGCAATACCACTATAGAAgacatgaggatc acccatgtctTTTC-3') BC200 unirev (5'-ATCCTCATGTCTTTTTTAAGGGG-3') BC200rev (5'-GACTTTCTTTGA*cttaag*CTAT-3'). For BC200MS2 RNA gene, the human BC200 RNA gene was used as a template for amplification. In the first step, PCR was performed using BC200 forv- BC200 unirev oligonucleotide pair, amplified the upstream of BC200 RNA locus until the very 3' end of the BC200 RNA gene (i.e. before UUU [3U's]). In another PCR amplification step using BC200 uniMS2 - BC200 unirev oligonucleotide pairs, downstream part of the BC200 gene locus was amplified including the two MS2 stems at the 3' end of the BC200 RNA gene in place of UUU and the UUU and the following were inserted after the two MS2 stems for termination of the BC200MS2 transcript followed by the wild type locus sequences (TCATATT). In the second PCR amplification, both the PCR products were mixed and

reamplified the entire 3.9 kb fragment including the two MS2 stems using the BC200 forv and BC200 rev oligonucleotides **(Fig 2.0 B)**. Thus, two MS2 RNA motifs were introduced at the very 3' end of BC200 RNA (proximal to the unique region). Both the mutagenized BC1MS2 and BC200MS2 RNA gene fragments were cloned into pCR TOPOII vector. BC1MS2 transgenic mice were generated by inserting the 1.5 kb *Sac*I - *Bam*HI fragment released from pCR TOPOII BC1MS2. To generating BC200MS2 transgenic mice the 3.9 kb fragment of *Eco*RI - *Eco*RI was released from pCR TOPOII BC200MS2.

2.2 Construction and generation of MS2TAP protein transgenic mice

For generating MS2TAP protein expressing mice, the MS2TAP fragment was cloned into the pNECKu vector (Blichenberg et al., 1999), a generous gift from Dr. Stephan Kindler, University of Hamburg (Blichenberg et al., 1999). The vector features the chicken β-actin promoter and the transcript has a ~1 kb intron in the 5' UTR (also derived from the chicken β-actin gene). The protein coding sequence of MS2TAP replaced the EGFP ORF of pNECKu between the *Eco*RV/*Xba*I restriction sites. HA and FLAG peptide sequences were fused to the N-terminus of MS2TAP coding sequence for detecting the protein expressed in cell lines or in mice by Western blots. The MS2TAP fragment was generated using a PCR based approach with appropriate restriction sites to clone into the pNECKu vector. Amplification of the MS2TAP fragment was achieved as follows. The fragment encoding the MS2 coat protein was obtained from plasmid pG14-MS2-GFP (Kaganman, 2008) (kindly provided by Rodert H. Singer, Albert Einstein College of Medicine, Bronx, New York) by PCR using the following oligonucleotide pairs : (5'-GGTCGCAT*gatatc*GAC<u>ACCATGG</u>TGGACTACAAAGAC-3') and (5'-CAGCGAATTAATTCCAGCACAC-3') was in frame MS2TAP protein. The ATG start codon is embedded within a Kozak consensus sequence (Kozak, 1987) (underlined) and, harbored HA and FLAG tag peptides at the N-terminus of the MS2 coat protein. The fragment encoding the TAP protein was obtained from pTAP (kindly provided by Bertrand Seraphin, EMBL, Heidelberg, Germany) by PCR amplification using the following oligonucleotide pairs:
(5'-TTAACCTTT*gcggccgc*TCTAGTAACAGCAGCAAGTGTGCTGGAATTAATTCGCTGTCTAGGA GGATCCATGGAAAAGAGAAGA-3'), (5'-GGTCGCAT*tctaga*CAAGCTTCAGGTTGACTTCCCC-3'). The italicized lower case letters were *Eco*RV, *Not*I and *Xba*I restriction enzyme sites respectively. The MS2 coat protein fragment was obtained by PCR amplification was restriction digested with *Eco*RV and *Not*I and TAP protein fragment was PCR amplified and restriction digested with *Not*I and *Xba*I. After restriction digestion, both the fragments were ligated to obtain the MS2TAP protein and cloned in pNECKu vector, which was previously digested with *Eco*RV and *Xba*I restriction endonucleases. The obtained plasmid was designated as pNEMS2TAP. The DNA insert was released with *Pvu*I - *Nru*I to generate transgenic mice.

2.3 Cloning of the combined (double gene) construct and generation of transgenic mice - BC1MS2+MS2TAP

Constructs that harbor both BC1MS2 RNA and MS2TAP protein-expressing cassettes were cloned into a single plasmid. The cloning scheme for the construct was as follows.

The BC1MS2 fragment (*Sac*I - *Apa*I) contained 423 bp upstream of the BC1MS2 RNA gene and about 200 bp downstream of the RNA polymerase III terminator of the RNA gene was PCR amplified from pCR TOPOII BC1MS2 (described above) and cloned into pBluescript II KS+. The MS2TAP fragment was released using *Pvu*I - *Nru*I from pNEMS2TAP plasmid (described above) and subcloned into pBS, which harbored BC1MS2 (with the MS2 domains located at the very 3' end of BC1 RNA). Insertion of the MS2TAP fragment was done by blunt end cloning resulting in two different orientations with respect to the plasmid backbone. The MS2TAP fragment inserted into pBS-BC1MS2 in the same transcriptional orientation as BC1MS2 [i.e. head to tail], resulting in plasmid as 11 dirTAP. The plasmid with MS2TAP inserted in the opposite orientation with respect to BC1MS2 [head to head] resulting in plasmid 3 revTAP **(Fig 3.2)**. The inserts carrying the RNA and protein coding genes were released with *Sfi*I and *Asc*I to generate the respective transgenic mice.

2.4 Construction and generation of BC1MS2A and BC1MS2B RNA (MS2 stems internal to the BC1 RNA) transgenic mice

As the MS2 stems at the very 3' end of BC1 RNA were post transcriptionally removed in most molecules in the *in vivo* mouse system, constructs, which harbor the MS2 stems were placed near the central Adenosine-rich domain of BC1 RNA. One was placed proximal to the ID stem of BC1 RNA (Rozhdestvensky et al., 2001) designated as - BC1MS2A, and the other was placed proximal to the 3' stem of BC1 RNA designated as BC1MS2B. Both constructs were cloned into pDrive cloning vector [Qiagen, Hilden]. The following oligonucleotides were used for PCR amplifications to generate the BC1MS2A and BC1MS2B DNA fragments. The lower case letters indicates the MS2 stem. A spacer with the sequence GCAATACCACTATAGAA was added between the two stems for proper folding.

BC1d1 (5'-CTCTCATTTTCAAAGAGCACCAT-3')

BC1d2 (5'-AAAAAAGACAAAATAACAAAAAGACCA-3')

BC1r1 (5'-GATCCAAAGGGGCCTGTCCGA-3')

BC1r2 (5'-TTTTTTCGGAGCTGAGGACC-3')

BC1d2MS2A (5'-GGTCCTCAGCTCCGAAAAAAAgacatgaggatcacccatgtctGCAATACCACTATAGAA gacatgaggatcacccatgtct$(A)_{22}$GACAAAATAACAAAAAGACCA-3')

BC1d2MS2B (5'-GTCCTCAGCTCCGA$(A)_{22}$AgacatgaggatcacccatgtctGCAATACCACTATAGAAgacat gaggatcacccatgtctAAAAAAGACAAAATAACAAA-3').

For generating the BC1MS2A RNA constructs, the rat BC1 RNA gene fragment (1.4 kb *Sac*I-*Bam*HI) was used as a template for PCR reaction. PCR amplifications were performed as two steps reaction. In the first the BC1d1/BC1r2 oligonucleotides, were used to amplify the upstream part of BC1 locus. The BC1d2/BC1r1 oligonucleotides, were used to amplify the downstream part of BC1 locus. In the second step of the PCR both PCR product from the first step were combined with the oligonucleotide BC1d2 MS2A containing two MS2 stems, and re-amplified the entire BC1 locus including two MS2 stems using BC1d1 and BC1r1 oligonucleotides **(Fig 2.1 A)**.

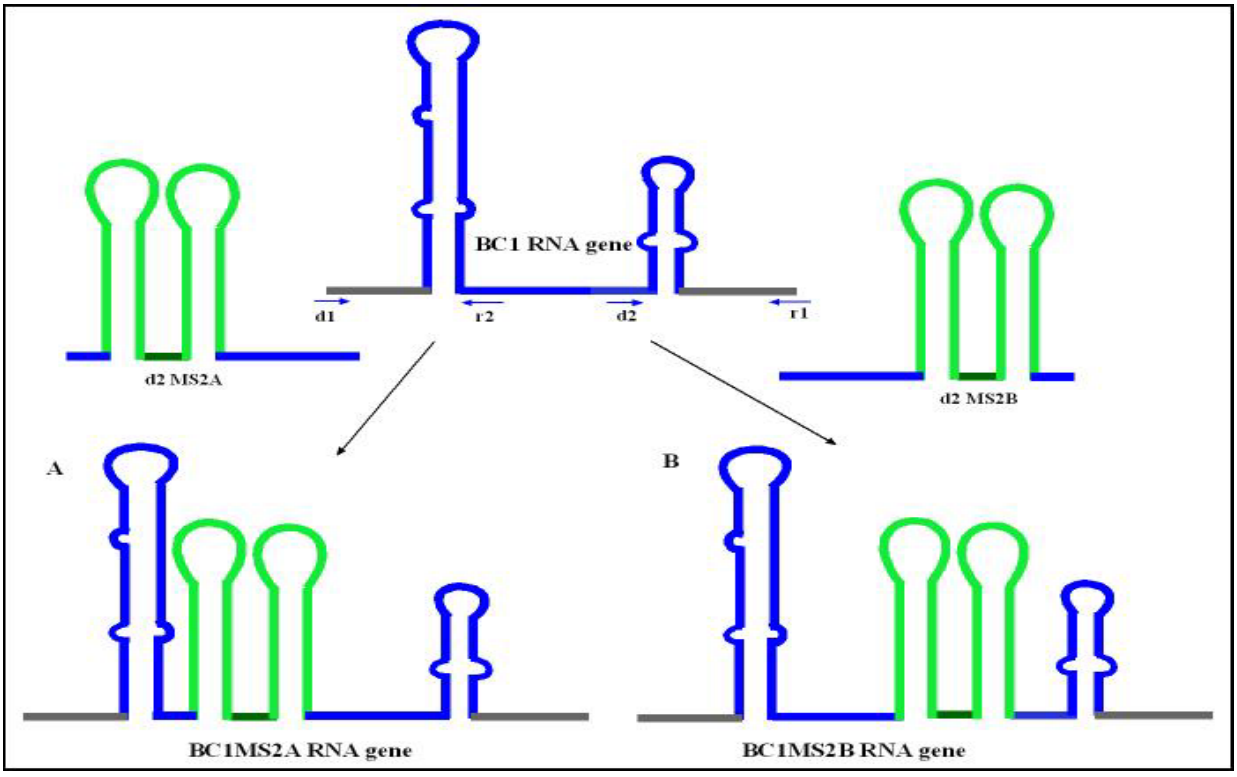


Figure 2.1 Scheme of the cloning strategy for insertion of two MS2 stems [green] with a spacer between two stems [dark green] into the BC1 RNA gene [blue]. The flanking regions that include control regions for transcription are shown in gray. Arrows denote PCR amplification schemes and primers. MS2 stems were placed closer to the ID domain (BC1MS2A) and closer to the 3' stem (BC1MS2B) followed by RNA polymerase III transcription termination signals as occurring in the wild type loci. Arrows denote PCR amplification schemes and primers (for details, see text).

The BC1MS2B RNA construct generated analogously. In the first round, the BC1d1/BC1r2 (upstream part) and BC1d2/BC1r1 oligonucleotides (downstream part of BC1 locus), were used to amplify the corresponding regions. In the second step, both PCR product were combined with the oligonucleotide BC1d2 MS2B containing two MS2 stems, and re-amplified the entire BC1 locus including two MS2 stems using BC1d1 and BC1r1 oligonucleotides **(Fig 2.1 B)**.

BC1MS2A and BC1MS2B transgenic mice were generated by inserting a 1.5 kb DNA fragment released with *Eco*RI from pDrive BC1MS2A/B plasmids.

All cloned sequences including mutagenesis in this study were confirmed by sequence analysis, before generating transgenic mice. In brief, for generating transgenic mice, the DNA fragments were diluted up to 2 ng/µl and injected into the pronuclei of FVB mice according to the protocol (Hogan et al., 1986) and subsequently transferred into the uteri of the 2.5 - day pseudopregnant CD-1 foster mice.

2.5 Generation of BC1MS2B+MS2TAP transgenic mice in BC1 RNA$^{-/-}$ background

In order not to competing with endogenous BC1 RNA for binding proteins, BC1MS2B RNA expressing transgenic mice were crossed with BC1 RNA deficient mice [BC1 RNA$^{-/-}$]. The resulting transgenic mice were designated BC1MS2B in BC1 RNA$^{-/-}$. Similarly MS2TAP protein expressing mice were crossed into BC1 RNA deficient mice [BC1 RNA$^{-/-}$], designated as MS2TAP in BC1 RNA$^{-/-}$. Finally both BC1MS2B in BC1 RNA$^{-/-}$ and MS2TAP in BC1 RNA$^{-/-}$ were crossed to yield mice, that express both BC1MS2B RNA and MS2TAP protein in a background depleted of endogenous BC1 RNA. The resulting mouse strain was labeled BC1MS2B+MS2TAP_BC1 RNA$^{-/-}$. The heterozygous [BC1MS2B$^{Tg/-}$ +MS2TAP$^{Tg/-}$ in BC1 RNA$^{+/-}$] mice were interbred for about 4-5 generations monitoring their status with the goal of obtaining homozygosity BC1MS2B$^{Tg/Tg}$+MS2TAP$^{Tg/Tg}$ in the BC1 RNA$^{-/-}$ background. The resulted mice were screened for BC1MS2B$^{Tg/Tg}$+MS2TAP$^{Tg/Tg}$ in BC1 RNA$^{-/-}$ were homozygous to both transgenes and expressed no endogenous wild type BC1 RNA. This was confirmed by Southern blotting, Real Time PCR analysis and by test crossing with non-transgenic mice [wild type], where all offspring was positive (heterozygous) for both transgenes. Thus, we obtained a stable homozygous mouse line with the desired traits.

2.6 Generation of BC200MS2+MS2TAP transgenic mice

BC200MS2 RNA expressing mice were crossed with MS2TAP protein expressing mice, to obtain transgenic mice, which harbor both BC200MS2 and MS2TAP expressing genes. The resulting transgenic mouse line was designated as BC200MS2+MS2TAP and were interbred for about 4 - 5 generations to obtain homozygous to both transgenes.

2.7 Generation of BC200 RNA transgenic mice

Transgenic mice expressing wild type BC200 RNA were previously generated in our laboratory and used as a control. They contained the 3.2 kb *Eco*RI – *Hind*III DNA fragment from the human BC200 RNA locus [–2271 bp upstream and 766 bp downstream of the BC200 coding region] (Khanam et al., 2007b).

2.8 Generation of BC1 KO (BC1 RNA$^{-/-}$) mice

BC1 RNA gene depleted (knockout) mice were available in our laboratory by targeted gene deletion as described before (Skryabin et al., 2003) and used as negative control and for crossing with transgenic mice that express BC1MS2B RNA and MS2TAP protein, in order to eliminate competition for expression and /or for binding proteins.

2.9 Construction of pMKS plasmid for stable transfection

The goal for constructing this mammalian vector that allow co-expression of novel npcRNA-candidates fused with an MS2 stem sequence and of an MS2TAP protein that specifically binds to the MS2 motif to establish a general system to test subcellular-localization of any npcRNAs by transfecting mammalian cells and using immunocytochemistry with an antibody against the MS2 fusion protein or by *in situ* hybridization. This general method could be used in transgenic mice but also in high throughput, employing transfected-cell array technology. This construct also provides the means for *in vitro* transcription of any npcRNA together with 3' terminal MS2 RNA tag. RNAs are transcribed by CMV promoter, an RNA polymerase II promoter and was terminated by CMV terminator.

The plasmid was constructed as follows. In a first step the previuosly present *Sal*I and *Not*I restriction sites, in pNEMS2TAP were eliminated so they could be used at the insertion sites for npcRNA coding regions. PCR based [Quick Change Site-Directed Mutagenesis kit from Stratagene, La Jolla, CA] site directed mutagenesis was carried out to mutate the *Sal*I and *Not*I sites using the following oligonucleotides:

F*Sal*I mut (5'-CTCAGTTCGTTCTC*gtagat*AATGGCGGAACTG-3')

R *Sal*I mut (5'-CAGTTCCGCCATT*atctac*GAGAACGAACTGAG-3')

F*Not*I mut (5'-CTCCGGCATCTAC*gcagccgc*TCTAGTAAC-3')

R*Not*I mut (5'-GTTACTAGA*gcggctgc*GTAGATGCCGGAG-3').

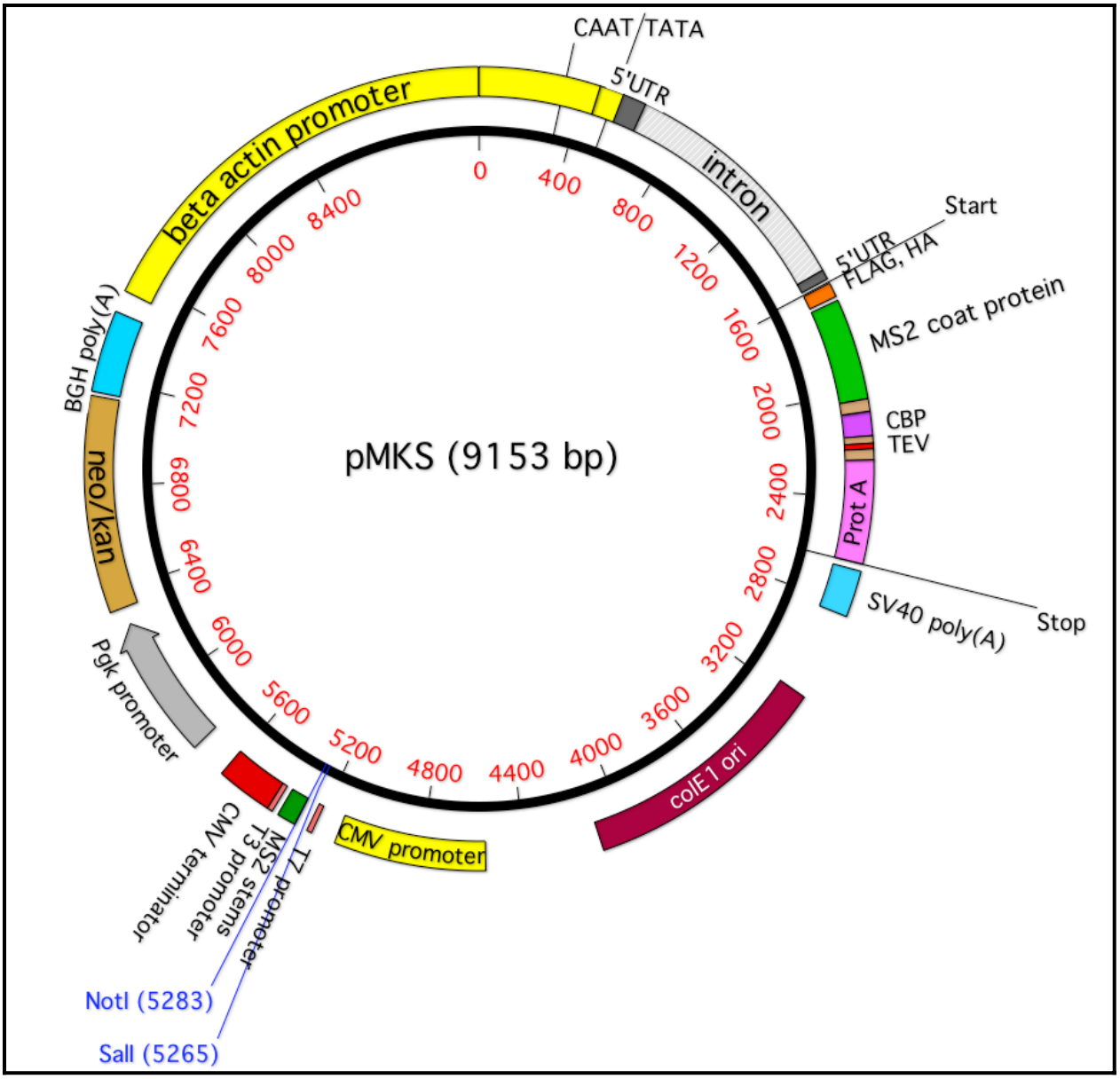


Figure 2.2 Construct pMKS harbors MS2TAP protein expressing the cassette from pNEMS2TAP, features the chicken β-actin promoter [yellow] was used to drive transcription of the MS2TAP fusion protein. The β-actin promoter is followed by the 5' UTR [gray] which, inturn is interrupted by a ~1 kb intron [striped area], also both derived from the chicken β-actin locus. The MS2TAP fusion polypeptide has the following components: orange, FLAG and HA peptide as antigenic determinants, green, MS2 coat protein; purple, calmodulin binding peptide [CBP]; red, tobacco etch virus [TEV] protease site; magenta, protein A domain [Prot A]. The protein coding sequence is followed by polyadenylation signals from SV40 [turquoise]. The vector also harbors two MS2 RNA stem motifs [small green box] driven by the CMV promoter [yellow] preceded by *Sal*I and *Not*I restriction sites that allow cloning novel npcRNAs, followed by a CMV transcription terminator [red]. The vector features the neomycin resistance gene cassette [ocher] flanked by the Pgk promoter [gray] and the bovine growth hormone [BGH] poly(A) addition signal for stable transfection in mammalian cells.

In the second step, the two MS2 stems transcribed by the CMV promoter and terminated by the CMV terminator were PCR amplified. The *Sal*I and *Not*I restriction sites were introduced proximal to the MS2 stems to clone inserts for example npcRNA-derived cDNA libraries. Both CMV promoter and CMV terminator were amplified from pCDNA 3.1. The PCR fragment was cloned into PL452 from Pentao Liu to include a Pgk Neo cassette. The plasmid was digested with *Asp*18 - *Bam*HI, and the fragment containing CMV promoter including the two MS2 stems, CMV terminator and Pgk Neo cassette was sub-cloned into pNEMS2TAP vector.

The following oligonucleotides were used to amplify the CMV promoter and CMV terminator with two MS2 stems.

fCMVpr : (5'-ATATTAAAGGTACCAAGGCGCGCCTGTAGCGGCACGATATACGCGTTGACAT-3')

rCMV pr : (5'-GGCCACCCGGGCCAAAAATGTACATTTagacatgggtgatcctcatgtc TTCTATAGTGGTATTGCagacatgggtgatcctcatgtcTTTTTTGCGGCCGCTCGAGTCTAGAGTCGACGTTTAAACGCTAGCCAGCTTGGGTCTC-3')

fCMV ter: (5'-TGTACATTTTTGGCCCGGGTGGCCCCCTTTAGTGAGGGTTAATTCCTCGACTGTGCCTTCTAGTTGCCAG-3')

rCMV ter: (5'-ATTATATAGAATTCCCCAGCATGCCTGCTATTGTCTTCCCA-3').

The cloned plasmid was confirmed by sequence analysis.

Note: Lower case letters indicates the MS2 stem between the two stems spacer was added.

2.10 DNA transfection

NIH 3T3 cells were grown on 10 cm plates with Dulbecco's modified eagle medium [Gibco BRL]. Plates with ~60-70% cell density were chosen for transient transfection [lipofection] with DNA. Before transfection, cells were changed into Opti-MEM [Gibco BRL] medium. Liposomes were prepared by separately adding 5 and 10 µg of DNA (~1 µg/µl) and 15 µl of DAC-30 cationic lipid reagent [Eurogentec, Seraing, Belgium] to 2.5 ml of Opti-MEM, gently mixed, and incubated at room temperature for 20 min. This liposome mixture was then added to the cells. After incubation for 3 hours, the medium was changed to Dulbecco's modified eagle medium and cells grown for an additional 8-12 hours. Cells were washed twice with 1x PBS (phosphate-buffered saline) and harvested by centrifugation at 2,000 x g for 5 min. Cytoplasmic extracts were generated by lysing the cells in non-denaturing hypotonic buffer [10 mM HEPES pH 7.9, 10 mM KCl, 0.1 mM EDTA, 0.1 mM EGTA, 1 mM DTT, one tablet/50 ml protease inhibitor mix "complete", Boehringer Mannheim], and by subsequently adding NP-40 to a final concentration of 0.05%]. This extract was used for Western blotting. For Northern blotting, total RNA was extracted by using TRIzol reagent [Gibco BRL] according to the manufacturer's instructions.

2.11 PCR analysis of transgenic mice

DNA from 3-week-old mouse-tails was extracted as described by (Laird et al., 1991). Approximately 500 ng of genomic DNA was taken for PCR analysis. The following oligonucleotides were used to screen transgenic mice.

For analyzing BC1MS2 transgenic mice [BC1MS2 - 3' end, BC1MS2A and BC1MS2B] the oligonucleotide pair BC1 ScD (5'-TACCATCTGATACTTGACTGTGT-3') and BC1 ScR (5'-GGGTGATCCTCATGTCTTCTAT-3') were used.

For BC200MS2 transgenic mice analysis the oligonucleotides BC200 ScD (5'-TGCGAACTTCTTACAGTTTAGG-3') and BC200 ScR (5'-AACCTTCCATAATGAAGGC-3') were used.

For MS2TAP transgenic mice analysis the following oligonucleotides fMS2 (5'-TTAACCTTTGCGGCCGCTCTAGTAACAGCAGCAAGT-3') and rTAP (5'-GGTCGCATTCTAGACAAGCTTCAGGTTGACTTCCCC-3') were used.

PCR was performed in a solution containing 50 mM Tris HCl pH 9.5, 20 mM $(NH_4)_2SO_4$, 1 mM dithiothreitol, 1.5 mM $MgCl_2$, 0.005% NP-40, 200 µM deoxynucleoside triphosphates, and 0.5 µM of each primer; annealing was performed at 94 °C for 4 min, and 35 cycles of 93 °C for 45 s, 55 °C for 30 s, 72 °C for 1 min, and with a final 5 min extension at 72 °C. The resulting PCR products were fractionated by 1% agarose gel electrophoresis.

2.12 DNA tail biopsy and Southern blot analysis

Approximately 5 µg of genomic DNA from mouse tails was digested with *Pst*I, *Eco*RI or *Bam*HI fractionated on 0.8% agarose gels, and transferred to GeneScreen nylon membranes [NEN DuPont] by positive pressure. Membranes were hybridized with a ^{32}P-labeled probe specific for the inserted genes, and washed twice with 0.5 x SSPE - 0.5% sodium dodecyl sulfate, [1 x SSPE consists of 0.18 M NaCl, 10 mM NaHPO4, and 1 mM EDTA pH 7.7] at 65 °C for 15 min each, and exposed to MS film [Kodak] at -80 °C overnight.

2.13 Preparation of tissue extracts

Extracts were prepared from adult mouse brain [excluding cerebellum as this part usually is low BC1 RNA expression]. Brain tissue was homogenized in 10 volumes of extraction buffer [20 mM Tris-HCl, pH 7.5, 100 mM KCl, 5 mM $MgCl_2$, 10% sucrose, 1 mM DTT, 0.25% NP40 and one tablet/50 ml protease inhibitor mix “complete”, Boehringer Mannheim]. The extract was centrifuged at 1,750 x g for 11 min at 4 °C. The supernatant was collected (S1) and centrifuged at 15,000 x g for 1 hour at 4 °C. The

resulting supernatant (S2) was stored at -80 ºC or further ultra centrifuged at 60,000 x g for 2 hours at 4 °C. The resulting supernatant (S100) was stored at -80 ºC.

The following mice were used for preparing the brain extracts:

Wild type extracts were prepared from NMRI/HWD outbred mice. BC1 KO (BC1 RNA $^{-/-}$)

BC1MS2 (MS2 at the very 3' end of BC1 RNA), 11 dirTAP, 3 revTAP, BC200MS2TAP (MS2 at the very 3' end of BC200 RNA), BC1MS2A (MS2 stems proximal to the ID region of BC1 RNA), BC1MS2B (MS2 stems proximal to the 3' stem of BC1 RNA) MS2TAP, BC1MS2B_MS2TAP extract were prepared from the corresponding transgenic mice.

All extracts were used for isolating RNA/RNP of BC1 or BC200 with MS2 tag or for identifying the MS2TAP protein.

2.14 RNA blot analysis

Total RNA was extracted from different mouse tissues using TRIzol reagent [Gibco BRL] according to the manufacturer's instructions. RNA was separated on 6% polyacrylamide/7 M urea gels and electro-transferred to positively charged nylon membranes [Ambion]. After UV-cross linking, membranes were hybridized at 58 °C overnight in 1M sodium phosphate buffer pH 6.2, containing 7% SDS and ^{32}P labeled oligonucleotide probes. After hybridization, membranes were washed twice with 0.2M sodium phosphate buffer pH 6.2, containing 1% SDS for 20 min each at 58 °C and exposed to MS film [Kodak] at -80 °C overnight to several days [2-8 days]. In case of LNA oligonucleotide probes, membranes were washed three times with 0.1M sodium phosphate pH 6.2 and 0.5% SDS for 30 min each at 65 °C and exposed to MS film [Kodak] at -80 °C overnight to several days [2-8 days].

(a) 5' end labeling of oligonucleotides

Oligonucleotides either DNA or LNA were radiolabeled, using T4 polynucleotide kinase [PNK], which adds a phosphate group to the 5' end of the oligonucleotide. Kinase reactions were carried out in a 15 µl volume, containing 20 pM of oligonucleotide, 1 x PNK buffer [New England Biolab], 10 U of PNK [New England Biolab] and 20 µCi of γ^{32}P ATP and volume was adjusted with water. The reaction was carried out at 37 °C for 1 hour and was stopped by adding 1µl 0.5M EDTA. Entire reaction mixture was directly added to the prehybridization solution.

Table 1

DNA oligo designation	oligonucleotide sequence
BC1MT (27 mer)	5'-GGTTGTGTGTGCCAGTTAACCTTGTTT-3'
MS2 (35 mer)	5'-CCTCATGTCTTGTAGACATGGGTGATCCTCATGTC-3'
BC207 (36 mer)	5'-CTTGTTGCTTTGAGGGAAGTTACGCTTATTTGGTAC-3'
7SL (36 mer)	5'-CGAGGTCACCATATTGATGCCGAACTTAGTGGGTAC-3'

LNA oligo designation	oligonucleotide sequence
BC1_uni_ LNA (21 mer)	5'-GTGtGCcAGtTAcCTtGTtTT-3'
MS2_LNA (23 mer)	5'-ATGgGTgATcCTcATgTCTtCTA-3'
BC200_LNA(21 mer)	5'-GGgGGtTGtTGcTTtGAgGGA-3'
RBC1(17 mer)	5'-TTTtTTCgGAGcTGAgG-3'
MBC1(17 mer)	5'-TTTtTCCaGAGcTGAgG-3'

*Note: lower case letters indicate LNA (Locked Nucleic Acid) positions.

2.15 Native gel electrophoresis

The native ribonucleoprotein (RNP) complexes from the S2 brain extract were separated on a composite gel containing 4% acrylamide/bisacrylamide [80:1, w/w] and 0.4% (w/v) agarose as described by (Lamond et al., 1994). Following electrophoresis, the RNPs and RNAs were electroblotted onto Nylon membrane [Ambion]. RNA was immobilized by UV cross-linking, the membrane was then hybridized with ^{32}P-labeled BC1MT, BC207, or MS2 DNA oligonucleotide probes. Hybridization was performed in 1 M sodium phosphate buffer [pH 6.5] containing 7% [w/v] SDS at 58 °C. Two washes of 20 min each was carried out at 58 °C in 0.2 M sodium phosphate buffer [pH 6.5] containing 1% SDS. Then the blot was exposed to X-ray films [Kodak MS] or screened by phosphoimager.

2.16 Localization studies and dendritic transport of tagged BC RNAs

(a) Immunocytochemistry on coronal sections of wild type, BC1 KO (BC1 RNA$^{-/-}$) BC200, BC1MS2B+MS2TAP in BC1 RNA$^{-/-}$ and BC200MS2+MS2TAP transgenic mouse brain

BC1MS2B+MS2TAP in the BC1 RNA$^{-/-}$ back ground, BC200MS2+MS2TAP transgenic mice, and as a control FVB [wild type], BC200 transgenic mice for BC1 and BC200 RNA respectively and as a negative control BC1 RNA deficient mice were used. Mice aged around six to ten weeks- were sacrificed using dry ice and transcardially perfused with freshly prepared 4% (w/v) paraformaldehyde in phosphate-

buffered saline, pH 7.2. The brains were removed carefully and fixed overnight at 4 °C in 4% (w/v) paraformaldehyde. Brain coronal sections of 30 μm thickness through the ascending arc of the hippocampus were cut on a vibrating microtome [Leica] and to facilitate handling they were further fixed in PFA for 3-7 days at 4 ºC. The free-floating sections were washed with 1 x PBS [3 x 5 min each] followed by permeabilisation for 10 min each in 10, 20, 40, 20 and 10 % (v/v) ethanol. Following blocking for 60 min [10% (v/v) normal goat serum (NGS), 0.2% (w/v) bovine serum albumin (BSA) in PBS] the sections were incubated overnight at 4 °C with primary antibody [rabbit polyclonal antibodies against PABP, HA and MS2] at 1:500 or 1:1000 dilution in 1% NGS, 0.2% BSA in PBS. The unbound antibody was removed by washing [3 x 5 min each] with PBS. This was followed by secondary antibody incubation at RT with biotinylated sheep anti-rabbit IgG for 1 hour [(1:1000), Sigma] and washing with 1 x PBS (3 x 5 min each). All washing steps were carried out at room temperature. Next, the sections were incubated with pre-formed ABC (Avidin : Biotinylated enzyme Complex) complex [ABC Elite Kit, Vector Laboratories, Inc., Burlingame, USA]. The ABC complex should be prepared 1 hour prior to incubation. The peroxidase reaction was done by incubating the sections for 10 min in 3,3'-diamino-benzidine (DAB) and hydrogen peroxide [Sigma Fast, Sigma] with nickel ammonium sulfate enhancement [0.03% (w/v)]. After mounting and cover slipping, the stained sections were photographed with a NIKON E600 microscope using a Cool SNAP digital camera [Photometrics] and Open Lab 3.0 software [Improvision, Coventry UK].

2.17 DIG-labeled *in situ* Hybridization

(a) Generation of DIG (Digoxigenin) labeled probes.

The 5' part of BC1 RNA shares sequence similarity with ID repetitive elements (Tiedge et al., 1991). In order to prevent hybridization to other RNAs that contain ID elements e.g., 3'UTRs of mRNA, we used probes that correspond to the unique 3' part of BC1 RNA generated from plasmid pMK1, which harbors 60 nt of the 3' unique part of BC1, cloned between *Kpn*I and *Sac*I restriction sites of pBluescript II KS+. A corresponding BC200 RNA probe was generated from plasmid pKK536-6, which harbors 71 nt (position 119-189) from the 3' unique portion of the BC200 gene cloned between the *Kpn*I and *Sac*I restriction sites of pBluescript II KS+. Both plasmids were linearized with either *Sac*I or *Kpn*I restriction endonucleases. DIG-labeled RNA probes were in vitro transcribed using T3 or T7 polymerases for sense (S) or antisense (AS) RNA probes, respectively, using the DIG RNA Labeling Kit (Roche Molecular Biochemicals) according to the manufacturer's instructions. The Dig-labeled RNAs were quantified in reference to the standards provided in the kit, as described by the manufacturer.

BC1MS2B+MS2TAP inBC1 RNA$^{-/-}$ transgenic mice, BC200+MS2TAP transgenic mice, wild type mice, BC200 transgenic mice and BC1 KO (BC1 RNA$^{-/-}$), all 6-10 weeks of age, were perfused

transcardially with 10 mM phosphate-buffered saline (PBS), pH 7.2 followed by freshly prepared 4% paraformaldehyde (PFA) in PBS. The brains were carefully removed and fixed overnight at 4 ºC. The next morning, coronal sections of 30 µm thickness, through the ascending arc were cut on a vibrating microtome [Leica] and, to facilitate handling, they were further fixed in PFA for 3-7 days at 4 ºC. The free floating sections were washed with PBS [3 x 5 min each], treated with Proteinase K (1 µg/ml) for 30 min at 37 ºC, incubated in freshly prepared 0.1 M glycine [2 x 5 min each] and finally rinsed with 2 x SSC [2 x 15 min each]. Subsequently the sections were pre-hybridized [50% Formamide, 1.2 M NaCl, 0.02 M Tris pH 7.5, 2 x Denhardts, 2 mM EDTA, 1 mg/ml denatured salmon sperm DNA, 1 mg/ml total yeast tRNA] at 50ºC for minimum 1 hour, in a humid chamber. This was followed by hybridization with 200 ng/ml of DIG-labeled RNA probe in hybridization buffer [pre-hybridization buffer plus 20% dextran sulfate] at 50 ºC overnight, in a humid chamber. The hybridized sections were washed with 2 x SSC (2 x 5 min each), incubated with RNase A (2 µg/ml) for 30 min at 37 ºC to remove unbound probe, washed with 2 x SSC at 50 ºC (3 x 15 min). Final washes were performed under high stringency, with 0.1 x SSC, 0.1% sodium pyrophosphate (3 x 30 min each). The sections were washed and equilibrated with PBS and blocked with blocking buffer [1% blocking reagent, Roche Molecular Biochemicals] for 60 min at RT and incubated overnight at 4 ºC with secondary antibody sheep anti-DIG-AP [Roche Molecular Biochemicals], 1:500 in blocking buffer. The excess of antibody was removed by washing with 1 x PBS (4 x 5 min). Finally the sections were equilibrated with equilibration buffer (0.1 M NaCl, 0.1 M Tris pH 9.5, 0.05 M $MgCl_2$) [2 x 5 min each], the antibody was detected enzymatically using nitro blue tetrazolium and 5-bromo-4-chloro-3-indolyl phosphate as a substrate [Boehringer Manheim]. The sections were finally washed, mounted and cover slipped with Vecta Mount [Vector Laboratories, Inc., Burlingame, USA]. Hybridization signals were photographed using a Nikon E600 microscope and Open Lab Software [Improvision, Coventry UK].

2.18 Raising antibodies against MS2 coat protein

(a) Expression and purification of Recombinant MS2 coat Protein

The gene encoding MS2 coat protein was obtained from plasmid pG14-MS2-GFP (Bertrand et al., 1998) [kindly provided by Rodert H. Singer, Albert Einstein College of Medicine, Bronx, New York]. An MS2 coat protein gene cassette was generated by PCR amplification using the following oligonucleotides.

pETMS2 forv 5'-CGCGGCAGC*catatg*CTAGCCGTTAAATGGCTTCTAACTTTACTCAG-3' and

pETMS2 rev 5'-GTGCTCGAGTTCAA*gcggccgc*GTAGATGC-3' with *Nde*I - *Not*I restriction recognition sites. The PCR fragment was digested with *Nde*I/*Not*I and cloned into the pET28b+ vector (Invitrogen). Italicized small letters indicate *Nde*I/*Not*I restriction recognition sites.

Recombinant MS2 coat protein was expressed with an N-terminal histidine tag (Hochuli, 1990) for convenience of purification. In order to achieve this, plasmid pET28b+ MS2 was transformed into *E.coli* BL21 (DE3) pLysS cells [Novagen]. A single colony was inoculated in 10 ml of Luria-Bertani broth [10 grams of tryptone, 5 grams of yeast extract and 10 grams of sodium chloride in one liter of deionized water, with antibiotic kanamycin (100 µg/ml)] overnight. The overnight culture was transferred into one liter LB containing kanamycin (100 µg/ml) and chloramphenicol (34 µg/ml) and grown to an optical density reached of between 0.4 and 0.6. Transcription was induced with 100 mM IPTG (isopropyl-β-D-thiogalactopyranoside) and cells were grown for an additional 3 hours, followed by harvesting by centrifugation at 5,000 x g for 10 min at 4 °C, resuspension in 15 ml of buffer containing 50 mM HEPES [pH 7.5], 250 mM NaCl and 1 mM $MgCl_2$ and once more pelleting at 5,000 x g for 10 min at 4 °C and storage at -80 °C.

Expressed protein was affinity purified with a Ni-NTA matrix according to the recommendations of the manufacturer [Qiagen, Hilden]. Briefly, the bacterial pellet, which expresses histidine tagged MS2 recombinant protein was suspended in 10 ml of lysis buffer [20 mM HEPES pH 7.5, 250 mM NaCl, 1 mM $MgCl_2$, 20 mM imidazole one tablet/50 ml protease inhibitor mix “complete”, Boehringer Mannheim]. Cells were sonicated with 2 - 3 pulses (100 W - 20 kHz), and the soluble and insoluble fractions were separated by centrifugation at 15,000 rpm (Sorvall) for 1 hour at 4 °C. The resulting supernatant was bound to Ni-NTA resin (previously equilibrated with lysis buffer) for about 1 hour at 4 °C with gentle mixing. The resin [matrix] was washed with 30 ml buffer containing 20 mM HEPES pH 7.5, 500 mM NaCl , 1 mM $MgCl_2$, 40 mM imidazole and the protein was eluted from the resin with 10 ml of elution buffer [20 mM HEPES pH 7.5, 200 mM NaCl, 1 mM $MgCl_2$, 500 mM imidazole]. The protein was identified by SDS-PAGE. Contaminant free fractions were pooled, dialyzed against 20 mM HEPES, pH 7.5, 100 mM NaCl, 1 mM $MgCl_2$, 50% glycerol and concentrated by dialysis and stored at -80 °C.

(b) Generation of MS2 antibody

Purified histidine tagged MS2 coat protein [300 µl, 5 - 10 mg of protein] was mixed with equal volume of Freund’s incomplete adjuvant (Sigma-Aldrich). The homogenized mixture (water-in-oil emulsion) was injected [subcutaneous/intramuscular] into two New Zealand white rabbits to raise antibodies. After 45 days, the sera were collected and designated as JB 43 and JB 44. This serum was tested for its immunoreactivity and specificity against MS2 coat protein by immunoblot analysis, as well as with brain extracts (S2 fraction) of MS2TAP transgenic mice, BC1MS2B+MS2TAP_in BC1 $RNA^{-/-}$ background and BC200MS2+MS2TAP transgenic mice, as well as wild type and BC1 $RNA^{-/-}$ knockout mice. Antibody was purified using the protein G midi antibody purification kit [Proteaus] for immunohistochemistry studies.

2.19 Purification of BC RNPs using TAP methodology from [BC1MS2B+MS2TAP and BC200MS2+MS2TAP transgenic mouse brain extracts] – an initial attempt

The TAP [Tandem Affinity Purification] protocol (Rigaut et al., 1999) was used with some modifications. Mice brain extracts were prepared as described (Materials & Methods, section 1.13) with some modifications i.e. buffer [20 mM Tris-HCl, pH 7.5, 100 mM KCl, 3 mM $MgCl_2$, 10 mM β-mercaptoethanol, 20 mM NaF, 1.mM phenylmethanesulfonylfluoride (PMSF), 1 mM DTT, 0.25% NP40 and one tablet protease inhibitor mix "complete", Boehringer Mannheim/50 ml of buffer]. The extract was centrifuged at 1,750 x g for 11 min at 4 °C. The supernatant was collected (S1) and centrifuged at 15,000 x g for 1 hour at 4 °C. The resulting supernatant (S2) was stored at -80 ºC or further centrifuged at 60,000 x g for 2 hours at 4 °C. The resulting supernatant (S100) was stored at -80 ºC until further use. The centrifuged (S100) supernatant was mixed with 500 µl of IgG Sepharose beads [Amersham Biosciences, Piscataway, NJ, USA] and incubated at 4 °C for 4 hours with gentle rocking. The mixture was loaded onto a disposable polyprep chromatography column [10 ml bed volume, Bio Rad laboratories, Hercules, CA,USA] and washed once with 15 ml buffer B [25 mM Tris-HCl, pH 7.5, 150 mM KCl, 3 mM $MgCl_2$, 10% sucrose, 10 mM β-mercaptoethanol, 20 mM NaF, 1.mM phenylmethanesulfonylfluoride (PMSF), 1 mM DTT, 0.25% NP40 and one tablet/50 ml protease inhibitor mix "complete", Boehringer Mannheim]. The column was equilibrated with 5 ml TEV cleavage buffer [25 mM Tris-HCl, pH 7.5, 150 mM KCl, 3 mM $MgCl_2$, 1mM DTT, 0.5 mM EDTA, 0.1% NP-40 and protease inhibitor mix "complete", Boehringer Mannheim] and the TAP-tagged protein (minus the IgG domain) was released from the resin by digestion with 500 U of rTEV protease [Invitrogen] in TEV cleavage buffer containing 1 µM E-64 protease inhibitor for 1 hour at 16 ºC and overnight at 4 ºC. The eluate was then bound to the calmodulin (CAM) agarose beads [Stratagene, La Jolla, CA, USA] in CAM-binding buffer [CBB; 25 mM Tris-HCl, pH 7.5, 150 mM KCl, 3 mM $MgCl_2$, 1mM Mg acetate, 1 mM imidazole, 2 mM $CaCl_2$, 10 mM ß-mercaptoethanol, 0.1% NP-40 and protease inhibitor mix "complete", Boehringer Mannheim] (Rigaut et al., 1999) and eluted with CBB buffer containing 10 mM EGTA. The eluate was TCA precipitated and loaded onto a 10% polyacrlamide gel for SDS-PAGE. Proteins were visualized either by commassie or by silver staining.

2.20 TEV protease expression and purification

(a) Expression of histidine tagged TEV protease

The TEV protease expression plasmid [pMHTdelta238] was obtained from the Harvard plasmid collection, (Blommel and Fox, 2007). TEV protease was expressed in *E.coli* BL21 RILP strain (Stratagene, La Jolla, CA) with some modifications. Starting inoculums were grown over night at 25 °C in 10 ml LB containing kanamycin (100 µg/ml), chloramphenicol (34 µg/ml), 0.375% aspartic acid and 25

mM sodium phosphate. About 2 ml of overnight culture were transferred to one liter of terrific broth [12 grams of tryptone, 24 grams of yeast extract and 4 ml of 100% glycerol in 900 ml of deionized water, sterilize by autoclaving and adjust volume to one liter with sterilized 100 ml solution of 0.17 M KH_2PO_4 and 0.72 M K_2HPO_4 grams of sodium chloride in one liter of deionized water Expression was carried out in terrific broth containing kanamycin (100 µg/ml), chroamphenicol (34 µg/ml), 0.8% glycerol with 2 mM $MgSO_4$ and 0.375% aspartic acid. For induction final concentration of 0.5 mM IPTG was added and the induction was carried out for about 10-12 additional hours at 25 °C. The cells were harvested by centrifugation at 5,000 x g for 10 min. The pellet was resuspended in 10 ml of a buffer containing 50 mM sodium phosphate pH 7.5, 300 mM NaCl and 20% ethylene glycol, and centrifuged again to recover the washed cell paste, which was stored at -80 ºC until further use.

(b) Purification of histidine tagged TEV protease

The cell paste was resuspended in 10 ml of 20 mM Tris-HCl, pH 8.0, containing 150 mM NaCl, and 0.3 mM DTT (dithiothreitol). The cells were lysed using sonication (two times about 2 min pulses, 100 W – 20 kHz) and centrifuged at 15,000 x g for about 1 hour at 4 ºC. Supernatant was collected, passed through the Ni-NTA matrix as described under Materials and Methods section 1.18. Pure protein fractions were identified by SDS-PAGE, pooled, dialyzed in 20 mM Tris-HCl, pH 7.5, 100 mM NaCl, 1 mM $MgCl_2$, 50% glycerol and 5 mM DTT, concentrated by dialysis and stored at -80 °C.

2.21 Silver staining of protein gels

After electrophoresis, protein gels were kept in fixation solution [40% methanol, 5% acetic acid and 55% water] for 30 min at room temperature, and then washed with water for one hour with several changes (about 3-4 changes). After sensitisation using 0.02% sodium thio-sulphate solution for 1 min, gels were impregnated with 0.1% silver nitrate solution at 4 ºC for 30 min. Gels were developed using 2% sodium carbonate solution with 0.04% formaldehyde. The reaction was stopped using 1% acetic acid solution.

2.22 Coomassie staining of protein gels

The protein gels were incubated with coomassie staining solution [1% Coomassie Brilliant Blue (CBB) R-250, 45% methanol and 5% acetic acid] for 1-2 hours at room temperature and destained using 45% methanol and 5% acetic acid solution for several hours at room temperature.

2.23 Immunoblotting of protein gels

After electrophoresis, gels were blotted onto the PVDF membranes [polyvinylidene difluoride] (Hybond-P, Pharmacia] using a semi dry transfer system [Bio-Rad]. The blots were blocked using 1x TBST [20 mM Tris-HCl pH 7.4, 0.8% NaCl and 0.1% Tween-20) containing 5% non-fat dry milk powder, overnight at 4 ºC. The membranes were incubated with primary antibody for 1-2 hours followed by 3 x 15 min wash with 1x TBST, followed by HRP conjugated secondary antibody incubation (1:5000) for 45 min at room temperature in blocking solution (1 x TBST + 5% non-fat dry milk powder). Membranes were washed 3 x 15 min with 1 x TBST. Membranes were developed using the Enhanced Chemiluminescence ECL Plus kit [Pharmacia] according to the manufacturer's instructions and exposed to X-ray films from 2 - 10 mins depending on the sensitivity and signal intensity.

2.24 ELAV/Hu [Embryonic Lethal, Abnormal Vision] protein expression and purification

(a) Expression and purification of histidine tagged ELAV or mouse Hu proteins

cDNAs encoding each mHu family members [mHuB, mHuC and mHuD] were cloned into pET28a (+) [Novagen, Madison, WI] such that each coding sequence of Hu proteins was protein was preceded by a histidine tag at the N-terminus. After transformation of the construct into *E coli* BL21 (DE3) pLysS, bacteria were grown in LB broth at 37 °C for 3 hours. Following induction with 1 mM isopropyl-β-D-thiogalactopyranoside [IPTG], the cells were grown for an additional 3 hours. Cells were harvested by centrifugation at 6,000 x g for 10 min at 4 °C and the cell pellet was stored at -80 °C. Fusion protein was purified by nickel chelation chromatography according to the manufacturer's instructions (Qiagen) as described under Materials and Methods, section 1.18 a.

(b) Purification of histidine tagged Hu proteins

The cell pellets were re-suspended in small volume of 20 mM sodium phosphate, pH 8.0, containing 150 mM NaCl, 0.3 mM DTT [dithiothreitol]. The cells were disrupted by sonication for about 2 min pulses for two times and the cells were centrifuged at 13,000 rpm for about 1hour at 4 ºC. Supernatant was collected and passed through the Ni-NTA resin. The proteins bound to resin was washed with 20 mM sodium phosphate buffer, 500 mM NaCl, 20 mM imidazole and 0.3 mM DTT [dithiothreitol] to remove contaminants. Finally, recombinant protein was obtained with elution buffer containing 20 mM sodium phosphate buffer containing 500 mM NaCl, 500 mM imidazole and 0.3 mM DTT [dithiothreitol]. The purified protein fractions were identified by sodium dodecyl sulfate gel electrophoresis, pooled, dialyzed, and concentrated.

2.25 Immunoprecipitation

Anti-HuB, anti-HuC and anti-HuD antibodies were immobilized on protein G-Sepharose beads [Pharmacia] by gently mixing at 4 ºC overnight, washed three times with buffer [12.5 mM HEPES pH 7.9, 150 mM NaCl, 2 mM $MgCl_2$, 0.25 % NP-40, 5% glycerol]. Followed by incubation with brain extracts (S2) of FVB wild type mice and BC200 [-2271] transgenic mice at room temperature. The unbound sample was washed with same buffer, RNA was extracted from the bound complexes using TRIzol reagent (Gibco BRL). The RNA was separated on 8% polyacrylamide/7 M urea gels and transferred to nylon membranes.

2.26 Electrophoretic Mobility Shift Assays (EMSA)

(a) *In vitro* transcription of BC1 RNA and BC200 RNA

DNA templates for *in vitro* transcription of BC1 and BC200 RNAs were generated from pBCX 607 (Cheng et al., 1996) and pBC200 (Bovia et al., 1997) plasmids, respectively. Plasmid DNA was digested with *Dra*I restriction endonuclease that leaves a TTT sequence at the 3' end, yielding genuine RNAs that terminate in UUU. Transcription with T7 RNA polymerase was performed in a 100 µl reaction volume containing 20 µl 5 x transcription buffer [MBI Fermentas], 2.0 mM (final concentration) of rGTP, rCTP and rATP, each and 0.5 mM rUTP, 100 µCi of [α-^{32}P] UTP (Perkin Elmer), 20 units of RNasin [Fermentas] and 200 units of T7 RNA polymerase [Fermentas]. After incubation at 37°C for 2 hours and 30 min, the DNA template was digested by adding 20 units of *DNase*I [RNase free, Roche] for 15 min and precipitated with EtOH. ^{32}P-labeled RNAs were separated on 8% polyacrylamide/7 M urea denaturing gels, and passively eluted into 0.3 M NaOAc buffer pH 5.2, 1.0 mM EDTA, 0.2% phenol at 4 ºC overnight. Subsequently, RNAs were ethanol precipitated and dissolved in 30 µl H_2O.

(b) Gel shift assays.

Radiolabeled RNAs were heat denatured at 65 ºC for 5 min and placed on ice for 5 min. Gel shift assays were performed in 20µl of total reaction volume, with the following buffer: 10 mM Tris-HCl [pH 7.5], 2.5 mM $MgCl_2$, 100 mM KCl, 15% (v/v) glycerol, 0.5% Triton X-100, 1 µg of BSA (Sigma-Aldrich), 1 µg of total yeast tRNA (Sigma-Aldrich), 20 µg of heparin and 5 units of ribonuclease inhibitor (RNasin, MBI). 10 nM of RNAs were incubated with increasing concentrations of Hu proteins [mHuB, mHuC and mHuD] and as a positive control 15 nM of PABP [Poly(A)-Binding Protein]. Reactions were incubated for 30 min at 4 ºC or on ice, followed by 10 min at room temperature. Reactions were separated on non-denaturing, 6% polyacrylamide/bis-acrylamide (29:1, w/w) gels in 0.5 x TBE [45 mM Tris-borate pH 8.3, 1 mM EDTA] at 200 V for 120-180 min depending on RNA length. Gels were dried using vacuum and heated slab gel dryer (Bio-Rad) at 80 ºC for 2 hours and exposed to X-ray films.

2.27 DNA extraction from different rodents

DNA was processed by standard protocols (Maniatis et al., 1982) from tissues of the following rodent species: Cricetidae: *Cricetus cricetus* (black-bellied hamster), *Ondatra zibethicus* (muskrat); Muridae: *Meriones unguiculatus* (Mongolian gerbil); Spalacidae: *Nannospalax ehrenbergi* (Ehrenberg's mole-rat); Dipodidae: *Jaculus jaculus* (lesser Egyptian jerboa); Pedetidae: *Pedetes capensis* (springhare); Anomaluridae: *Anomalurus* sp. (scaly tailed squirrel); Castoridae: *Castor fiber* (Eurasian beaver); Heteromyidae: *Heteromys gaumeri* (Gaumer's spiny pocket mouse); Ctenodactylidae: *Ctenodactylus gundi* (northern gundi); Thryonomyidae: *Thryonomys swinderianus* (greater cane rat); Echimyidae: *Proechimys cuvieri* (Cuvier's spiny rat); Myocastoridae: *Myocastor coypus* (nutria); Octodontidae: *Octodon degus* (degu); Caviidae: *Cavia porcellus* (domestic guinea pig), *Dolichotis patagonum* (Patagonian mara); Erethizontidae: *Coendou prehensilis* (Brazilian porcupine); Sciuridae: *Sciurus vulgaris* (Eurasian red squirrel), *Marmota marmota* (European marmot); Gliridae: *Glis glis* (fat dormouse).

2.28 PCR amplification of informative presence/absence loci

PCRs were performed for 4 min at 94 °C followed by 35 cycles of 30 s at 94 °C, 30 s at the primer-specific annealing temperature and 60 s at 72 °C. The PCRs were finished with 5 min at 72 °C. For complicated PCR amplifications, gradient PCRs were used with annealing temp from 45 °C to 65 °C for about 30 s to 60 s depending upon the fragment length. For most of the PCR amplifications were carried out using my own lab made Taq polymerase. The PCR fragments were purified on agarose gels, ligated into the pDrive cloning vector [Qiagen, Hilden] and electroporated into TOP10 cells [Invitrogen, Groningen]. Inserts from each of three individual clones were sequenced using the Ampli Taq FS Big Dye Terminator Kit [PE Biosystems, Foster City] and standard M13 forward and reverse primers.

2.29 Computational Strategies to identify the informative loci

To identify phylogenetically informative presence/absence patterns of retroposed elements, we applied five different in silico search strategies, which will be discussed in this chapter.

The high rate of substitution in rodent genomes presents a though challenge in the development of special strategies to identify phylogenetically informative presence/absence loci. Highly conserved flanking regions are necessary for the successful PCR amplification of orthologus loci in diverged species. So previuosly developed CPAL [Conserved Presence/Absense Locus-finder], a fully automated computational search program utilizing Bioperl, which was discussed in detail (Farwick et al., 2006). The objective of the program is to locate mouse SINEs inserted in introns and to identify conserved regions in the flanking exons in order to design universal rodent PCR primers. CPAL is organized in four main

processes. 1. An NCBI GenBank search for annotated exon-intron-exon structures in the species of choice, in our case *Mus musculus*. Intron sequences no larger than 1 kb with their flanking exons are extracted for further analyses. (2) A local RepeatMasker screen of the selected sequences for interspersed repeats, in our case all SINEs known in rodents (3) A StandAloneBlast alignment of the mouse exons against sequences of a reference species of choice, in our case all human chromosomes, to reveal loci conserved in both mouse and human. (4) A Clustal W (Thompson et al., 1997) alignment of all mouse exon-flanked repeat loci against the human reference sequences from point 3.

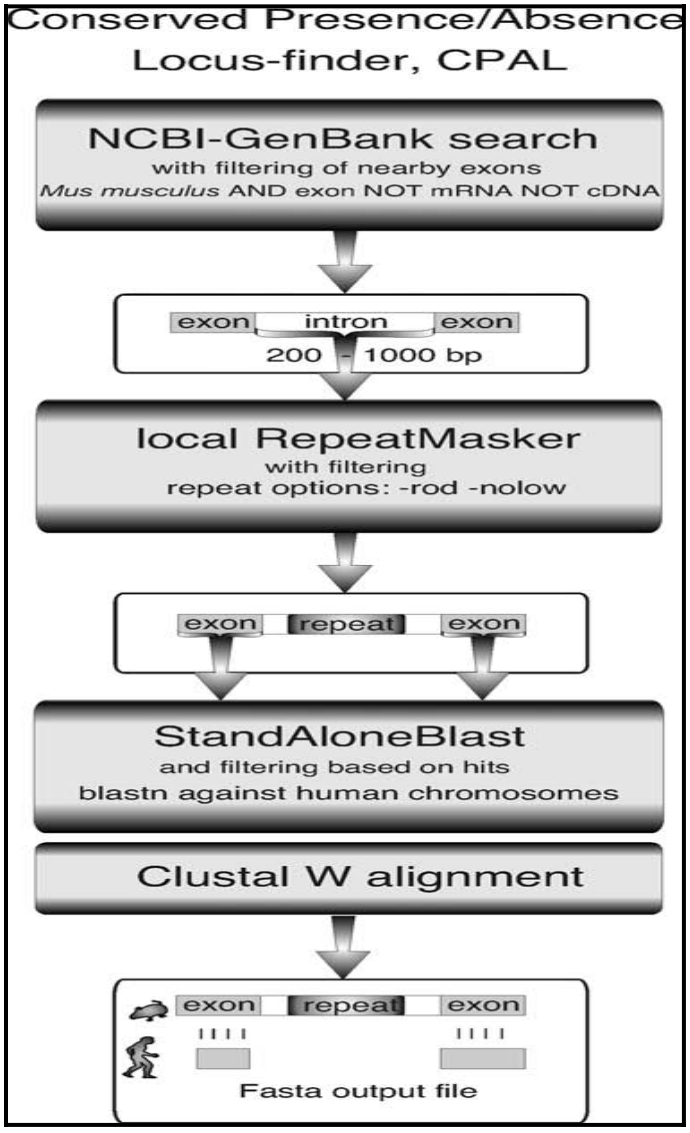


FIGURE 2.3 Flow chart of the Conserved Presence/Absence Locus finder computational strategy. The CPAL Bioperl script begins with a GenBank search using the query "*Mus musculus* and exon" and excludes mRNAs and cDNAs. All hits are scanned for introns not longer than 1 kb, flanked by exonic sequences. With a local version of the RepeatMasker, CPAL selects introns with rodent-specific SINEs (-rod) and excludes sequences of low complexity (-nolow). A StandAloneBlast search of the exonic flanks against all human chromosomes reveals loci conserved in mouse and human. A CLUSTAL W alignment produces a FASTA output file with potential presence/absence loci in rodent species flanked by highly conserved regions facilitating the generation of conserved primers for Zoo-PCR in rodents. Figure was adapted from (Farwick et al., 2006).

Strategies II, III and IV

Mouse intronic sequences were downloaded from the Santa Cruz Server (http://genome ucsc.edu/cgi-bin/hgTables).

After applying script based exclusion of duplicated sequences and the introns larger than 1 kb to facilitate PCR amplification, searched for the presence of retroposed SINE elements for rodents (RepeatMasker). Mouse introns with rodent specific elements were selected and/or low complexity repeats were excluded. The remaining intronic sequences were analyzed for the presence of conserved flanks (Santa Cruz Server) and were analyzed by searching against the trace data base of the guinea pig, kangaroo rat and spermophilus for the respective strategies II, III and IV at the NCBI server (http://ftp.ncbi.nih.gov/pub/TraceDB), and selected 50, 20 and 30 loci to generate PCR primers.

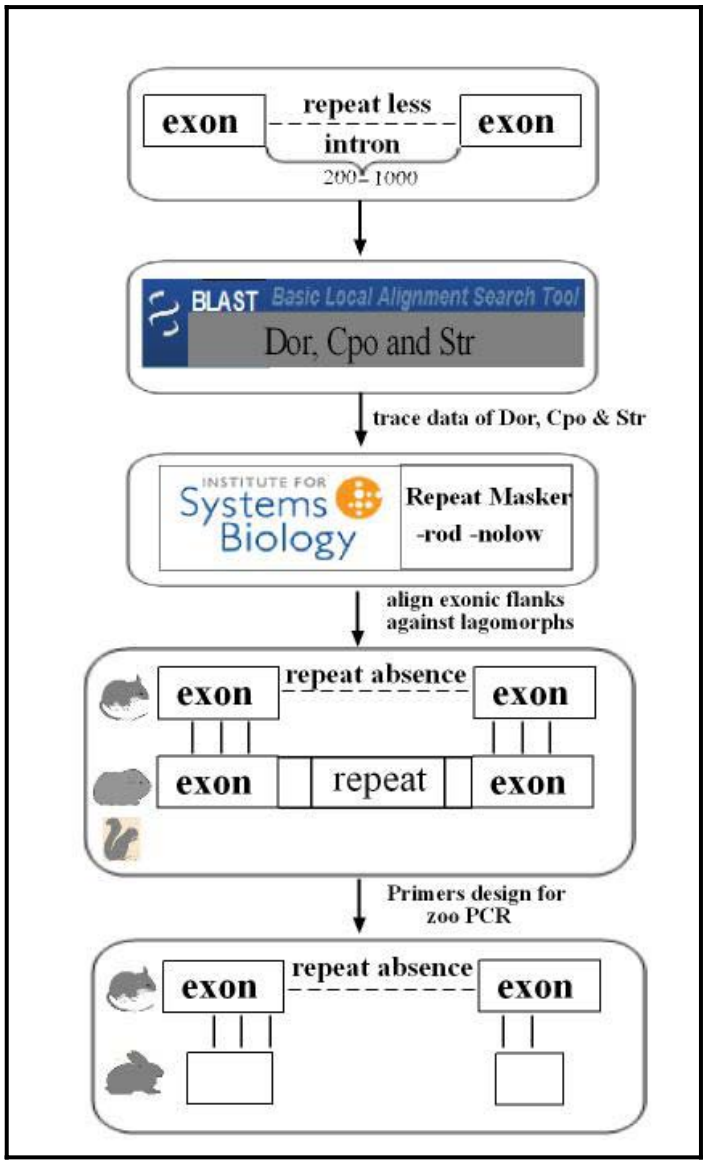


FIGURE 2.4 Alternative strategies to find informative loci. In this strategy 25,000 intronic sequences of *Mus musculus*, has taken with flanking exons and the intronic sequences lack repeats (SINE elements). All intronic sequences were blasted against *Dipodomys ordii* (Dor), *Cavia porcellus* (Cpo) and *Spermophilus tridecemlineatus* (Str). All hits were scanned for introns not longer than 1 kb, flanked by exonic sequences. With a local version of the RepeatMasker, strategies selects introns with rodent-specific SINEs (-rod) and excludes sequences of low complexity (-nolow). A Blast2 sequence alignment produces a file with potential presence/absence loci in rodent species flanked by highly conserved regions facilitating the generation of conserved primers for Zoo-PCR in rodents. Finally amplified sequences were aligned against lagomorphs (used as an our group).

For finding informative loci in other representative clades from guinea pig, kangaroo rat and spermophilus, where trace data is available in NCBI GenBank for searching informative loci. In these strategies, we have taken 25000 mouse intronic sequences of about 1 kb, with highly conserved exons of 150-200 nt length. We blast these intronic sequences against the trace data of these representative species (guinea pig, kangaroo rat and spermophilus). The resulting sequences were checked for repeats in local repeat masker where we selected SINEs in rodents. Next step is manual alignment of all available representative species [human and rabbit (lagomorphs)] sequences. Design the conserved primers for PCR amplification in rodents.

2.30 Presence/Absence Analysis of SINE retroposed elements

We applied the RepeatMasker option (-rod) to all our strategies in this study. Thereby focusing only on rodent-specific and mammalian wide repeats. Thus, all selected loci represents the presence state should contain SINE retroposed elements in the intronic sequence after PCR amplification. Conversely, the absence state should have only flanking exons without the retroposed elements therefore appear as smaller PCR products. All species contain a transposed element at the orthologous position were considered to have derived from a common ancestor.

For a handful of loci, the PCR patterns were difficult to interpret because of additional independent SINE insertions or random indels or deletions in different lineages. The sequencing of all PCR products was thus an important criteria to characterize specific presence/absence patterns.

Table 2

marker designation	oligonucleotide sequence (5'-3')	Tm
CPAL 25	CAGCCATGGTCCTTTTCACCAG	55 °C
	CCAGACCCWCCATTRGAGAGTGA	
CPAL 405	GATGGTTCTGCKGCCTATGTTCA/TGCTTGACCATCCTCTAGAC	54 °C
	GTCCCGTCTTCTARCTGYACTG/ATGTGCCCATACCTAAAAGTAGT	
CPAL 437	GATGAGAAACCTCTTCAGGACA	55 °C
	AGCACTTCTAGCTGAAGTGACAAGC	
CPAL 1406	YAGGAGCCGACCCYKCTGAG	54 °C
	AMGCCCTCCTTCACCATCCAGTGTCT	
CPAL 503	AGTGGAARTTCACAGRCTGCTCAG/ TAGGGCAGTTGAGATGTGTGAGA	52 °C
	GCAGGAAGRAGTCTTCACAAGG/AGTGGAGGGTGAGGGGGTAAG	
17 B3A	CTAGATACMCCMGAGATGAGRCG	58 °C
	GTGTGTTTTTCTTTACTCGTTCAGTG	
18 ID4	GGTCAAGACGAGATACATGAACTCTGC	55 °C
	CCAAGCGGAGAAAGGARGGCA	
19 IDL-Geo	AGACTGCCGGGTGTATGGGAG	56 °C
	TGGCCAGGGCCCGGATCT	
20 PB1-Geo	AGGGTGAACGAGAGCTGCA	55 °C
	GTAGGGATGGAATTCTCAGCTG	
21 PB1-Geo	TCTCAGAGACAGAGGAAGTGGC	54 °C
	GGAGGCTGAGCTATACTCTTGG	
22 PB1-Geo	GATCGAAATGGAGATGTACGAAA	51 °C
	AGGCTTTGCTGGCATGTTAG	
23 IDL-Geo	GATTCAGTTTTGCAGCCCTA	54 °C
	AAGAGTAATTGACGTGTGATATGG	
24 PB1-Geo	AATGGGGGTCACTGTCAGAA	55 °C
	TCGATGTAGCACTGAGCACC	
25 B1F	CTGTCAAGACAGCTGTGTCAG	54 °C
	ATGTCTCCGGATCTGAGAGAT	
26 B1 & B1F	GCTGTGGACATATTCGACTTCTG	54 °C
	GTCCGATGATCCCACAGAAGA	
27 B1 & ID2	AGCGTGGCTGCCAAGAACCT	56 °C
	CTGGCGCAAGGCATCATTGT	
28 ID4	TGGGAAAGATCTCAGCAATAC	50 °C
	ATGTATGTTAGCTCCTCAATTTC	
29 B1	AAATTGCCATGTTGCTGTGAT	53 °C
	CTGTCTTATAGACGCTTCCGTG	
30 ID4	GTTTCATCCCTACATCCCCTT	54 °C
	TCCATGAAGGCCTCATAGAAAT	
31 ID4	GTTTATGACCAGGTGGAGTTTGA	55 °C
	ACTGTTCATGATTCGAAGGGACT	
32 ID2& ID2& ID4	TGAGATTTGTCGGCTGTCTGTTTT	53 °C
	GTGCAACGATCAGCCAAATTTC	
33 ID-Str	CCCCGTGGTGAGCGAGAA	52 °C
	AACTTTTTGCCACTCTTCCCAG	
34 PB1	TTCATTGAAAACTCTTCGTTATGG	51 °C
	AGAAGAGATGGCCAATTGTGAT	
35 ID4	GATGGTGTGTGTGTCCGTGT	55 °C
	AGAGGATAAGGAACATTGAAGTAGT	

RESULTS

Establishing and characterization of the tagged BC RNAs in transgenic mice for localization and purification studies

In order to purify the BC RNP complexes and to examine the tagged RNA transport *in vivo*, we have to establish a transgenic mice, which express BC1MS2 RNA and BC200MS2 RNA in combination with MS2TAP [MS2 coat protein fused to TAP tag for the affinity purification of BC RNP complexes] protein-expressing transgenic mice. Therefore, we have tagged MS2 protein binding sites to the 3' portions of the BC RNAs in combination with constructs expressing an MS2 coat fusion protein for binding and as a handle for purification.

Results discussed in this section:

(I) How we generated the MS2 tagged BC RNAs in transgenic mice.

(II) MS2 tagged BC RNAs transgenic mouse model for dendritic BC RNAs transport studies.

(III) Transgenic mouse models for BC RNP purification studies and an initial test towards purification.

3.0 Affinity purification as a strategy for purification of BC RNPs.

Protein affinity tags have been used routinely and successfully for either purification or detection of the tagged proteins. To study the RNA or RNP [RNA-protein complexes], RNA affinity tags are potentially very useful to purify the RNP complexes. For purification or isolation of native *in vivo* RNP complexes, it is useful to have an RNA affinity motif that can be incorporated during the synthesis of the RNA *in vivo*.

Earlier RNA affinity tags have not been widely used [like protein affinity tags] successfully to isolate or to purify the protein subunits of the RNP complexes and for localization or transport studies. Recently researchers started using RNA based affinity tags to purify the RNP complexes or to locate the tagged RNAs (Hogg and Collins, 2007, Srisawat and Engelke, 2001). These RNA tags can be classified into two categories. In the first category, tags that occur naturally were based on RNA - protein interactions found in nature. RNA motifs bound by MS2 phage coat protein or λ-phage N anti-terminator protein (N-peptide) are the prominent examples of this type. In second category, tags were developed based on the artificially selected RNA motifs that have small-defined structures and were particularly useful for RNP isolation. In this category, *in vitro* synthesized tags like streptavidin, tobramycin and D8 sephadex were the most widely used RNA tags (Bachler et al., 1999, Hartmuth et al., 2004, Srisawat and Engelke, 2001, Srisawat et al., 2001). Since our goal is to purify BC RNP complexes to near homogeneity and to study the dendritic transport of the tagged RNAs using antibodies in electron microscopy in the future, we had developed RNA based affinity tag. By utilizing the MS2TAP tag, we are assuming to purify BC RNP complexes *in vivo* or to identify the proteins bound to these RNPs and to locate the tagged RNAs transport for assigning the function to these RNAs.

Several factors were considered, when we chose the MS2 tag over other RNA tags for purification and localization of BC RNPs. MS2 tag is a naturally available tag. It is very selective and efficient in binding to MS2 coat protein. Singer et al speculating, that MS2/MS2 coat protein can bind to one RNA/RNP at a time compared to PP7 tag (Chao et al., 2008, Kaganman, 2008), which is the utmost important thing for selecting MS2 tag. The interaction between MS2 coat protein and its cognate RNA has been extensively characterized (Horn et al., 2004, Johansson et al., 1998, van den Worm et al., 1998).

3.1 Tagging the BC RNAs with MS2 binding sites for localization and purification

Several factors like folding, steric hindrance and not losing the tag from BC RNAs via processing for purification and localization were considered while tagging MS2 motifs to the BC RNAs. These issues were discussed briefly below.

(I) **Folding:** It is recommended that the RNA affinity tag should be inserted at either 3' end or 5' end of the interested RNAs/RNPs to be purified or to be visualized. So that, the access of tag for purification and visualization could be substracted. The folding problem is simply a matter of inserting the tag in such a way that both the tag and the RNAs should remain correctly folded. RNA folding algorithms; (Zuker, 2003) http://www.bioinfo.rpi.edu/applications/mfold) has been used to predict the folding of these BC RNAs to obtain the MS2 tagged RNAs for proper folding. Figure 1.8. depicting the possible folding of BC1 RNA with MS2 motif.

(II) **Steric hindrance:** This problem arises, when the tag is partially or fully covered the RNA or its associated proteins of our interest, thus obstructing access of the tag to the affinity matrix, to avoid any hindrance we had placed a spacer between the tag and the RNA.

(III) **Processing:** The major problem is keeping the tag on the RNA from being processing for purification or localization purpose, that should be tested in order to check for RNA processing in the cell or *in vivo* models like transgenic mouse models.

3.2 Placing the MS2 protein binding site at the 3' end of BC RNAs

Bacteriophage MS2 RNA motif is a 19 nt long, whose three-dimensional structure with the corresponding MS2 coat protein are well known from the crystal studies (Horn et al., 2004, van den Worm et al., 1998). Using mfold program for RNA folding we had inserted two MS2 motifs (stems) to the 3' end of the both RNAs **(Fig 3.1)**, and has been cloned in pCR TOPOII vector [invitrogen] under its native BC1 and BC200 RNA gene promoters respectively. The predicted model in the **(Fig 3.0)** showing the MS2 tag located at the very 3' end of both BC RNAs but proximal to the RNA polymerase III terminator.

Note: In case of BC1 RNA, we have cloned rat BC1 gene, in order to distinguish between rat and mouse BC1 RNA, expressed in the transgenic mice.

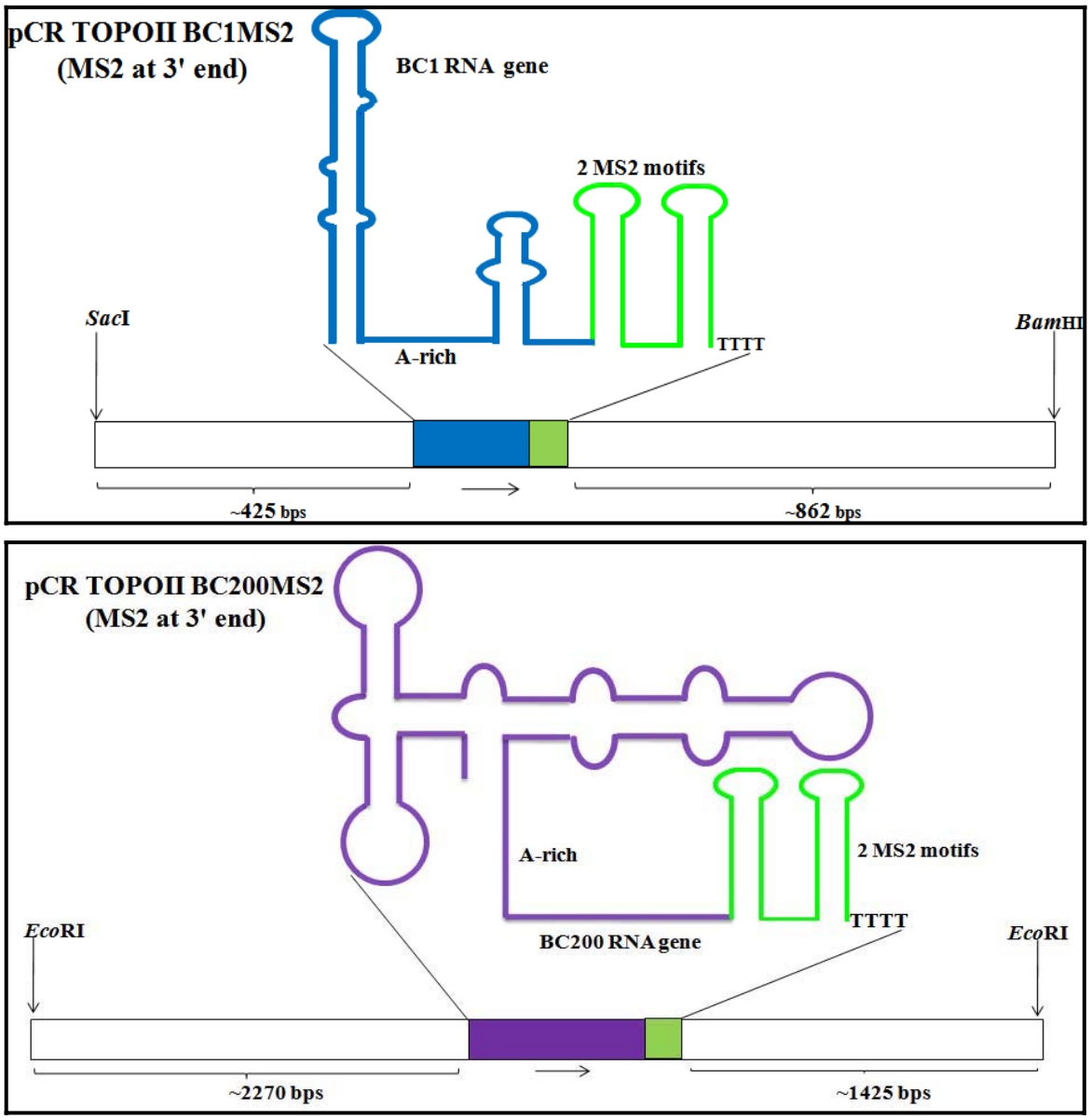


Fig 3.0 Structures of BC1MS2 RNA (MS2 at 3' end) and BC200MS2 RNA (two MS2 stems located at the 3' end). Cloned in pCR TOPOII vector for generating transgenic mice. Both RNAs are transcribed from their native promoters. The two MS2 motifs were placed proximal to the RNA polymerase III terminator (i.e. ending in UUU in BC1 and BC200 RNA).

Mfold application showed relatively efficient folding with minimum free energy (ΔG). From the mfold application, we are predicting the MS2 stem of BC RNAs might be free to bind the MS2 coat protein. Thus the MS2 tag will allow purification and dendritic transport studies in the transgenic mouse models employing antibodies for electron microscopy.

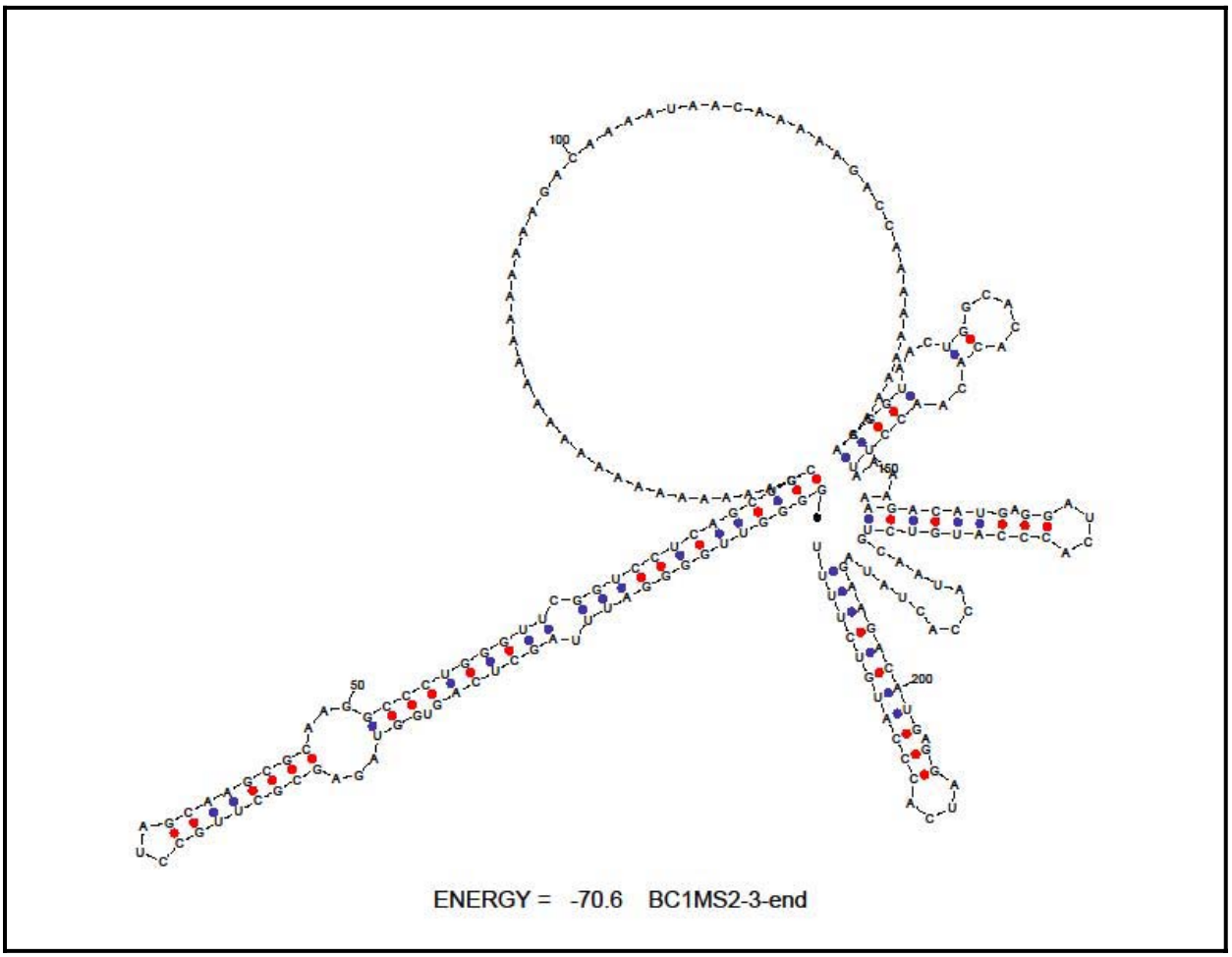


Fig 3.1 Secondary structure of BC1 RNA tagged with MS2 stem motifs. Two MS2 motifs separated by an unstructured spacer were inserted at the very 3' end of the BC1 RNA. This is followed by a stretch of U-residues for RNA polymerase III transcription termination. Figure depicting folding of BC1MS2 by mfold from Michael Zuker http://www.bioinfo.rpi.edu/applications/mfold.

3.3 TAP tag fused to MS2 coat protein for affinity purification of MS2 tagged BC RNPs

The MS2 coat protein fused to TAP tag [Tandem Affinity Purification] for affinity purification of BC RNPs, was cloned into pNE vector (a generous gift from Dr. Stephan Kindler, University of Hamburg) under the chicken β-actin promoter. TAP method involves the fusion of the TAP tag to the MS2 coat protein and introduction into the host cell and observes the expression of fusion protein in the mammalian cell lines (*ex vivo*) or in transgenic mouse system (*in vivo*) [mouse brian extract]. The fusion protein MS2 with TAP tag is approximately 357 amino acids long corresponding to 39,329 Da, and at the N-terminus of the MS2TAP fusion protein HA and FLAG tag peptides were fused to observe the protein expression by Western blot analysis.

3.4 Expression of BC1MS2 RNA and MS2TAP fusion protein from combined construct in NIH 3T3 cells

In order to facilitate our work and to avoid generating two different transgenic mice, we had generated a construct, which harbors both BC1MS2 and MS2TAP expressing modules in a single plasmid. The rat BC1MS2 fragment was PCR amplified from *Sac*I to *Apa*I from the pCR TOPOII BC1MS2 plasmid and subcloned in pBluescript II KS+ vector backbone. The MS2TAP fusion protein is driven by the chicken β-actin promoter from pNECKu vector was subcloned in the same vector that harbors BC1MS2 module. The MS2TAP insertion was done by blunt end cloning, resulting in two different kinds of orientations with respect to the plasmid backbone **(Fig 3.2)**. The MS2TAP fragment inserted into pBS-BC1MS2 was in the same transcriptional orientation as BC1MS2 [i.e. head to tail] resulted in plasmid designated, 11 dirTAP. The plasmid with MS2TAP inserted in the opposite orientation to the BC1MS2 [head to head], resulted the plasmid designated as 3 revTAP **(Fig 3.2)**. Both the plasmids were transiently transfected in NIH 3T3 cells and observed for protein and RNA expression **(Fig 3.3 & 3.4)**.

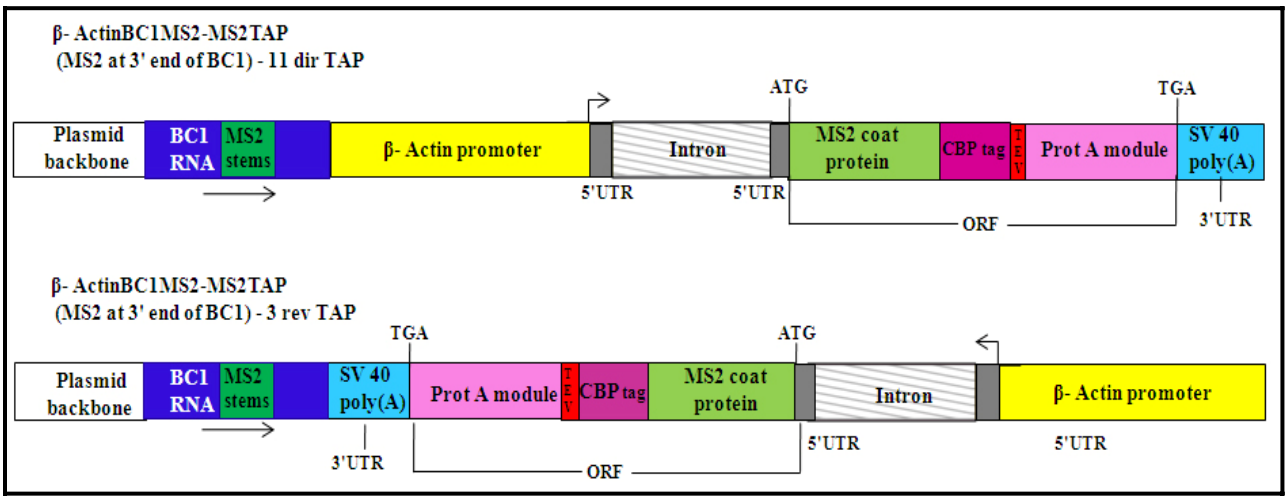


Fig 3.2 Schematic representation of BC1MS2+MS2TAP (not drawn to scale). The MS2TAP fusion polypeptide features the MS2 coat protein [green] fused to the TAP module with the following components: CBP tag [purple] and protein A tag [magenta] separated by a TEV protease cleavage site [red]. The MS2TAP fusion protein is driven by the chicken β-actin promoter [yellow] including its 5' UTR [gray] which, in turn, is interrupted by a ~1 kb intron [striped area] followed by a polyadenylation signal from a fragment from SV40 virus containing poly(A) additional signal [turquoise]. The BC1MS2 RNA gene [dark blue] has two MS2 stems [dark green] at the very 3' end of the BC1 RNA gene [blue - arrow showed the orientation of transcription] is driven by its native promoter with flanking regions [blue]. The MS2TAP fragment cloning resulted in two different orientations in the BC1MS2 plasmid backbone. One is in the same transcriptional orientation as BC1MS2, the other is in the opposite orientation to BC1MS2, and the plasmids were labeled as 11 dirTAP and 3 revTAP respectively. These fragments carrying the RNA and protein coding genes were (released with *Sfi*I/*Asc*I restriction endonucleases) used to generate transgenic mice.

As a first test in order to check expression of the generated constructs, we introduced the construct into an *ex vivo* system before generating transgenic mice (*in vivo*). We performed transient transfections in the mouse embryonic fibroblast cell line [NIH 3T3 cell line] for both the plasmids [3 revTAP and 11 dirTAP]. After transfections, the whole cell lysates were used for RNA and protein extraction. We observed that both plasmids [3 revTAP and 11 dirTAP] showed roughly equal quantity of expression of BC1MS2 RNA and MS2TAP protein. Northern blot **(Fig 3.3)** depicting BC1MS2 RNA expression of the expected size of ~220 nt in transient transfection in NIH 3T3 cells. As a control *in vitro* transcribed BC1 RNA was loaded and the expected size of BC1 RNA of 152 nt was observed in the blot.

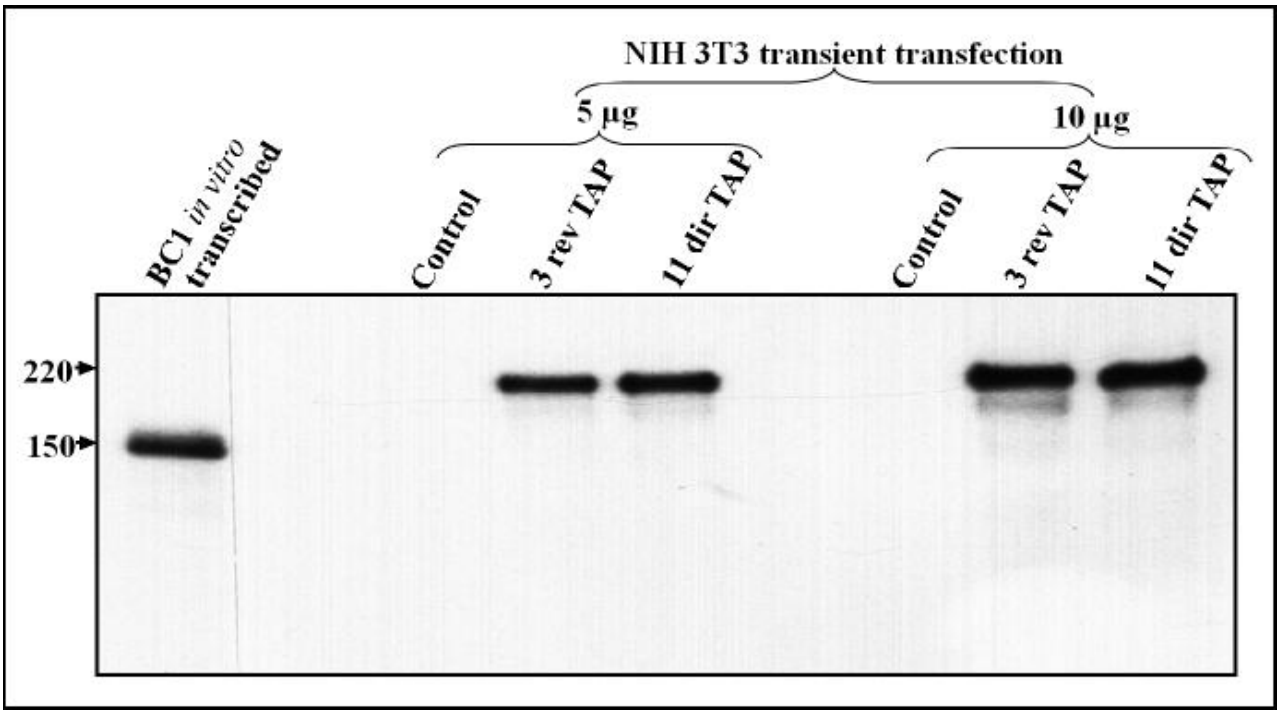


Fig 3.3 Northern blot showing the expression of BC1MS2 RNA in transient transfections from two plasmids in NIH 3T3 cells. Plasmids 3 revTAP and 11 dirTAP were transiently transfected in NIH 3T3 cells and total RNA was extracted and separated on a 6% polyacrylamide/7 M urea gel, transferred onto a nylon membrane and hybridized with a 27 nt long DNA probe complementary to the BC1 RNA unique region (Table.1). The tagged RNA exhibited the expected size of 220 nt. The first lane is *in vitro* transcribed BC1 RNA. Lane 3 & 6 were transfected with 3 revTAP [5 µg and 10 µg of total RNA loaded, respectively] and lane 4 & 7 were transfected with 11 dirTAP [5 µg and 10 µg of total RNA loaded, respectively]. The filter was reprobed for SRP RNA as a control for sample loading (not shown). Total RNA from untransfected NIH 3T3 cells (lanes 2 & 5) was used as a negative control. Exposure time for the blot was about 3 hours.

We checked the protein expression in transient transfections for both plasmids [3 revTAP and 11 dirTAP] in mouse embryonic fibroblast cell line [NIH 3T3 cells]. As the MS2TAP fusion protein exhibits an HA epitope at the N-terminus of the fusion protein, we used HA antibody to detect the fusion protein

expression. Immunoblot identifies a protein of about 39.3 kDa, as expected from both constructs used in transfection **(Fig 3.4)**. We also checked the protein expression using MS2 antibody raised in our lab and the experiment showed the same expected size [data has not shown].

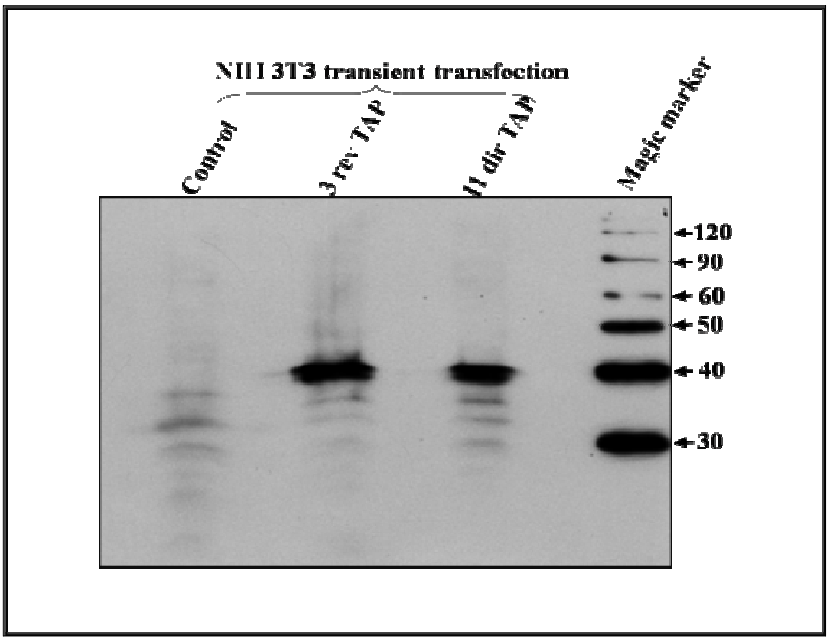


Fig 3.4 Immunoblot showing the expression of MS2TAP fusion protein in transient transfections of NIH 3T3 cells. Plasmids 3 revTAP and 11 dirTAP were transiently transfected in NIH 3T3 cells and the total proteins were extracted and separated on 12% polyacrylamide gel electrophoresis containing sodium dodecyl sulfate (SDS-PAGE), transferred onto a polyvinylidene difluoride (PVDF) membrane and probed with anti-HA antibody, the epitope which is located at the N-terminus of the fusion protein. The signal (band) in 3 revTAP and 11 dirTAP indicates that the MS2TAP fusion protein is expressed. Total protein from untransfected NIH 3T3 cells were used as a negative control. The lane "Magic marker" shows the molecular weight standards with fragment sizes of 120, 90, 60, 50, 40 and 30 kDa.

3.5 Expression of MS2TAP fusion protein but not BC1MS2 RNA in transgenic mice from combined construct

Transient transfections, indicated that both plasmids exhibited similar expression levels for the desired RNA and protein in the *ex viivo* system. Therefore, we had generated transgenic mice for both the constructs (3 revTAP and 11 dirTAP) with the MS2 stems at the very 3' end followed by the RNA polymerase III terminator for localization and purification *in vivo*. Total brain RNA from two transgenic mouse founders (3 revTAP mice # 21, #45; and 11 dirTAP mice # 14, # 19; see Appendix **Fig 8 A & B**) each was analyzed. Unexpectedly, we have not observed BC1MS2 RNA expression in any of the transgenic mouse founders examined **(Fig 3.5)**.

We have used rat BC1 LNA probe (a nucleotide difference between rat and mouse [C→U and A→G] near the ID region) to distinguish between the endogenous BC1 RNA expression and the transgenic BC1 RNA expression. We could observe a very small amount of cross hybridization with mouse endogenous BC1 RNA but not the larger transgenic BC1MS2 RNA. The reason for failure to detect BC1MS2 RNA might be that the MS2 stems at the very 3' end of the BC1 RNA transcript is destabilizing the RNA or impedes its transcription. It has been shown, for efficient transcription-termination and reinitiating of transcription by RNA polymerase III that the 3'-flanking sequence including the terminator mainly consisting of a stretch of U-residues is very important (Dumpelmann et al., 2003).

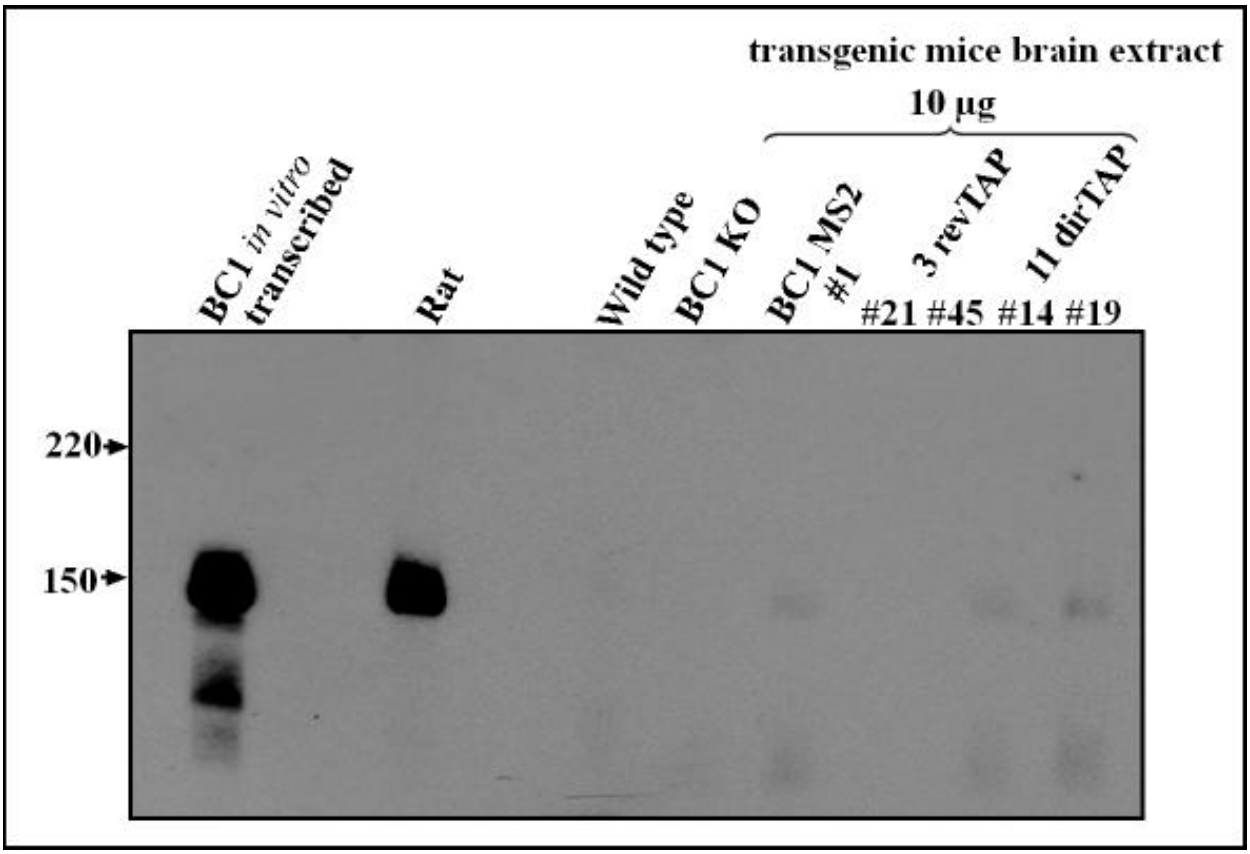


Fig 3.5 Northern blot revealing that BC1MS2 RNA was not expressed (MS2 3' end) in transgenic mouse brain. Total RNA extracted from mouse brain from 3 revTAP and 11 dirTAP transgenic mice was extracted and separated on a. 6% polyacrylamide/7 M urea gel, transferred onto a nylon membrane and hybridized with rat LNA oligonucleotide probe (Table.1) complementary to the rat BC1 RNA where it differs in two positions from the endogenous mouse BC1 RNA. Rat LNA oligonucleotide probe, distinguishes between the transgenic BC1 RNA from the endogenous mouse BC1 RNA. The first lane is *in vitro* transcribed BC1 RNA. As positive and negative controls, rat brain and wild type mouse total RNA was loaded, respectively. Total RNA from the transgenic mouse brain of two independent founders (from each construct) revealed that the BC1MS2 RNA was not or barely detected from both transgenic mice 3 revTAP and 11 dirTAP. The filter was reprobed for SRP RNA as a control for sample loading (not shown). Exposure time for the blot was about over night (12 - 14 hours).

In order to check the MS2TAP fusion protein expression in transgenic mice, we performed immunoblotting. As expected, we detected the fusion protein in 3 revTAP (data not shown) and 11 dirTAP transgenic mouse brain extracts (11 dirTAP mouse #21, see Appendix **Fig 8 B**). **Figure 3.6** depicts the MS2TAP fusion protein expression of the expected size using MS2 antibody in both transgenic mouse brain extracts.

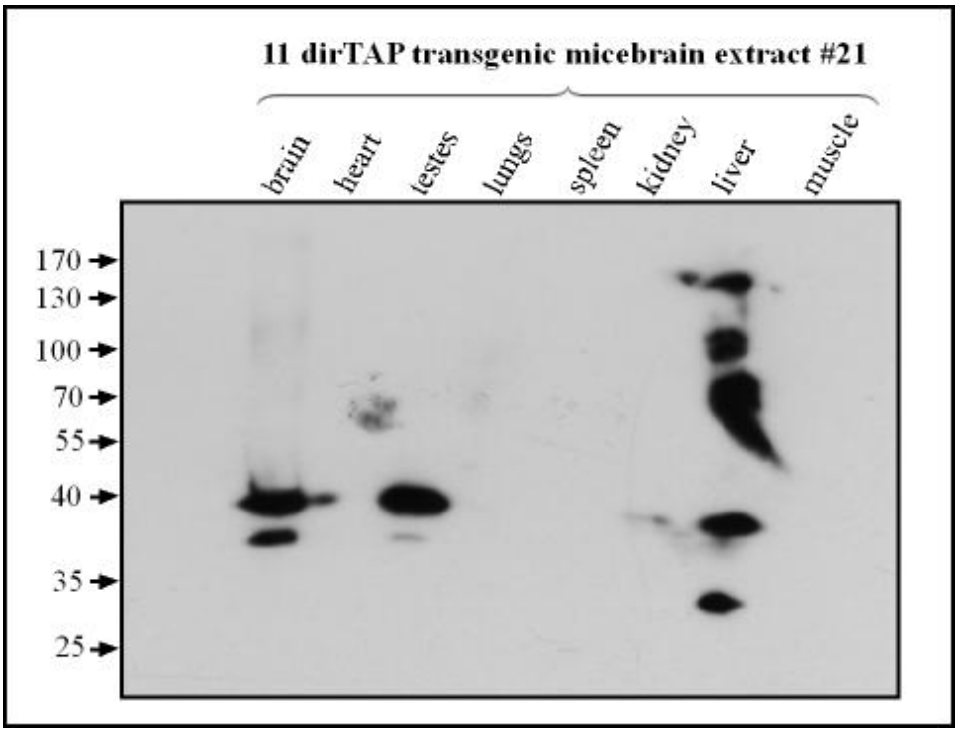


Fig 3.6 Immunoblot depiciting the expression of MS2TAP fusion protein in transgenic mouse tissue extracts. 11 dirTAP transgenic mouse tissue extracts were separated on 12% polyacrylamide gel electrophoresis containing sodium dodecyl sulfate (SDS-PAGE), transferred onto a polyvinylidene difluoride (PVDF) membrane and probed with anti-MS2 antibody. The signal in brain, testes and liver indicates MS2TAP fusion protein expression. The MS2TAP fusion protein showed the expected size of about 39.3 kDa. The molecular weight standard with fragment sizes of 120, 90, 60, 50, 40 and 30 kDa is indicated at the left margin.

3.6 Expression of BC200MS2 RNA and MS2TAP fusion protein from BC200MS2+MS2TAP transgenic mice

Parallel to BC1MS2 (MS2 at the very 3' end of BC1 RNA) tansgenic mice, we had generated BC200MS2 transgenic mice and investigated the RNA's expression. In contrast to the untagged BC200 RNA transgene (Khanam et al., 2007b), we observed a less pronounced expression of BC200MS2 RNA in transgenic mouse brain. In order to affinity purify the BC200 RNP from the MS2 tagged BC200MS2, we had crossed the BC200MS2 RNA expressing transgenic mice with MS2TAP protein expressing mice. From

BC200MS2+MS2TAP transgenic mice, we detected both MS2 tagged BC200 RNA and MS2TAP fusion protein expression. Interestingly, in addition to the 270 nt long MS2 tagged BC200 RNA, we observed a weaker band in the size of native BC200 RNA. Since normally mice are devoid of BC200 RNA and its gene, the smaller band is likely to correspond to an RNA that lost, presumably by processing, the MS2 domain at the 3' end. Both forms were detected only by using the BC200 unique DNA probe (Table 1) (**Fig 3.7,** mice # 11, 13 and 21, see Appendix **Fig 8 D)**. When we hybridized the membrane with the MS2 DNA probe (Table 1), we only observed the larger band (data not shown). From this experiment, we concluded that the BC200MS2 gene, when introduced into the *in vivo* system (transgenic mouse), the transcript showed a weaker expression and also is partially being processed.

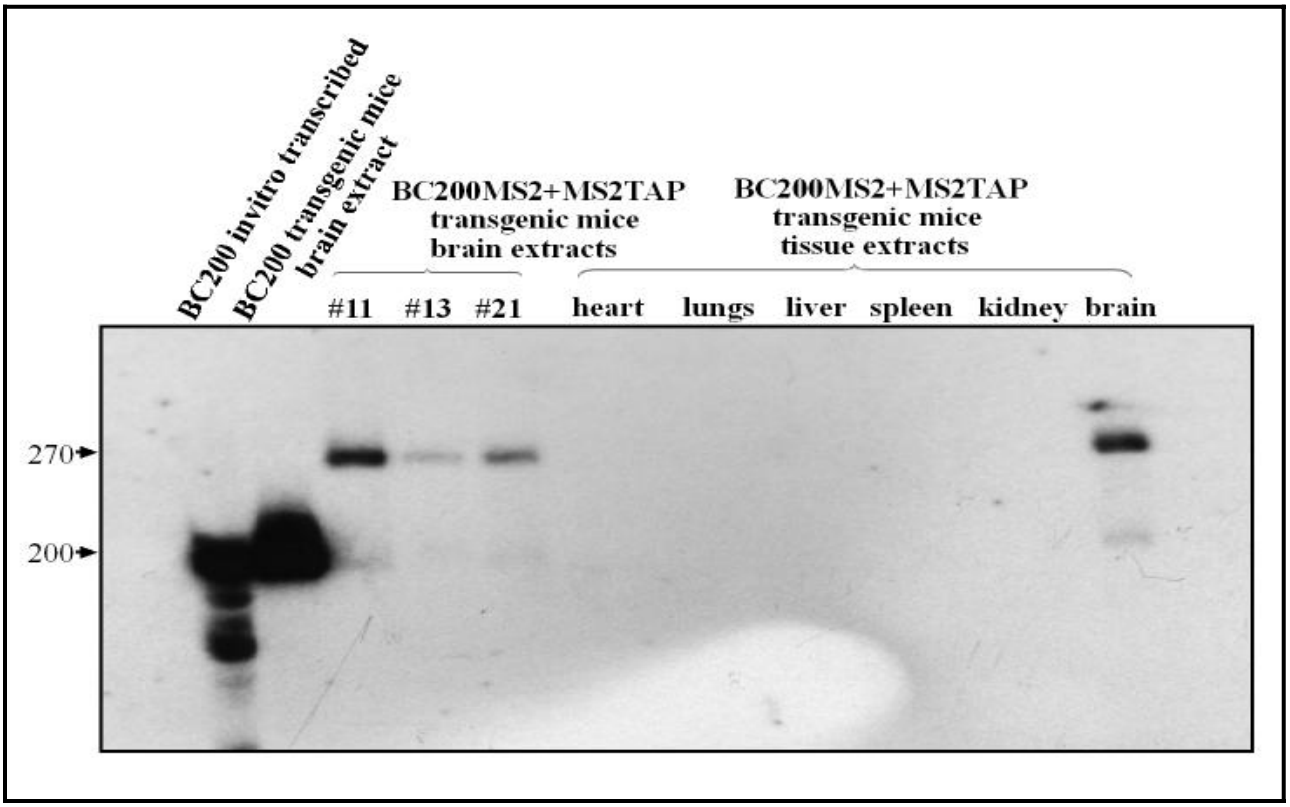


Fig 3.7 Northern blot showing the expression of BC200MS2 RNA in BC200MS2+MS2TAP transgenic mice. Total RNA was extracted from the brain of three independent transgenic mouse founders expressing BC200MS2+MS2TAP. The RNAs were separated on a 6% polyacrylamide/7 M urea gel, transferred onto a nylon membrane and probed against the BC200 unique DNA oligonucleotide (Table 1) and MS2 DNA probe (data not shown). BC200MS2+MS2TAP transgenic mice showed weaker signal of the expected size (270 nt) of BC200MS2 RNA compared to the 200 nt long untagged BC200 RNA. In BC200MS2 transgenic mouse brain total RNA, we observed two bands one at 270 nt long and a weak one of 200 nt length. About 10-20% of the BC200MS2 transcript was being processed. BC200MS2 transcript processing was observed only, when the membrane was hybridized with the BC200 unique probe. The filter was reprobed for SRP RNA as a control for sample loading (not shown). Exposure time for the blot was over night (about 12 hours).

3.7 Expression of BC1MS2 RNA in stable transfection

To further confirm or to understand the fate that the BC1MS2 RNA not expressed in BC1MS2 transgenic mice (MS2 at the very 3' end of the BC1 RNA), we had generated a plasmid containing BC1MS2 with Pgk neo gene to facilitate stable transfection in NIH 3T3 cells. In the stable transfection, we detected a weak expression of BC1MS2 RNA and that the transcript was partially processed **(Fig 1.15.)**. Processing is not certain, because NIH 3T3 cells have endogenous BC1 RNA expression (McKinnon et al., 1987). From these findings with BC1MS2 (MS2 at the very 3' end) in transgenic mice and stable transfection of BC1MS2, we decided on a different strategy, namely to place the MS2 tag internally with respect to BC1 RNA.

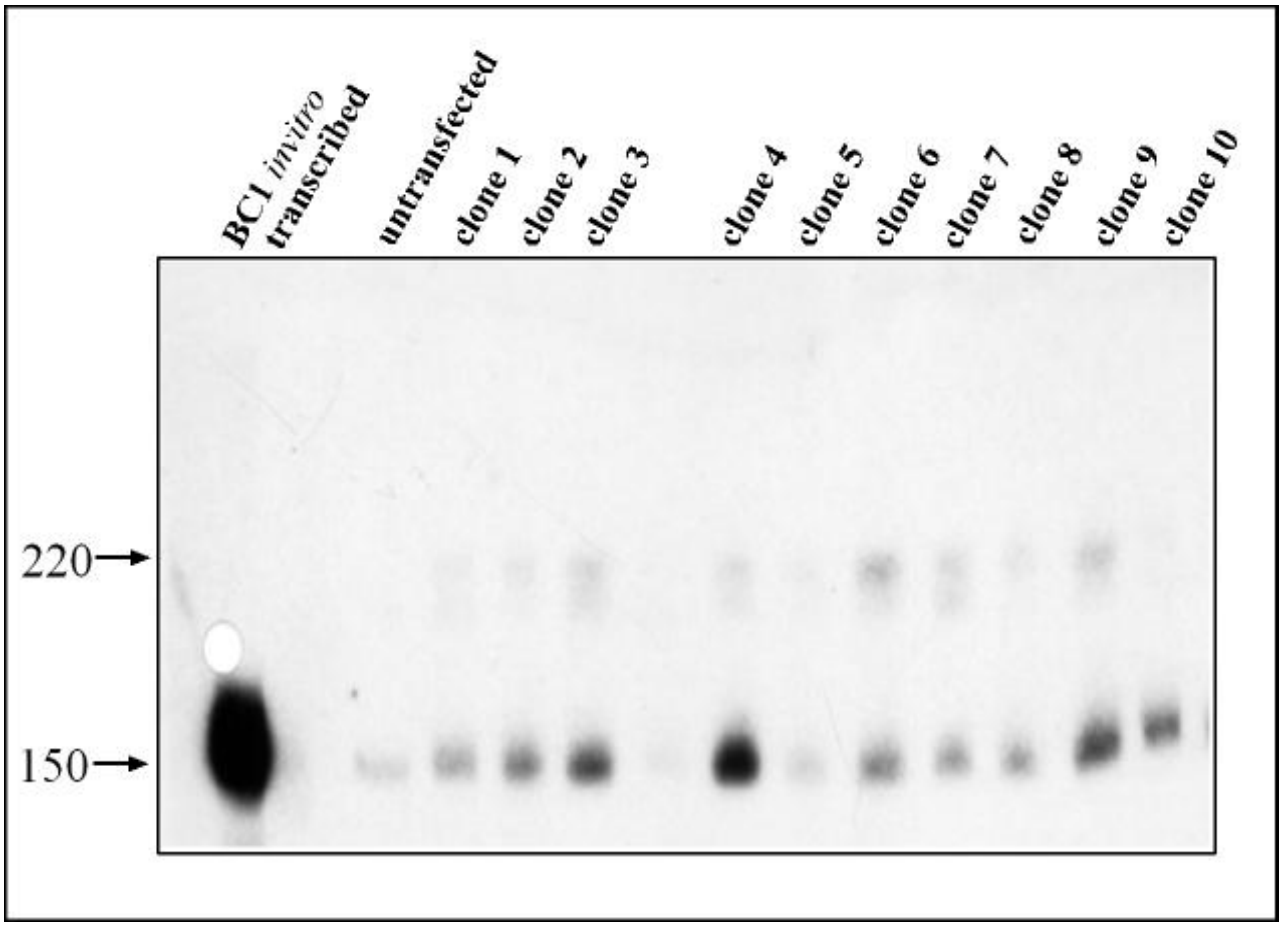
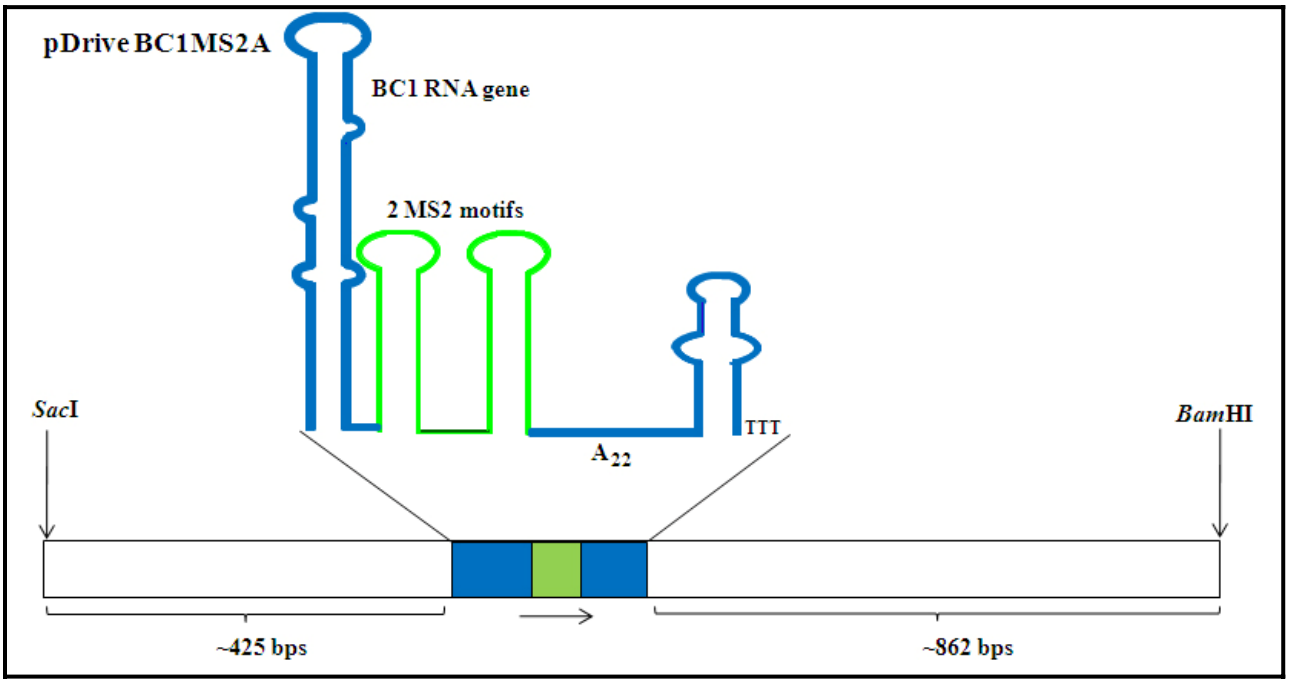


Fig 3.8 Northern blot showing the expression of BC1MS2 RNA during stable transfection. Total RNA was extracted from 10 clones of NIH 3T3 cells stably transfected pMKS-BC1MS2. Plasmid pMKS-BC1MS2 harbors Pgk neo cassette. Total RNA was extracted and separated on a 6% polyacrylamide/7 M urea gel, transferred onto a nylon membrane and hybridized with probe complementary to the BC1 region (rat LNA BC1 probe near to ID region, Table 1). The expected size of 220 nt was detected as a weak band. The first lane is *in vitro* transcribed BC1 RNA. Lane 2 was untransfected NIH 3T3 cell total RNA as a control. Exposure time for the blot was over night (about 12 hours).

3.8 MS2 tag placed internally of BC1 RNA to achieve BC1MS2 RNA expression.[MS2 stems proximal to the ID domain (MS2A) and proximal to 3' stem (MS2B) of BC1 RNA]

In order to achieve intact BC1MS2 RNA expression in the transgenic mice, as the MS2 stems at the very 3' end of the BC1 RNA likely was posttranscriptionally removed in most molecules in the *in vivo* mouse system, we moved the MS2 stems internal to the BC1 RNA gene i.e. immediately 3' to the ID (5' stem) or distal to the A_{22} stretch. **Figure (3.9)** represents the schemes of the new MS2 tagged BC1 RNAs. Both fragments were PCR amplified (described in Materials & Methods – section 1.4.) and cloned into pDrive vector (Qiagen, Hilden) and designated as BC1MS2A and BC1MS2B. From both constructs, we generated transgenic mice.



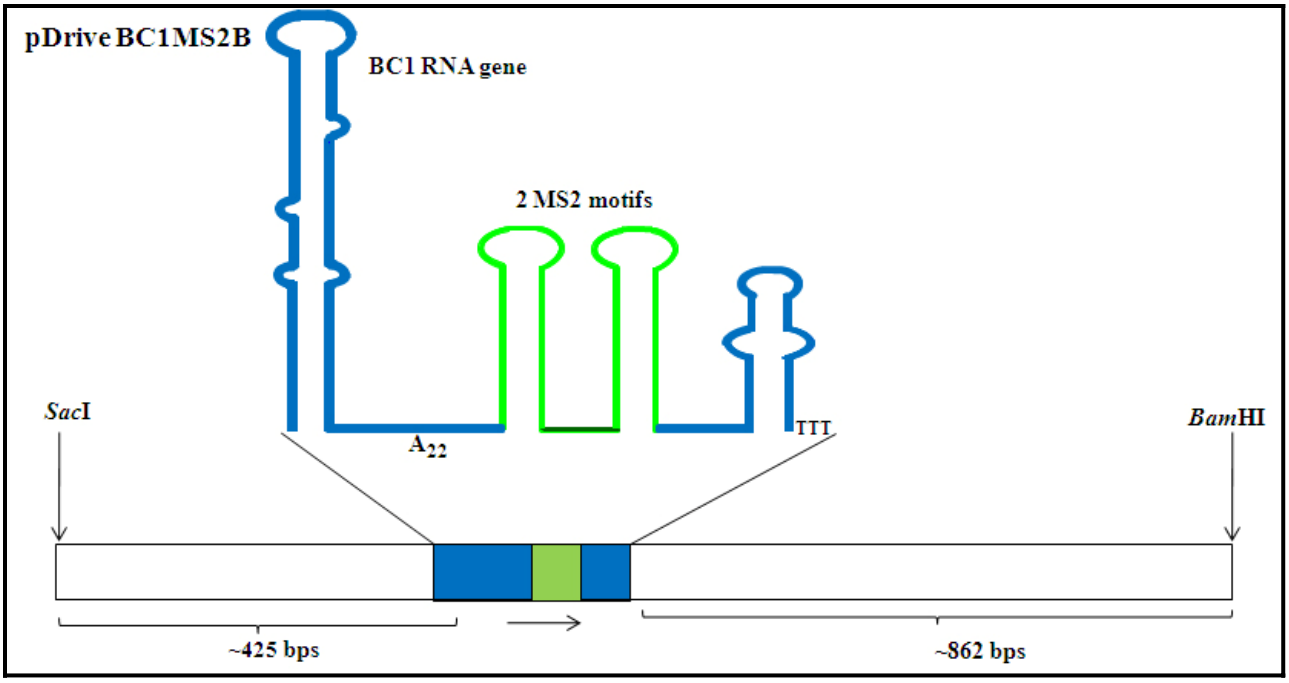


Fig 3.9 Schematic representation of BC1MS2A and BC1MS2B RNAs. The two MS2 stems motifs (MS2 protein binding sites) were placed internally of the rat BC1 RNA under their native promoter. In BC1MS2A, MS2 stems were placed distal to the 5' stem of BC1 RNA. For BC1MS2B the MS2 stems were located proximal to the 3' stem of BC1 RNA.

We have checked the RNA folding of BC1 RNA with MS2 tags placed internally, by using mfold from M. Zuker. http://www.bioinfo.rpi.edu/applications/mfold. The application showed no apparent interference with folding of the predicted BC1RNA structures with MS2 stem motifs. From the mfold application, we are predicting the MS2 protein binding sites might be accessible in binding the MS2 coat protein coexpressed as part of the BC1MS2+MS2TAP complex. We were assuming that the MS2 tag with BC1 RNA would express in the transgenic mice to allow for purification and dendritic transport studies in the transgenic mouse models (e.g., by employing antibodies in electron microscopy studies).

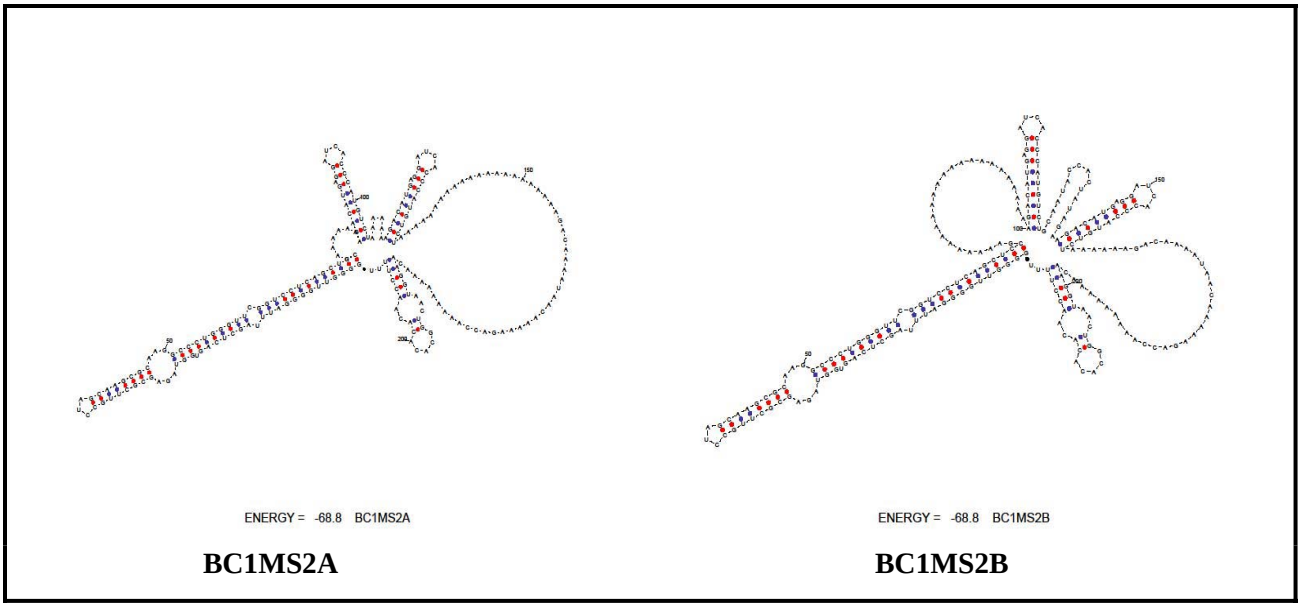


Fig 3.10 Secondary structure of BC1 RNA, two MS2 stems placed internal to the BC1 RNA gene in mfold. Folding of BC1MS2 using the Michael Zuker program http://www.bioinfo.rpi.edu/applications/mfold. Two MS2 RNA motifs separated by an unstructured spacer were placed internal to the rat BC1 RNA, close to the ID stem and the other closer to 3' stem, distal to the A_{22} stretch, designated BC1MS2A and BC1MS2B respectively.

Transgenic mice were generated using the two constructs BC1MS2A and BC1MS2B.The total RNA was analyzed by Northern blot from two independent mouse founders of BC1MS2A transgene and three independent mouse founders from BC1MS2B transgene **(Fig 3.11)**. We found little or no expression of BC1MS2A RNA in case of BC1MS2A transgenic mice. We found BC1MS2B RNA expression in BC1MS2B transgenic mice. The expression of this BC1MS2B RNA was weak in comparison with the expression of the wild type BC1 RNA. Among the offspring of three founder mice, one mouse **(**mouse #16, **Fig 3.11)** showed relatively high levels of BC1MS2B RNA expression. Therefore we expanded that transgenic mouse line.

3.9. Expression of BC1MS2B RNA from BC1MS2B transgenic mice crossed with MS2TAP fusion protein expressing mice for purification of BC1 RNP

Since our goal is to purify the BC1 RNP complex using MS2/MS2 coat protein complex, we had crossed BC1MS2B RNA expressing mice with MS2TAP protein expressing mice.We found that BC1MS2B RNA expression was further improved (mouse #18, **Fig 3.11**; see Appendix **Fig 9 D**). This difference between BC1MS2B #16 mouse and BC1MS2B+MS2TAP #18 mice may be due to the possibility that protein binding might stabilize the altered BC1 RNA.

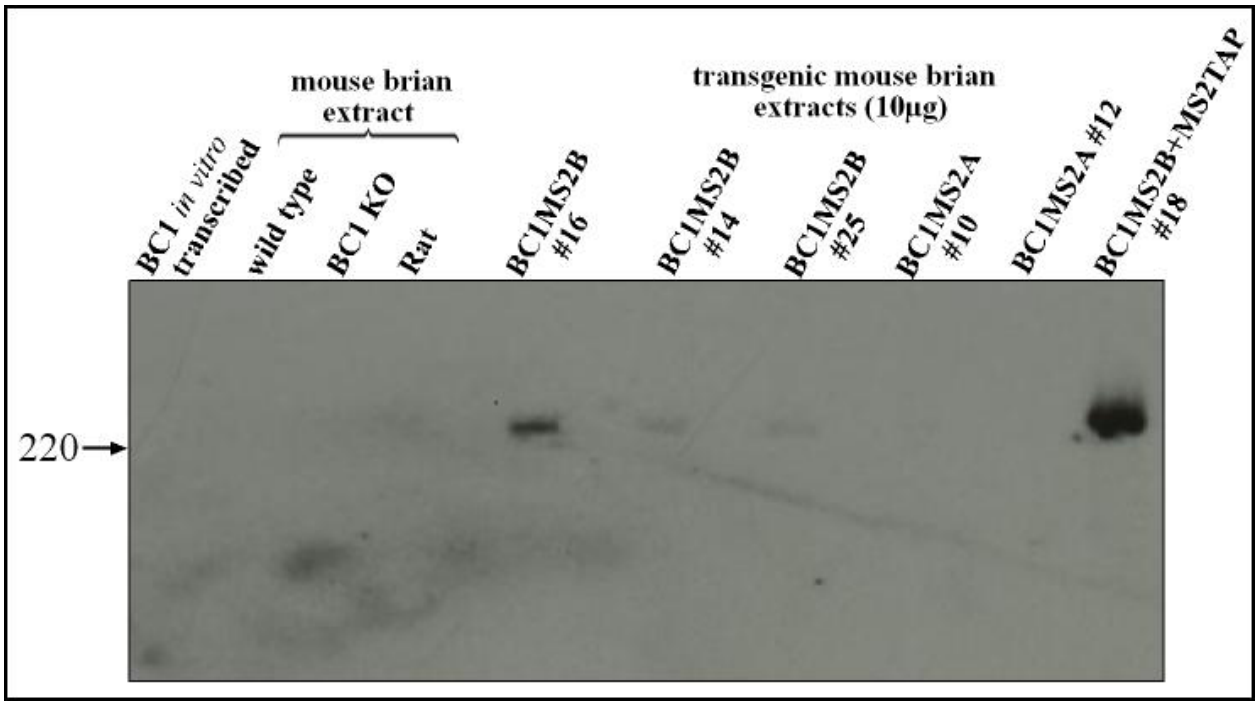


Fig 3.11 Northern blot showing the expression of BC1MS2A RNA and BC1MS2B RNA in BC1MS2A and BC1MS2B transgenic mice. Total RNA was extracted from the offsprings of three independent founder mice, expressing BC1MS2B RNA in BC1MS2B transgenic mice and two independent founders shows no expression of BC1MS2A RNA in BC1MS2A transgenic mice. The RNAs were separated on a 6% polyacrylamide/7 M urea gel, transferred onto a nylon membrane and probed against MS2 DNA oligonucleotide (Table 1). BC1MS2B transgenic mice showed only weak expression of expected size [~220 nt] (mouse #16, #14 & #25). In case of BC1MS2A transgenic mouse brain, we have not detected BC1MS2A RNA expression (mouse #10 & # 12) in any of the offsprings from two founder transgenic mice. Interestingly, BC1MS2B transgenic mice crossed with MS2TAP transgenic mice revealing increasing expression of BC1MS2B RNA (mouse #18). Prior to transfer, the gel was stained with ethidium bromide as a control for sample loading (not shown). Exposure time for the blot was about 3 days.

3.10 Crossing the BC1MS2B+MS2TAP in endogenous BC1 RNA$^{-/-}$ depleted background

When crossing the three independent BC1MS2B transgenic mice with MS2TAP transgenic mice, we found three different integration patterns of BC1MS2B fragment in the mouse genome by Southern blot analysis, using a probe specific to the 3' flanking region of BC1 RNA (~800 nt). Low copy number transgenic mice (mice #54, 67; **Fig 3.12**; see Appendix **Fig 9 D**) showed BC1MS2B RNA expression. Multi copy number transgenic mice, unable to detect BC1MS2B RNA expression.

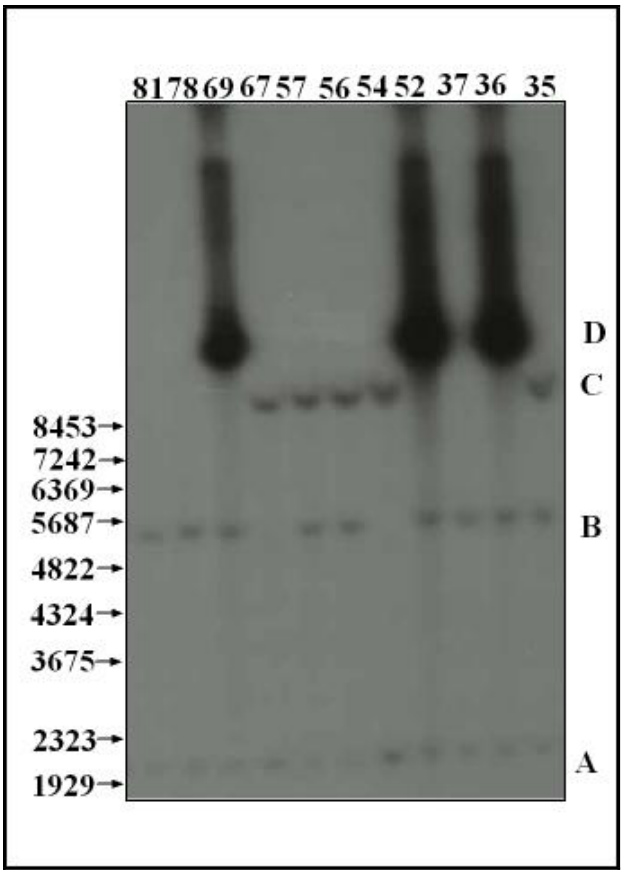


Fig 3.12 Southern blot patterns of BC1MS2B in mouse genome. Southern blot analysis of *Bam*HI-digested mouse genomic DNAs reveals three different types of integration patterns, when crossed into BC1 RNA$^{-/-}$ mice and probed with the 3' flanking region of BC1 RNA (~800 nt). Band B (around ~5 kb), is the endogenous wild type mouse DNA fragment (predicted size: 5,048 nt) containing the BC1 RNA gene and absent in some mice (# 54, 67) due to the BC1 RNA$^{-/-}$. Mice #54 and 67 show low copy number transgene integration at the larger size range. Mice #35, 36, 52, 56, 57 and 69 show multi copy number transgene integration at the larger size range.Exposure time for the blot was about 3 days.

3.11 Establishing a line of BC1MS2B RNA and MS2TAP protein expressing transgenic mice

In order to eleminate competition of native BC1 RNA expression in BC1MS2B RNA expressing transgenic mice, crossed BC1MS2B transgenic mouse (mouse # 16, **Fig 3.11**, see Appendix **Fig 9 C**) with BC1 RNA$^{-/-}$ mice. Parallel to the BC1MS2B transgenic mice, MS2TAP protein expressing transgenic mice were crossed into the BC1 RNA$^{-/-}$ background. The resulting transgenic mice BC1MS2B_BC1 RNA$^{-/-}$ and MS2TAP_BC1 RNA$^{-/-}$ were crossed to obtain a transgenic mouse line which harbor BC1MS2B RNA and the MS2TAP protein module. To achieve complete elemination of endogenous BC1 RNA in BC1MS2B+MS2TAP transgenic mice, we crossed BC1MS2B and MS2TAP transgenic mice to homozygosity (BC1 RNA$^{-/-}$) for about 3-4 generations. This was confirmed by performing Southern blot analysis with specific probes against BC1 (BC1 RNA 3' flanking region, data not shown) and MS2TAP (β-Actin MS2TAP, **Fig. 3.14**) genes. We also checked for maintenence of BC1MS2B RNA expression by Northern blot analysis (**Fig. 3.13)**.

Total RNA from homozygous BC1MS2B+MS2TAP_BC1 RNA$^{-/-}$ and heterozygous BC1MS2B+MS2TAP_BC1 RNA$^{+/-}$ transgenic mice was analyzed by Northern blot using BC1 unique DNA probe, revealing differences of BC1MS2B RNA expression in the different transgenic mice. In the

homozygous BC1MS2B+MS2TAP_BC1 RNA$^{-/-}$ (mice #5 &.#10), we detected only BC1MS2B RNA, whereas heterozygous BC1MS2B+MS2TAP_BC1 RNA$^{+/-}$ (mouse #18) showed BC1MS2B RNA along with endogenous BC1 RNA.

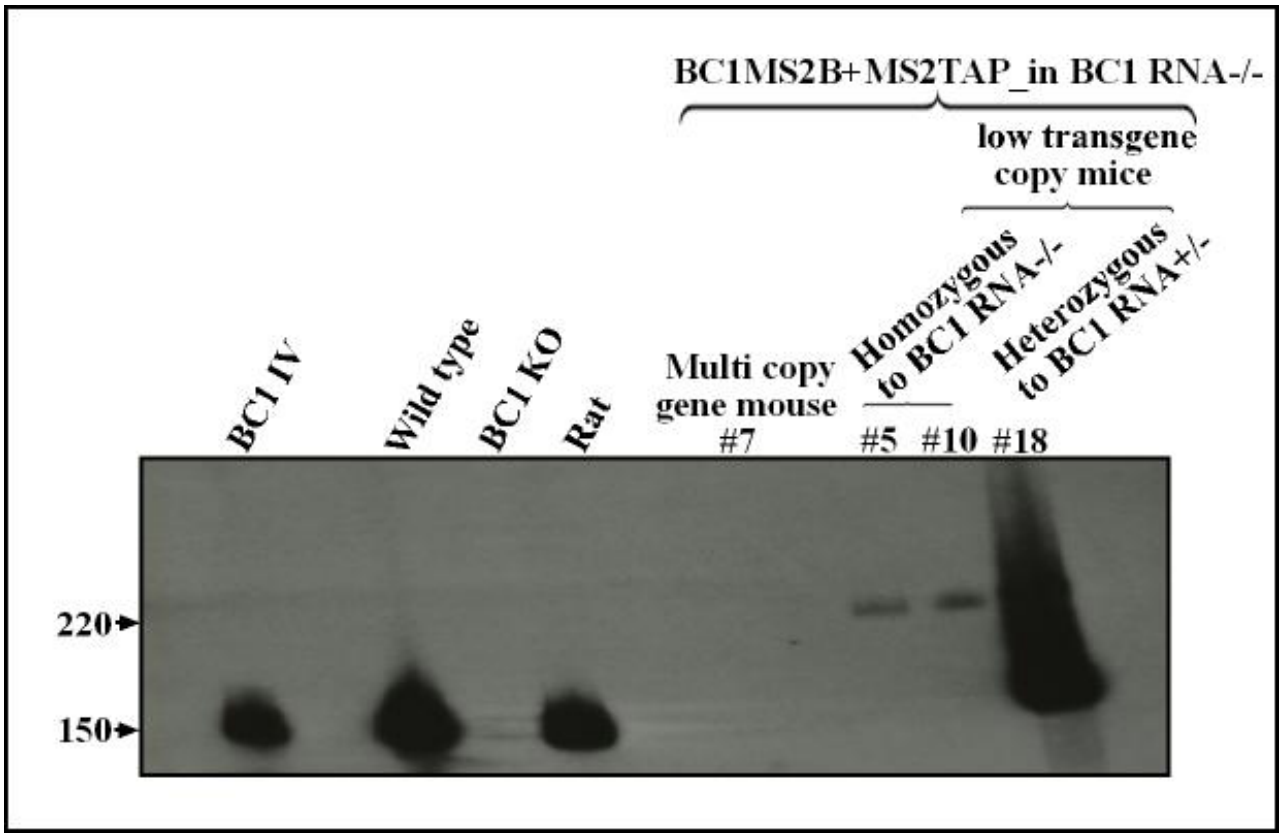


Fig 3.13 Northern blot shows the expression of BC1MS2B RNA in BC1MS2B transgenic mouse. Total RNA was extracted from two independent founders containing the gene encoding BC1MS2B RNA. RNAs were separated on a 6% polyacrylamide/7 M urea gel, transferred onto a nylon membrane and probed against BC1 unique oligonucleotide (Table 1). BC1MS2B transgenic mice showed weak expression of the expected size (~220 nt) (mouse #5, #10 & #18) of BC1MS2B RNA compared to the wild type mice. Multi-copy transgene integrated mice showed no expression of the BC1MS2B RNA (mouse #7) where as low transgene copy integrated mouse showed the expression of BC1MS2B RNA (mouse #5, #10 & #18). In order to avoid the native BC1 RNA competition, BC1MS2B transgenic mice were crossed with BC1KO (BC1 RNA$^{-/-}$) mice until the native BC1 RNA was absent from the transgenic mice. Northern blot also revealed the difference between the homozygous BC1 RNA$^{-/-}$ background vs BC1 RNA$^{+/-}$. Homozygous BC1 RNA$^{-/-}$ background transgenic mouse brain showed only BC1MS2B RNA expression (mouse #5 & 10), in heterozygous BC1 RNA$^{+/-}$ mice, endogenous BC1 RNA expression and transgenic BC1MS2 RNA expression was observed (mouse #18). Prior to transfer, the gel was stained with ethidium bromide as a control for sample loading (not shown). Exposure time for the blot was about 3 days.

Another approach in confirming the genotype of the presumptive homozygous to BC1MS2B $^{Tg/Tg}$ and MS2TAP$^{Tg/Tg}$ mice was based on test crossing (statistical genetics). In this approach of test crossing, confirmation was accomplished by mating between the presumptive homozygous mice with a non-

transgenic [wild type $^{(+/+)}$] partner. If the mouse in question is only a $^{Tg/+}$ heterozygote, one would expect equal numbers of $^{Tg/+}$ and $^{+/+}$ (wild type) offspring. If the transgenic mouse was homozygous, all the offspring would be heterozygous and could be easily distinguished by PCR analysis and were confirmed by Southern blot analysis.

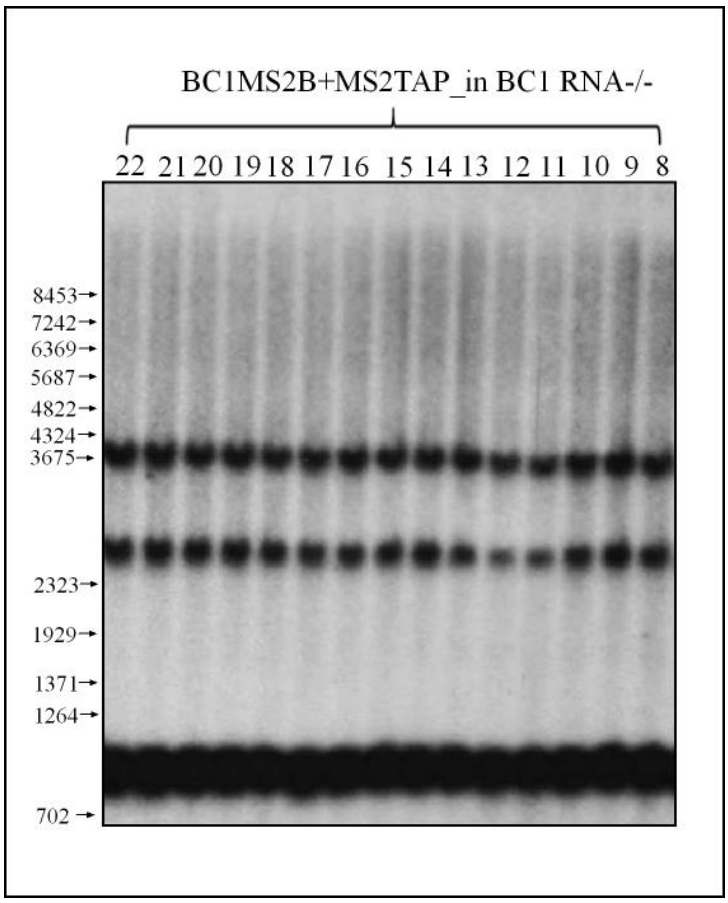


Fig 3.14 Southern blot demonstrating the integration of MS2TAP protein gene in BC1MS2B+MS2TAP_BC1 RNA$^{-/-}$ background. Southern blot analysis of *Bam*HI-digested mouse genomic DNAs revealed the MS2TAP fragment integration into the mouse genome. The probe used for identification of MS2TAP was 3' end of β-Actin MS2TAP gene (~1.2 kb). The upper band around ~3.6 kb correspond to the transgenic construct fragment to the 5' region and the 0.9 kb correspond to the 3' transgenic *Bam*HI- digested DNA fragment. Exposure time for the blot was overnight (about 12 - 14 hours).

3.12 Affinity purification of BC RNPs from BC1MS2B+MS2TAP in endogenous BC1 RNA depleted & BC200MS2+MS2TAP transgenic mice

Various biochemical methods have been used for purification and isolation of BC RNPs, most of them are not successful yet to purify or identify all the proteins bound to these RNP complexes. Therefore, we sought to develop a general method for purification of endogenously assembled BC RNPs *in vivo* and in transgenic mouse system. Therefore, we had tagged these RNAs with MS2 motifs and the corresponding MS2 coat protein was fused to TAP tag protein. This method represents itself as a RNA tag based affinity purification for purification of BC RNP complexes. Affinity purification has been used successfully to purify the affinity tagged proteins without disturbing the native structure. Therefore, we assumed affinity tags could serve the purpose of BC RNP purification.

To our knowledge, the purification of BC1 or BC200 RNPs using RNA tag based affinity purification is an advanced affinity purification strategy of endogenously assembled RNPs from the mouse brain extract. This generalized method of purification can be easily adapted to other non-protein coding RNAs (npcRNAs) to achieve high purification.

3.13 RNA/RNP integrity is the most important step in RNP affinity purification

We evaluated the MS2 tagged BC1MS2B+MS2TAP_in BC1 RNA$^{-/-}$ and BC200MS2+MS2TAP RNP complex formation in the mouse brain and compared with wild type and BC200 transgenic mouse brain extracts. We used DEAE ion-exchange chromatography, where BC1 and BC200 RNP complexes were partially purified using DEAE chromatography from wild type and BC200 transgenic mouse brain extracts. DEAE purification showed RNA/RNP integrity at 200 – 300 mM KCl concentration, where second round of 300 mM KCl concentration, the RNP complex was partially dissociated (Muddashetty et al., 2002). Therefore, S2 fractions of mouse brain extracts were used for binding to DEAE matrix for partial purification of BC MS2 RNP complexes. Same conditions were maintained for binding and washing. Complexes were eluted with 200 – 300 mM KCl concentration and checked for RNP integrity, by performing native gel eletrophoresis. **Fig 3.15** shows transgenic mouse brain extracts (BC1MS2B+MS2TAP_in BC1 RNA$^{-/-}$ and BC200MS2+MS2TAP) partially purified from DEAE elutes, showed the intact RNP migrates as of wild type and BC200 transgenic mouse brain DEAE elutes.

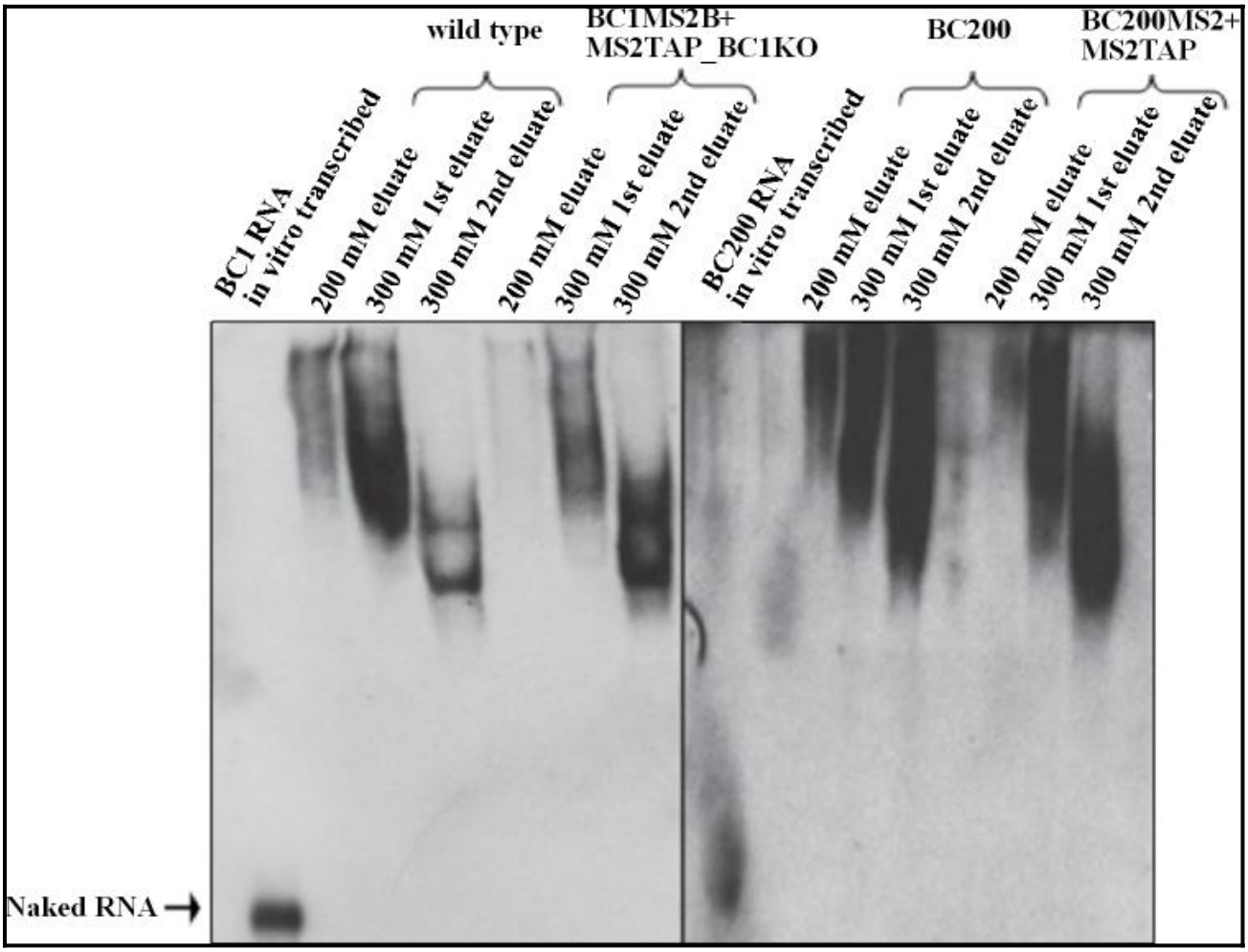


Fig 3.15 Autoradiograph of a native gel shows the RNP complex. RNP complex formation from brain extracts of wild type mice compared to BC1MS2B+MS2TAP transgene and BC200 transgene compared to BC200MS2+MS2TAP transgene where brain extracts were partially purified using DEAE ion–exchange resin and eluted with 200 mM and twice with 300 mM KCl concentrations. The RNP complexes were separated on a native gel, transferred on to a nylon membrane and probe with unique probes for BC1 or BC200 RNAs respectively (Table 1). The complex was stable at low concentrations of salt (200 mM). After the second elution with 300 mM KCl concentration, the BC RNP complex is partially dissociated.

Note: BC1MS2B+MS2TAP and BC200MS2+MS2TAP brain extracts were concentrated to attain a comparable signal with wild type and BC200 brain extracts. Exposure time for the blot was over night (about 12 hours).

3.14 Initial attempt in purification of BC RNPs using TAP method from BC1MS2B+MS2TAP_in BC1 RNA$^{-/-}$ and BC200MS2+MS2TAP transgenic mouse brain.

The RNP integrity is the most important step in purification of any RNP complex, Native gel electrophoresis showed intact RNP particles of BC MS2 RNP complexes. Therefore, we attempted to purify the BC MS2 RNP complexes using the TAP method. The TAP purification is a two-step affinity purification

developed by *Bertrand Seraphin* (Rigaut et al., 1999), it contains ProtA tag, TEV cleavage recognition site and CBP tag. The CBP tag allows for efficient selection and specific release from the affinity resin under mild conditions. In contrast, ProtA can be released from the matrix-bound IgG only under denaturing conditions at low pH, therefore, TEV protease recognition sequence was inserted, which has been specifically cleaving the ProtA from CBP tag under native conditions.

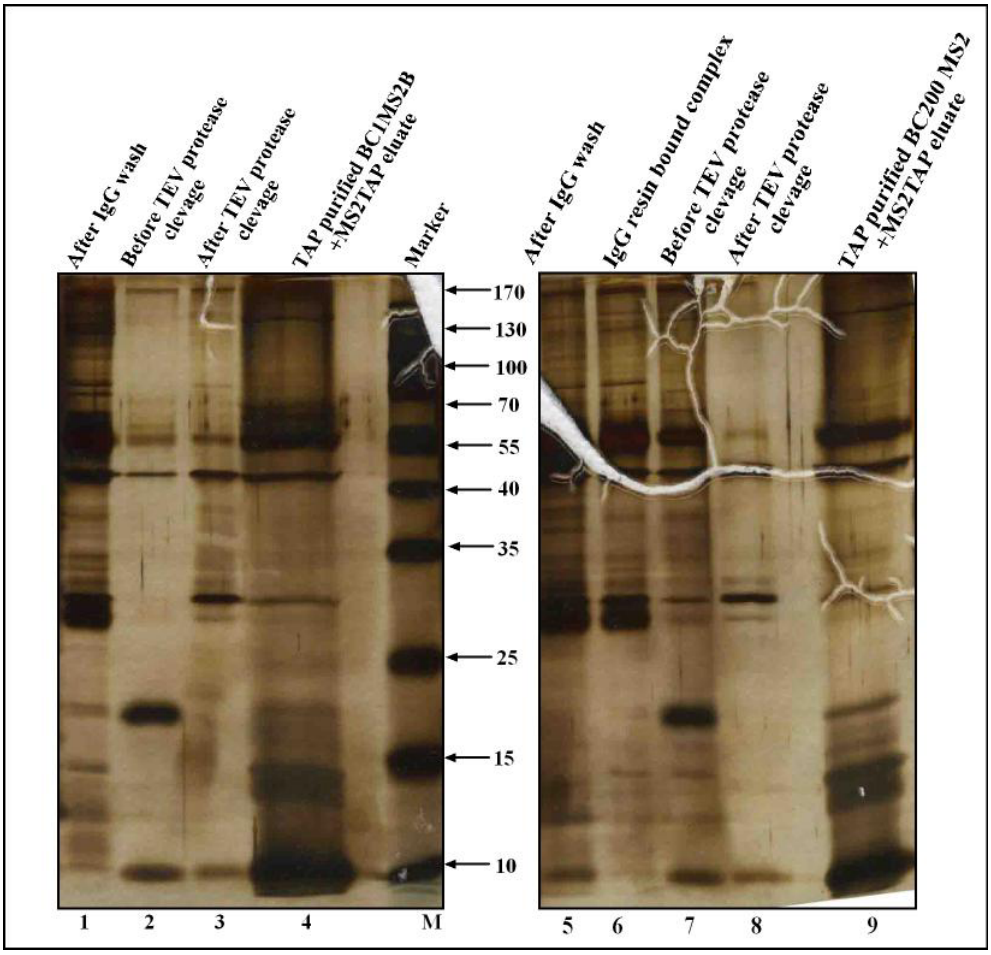


Fig 3.16 An RNA tag based affinity purification of BC1 and BC200 RNP complexes using TAP method. Mouse brain extracts (S100) were used for affinity purification. Proteins present in final TAP purification (lanes 4 and 9). First affinity purification of RNP complexes IgG bound resin after wash with IPP buffer. (lanes 1 and 5). Before TEV protease cleavage treatment (lanes 2 and 7). After TEV protease cleavage treatment (lane 3 and 8). Lane 6 is after TEV cleavage IgG bound resin.

Using the DEAE eluted fractions, (we exlcuded the second 300 mM KCl elution – as the RNP was partially dissociated) **(Fig 3.15)** of BC RNP complexes from the BC1MS2B+MS2TAP_ in BC1 RNA$^{-/-}$ and BC200MS2+MS2TAP brain extracts, we attempted to purify the complexes using TAP methodology. DEAE eluted fractions were first incubated with Protein-G sepharose for 2 - 3 hours at 4 °C, washed the Protein-G sepharose matrix with IPP buffer (15 ml) [25 mM Tris pH 8.0., 200 mM KCl, 0.1% NP-40, 1 mM DTT, 3 mM $MgCl_2$ and protease inhibitor]. Followed by TEV protease treatment for 2 hours at room temperature

(16-25 °C) and 4 °C over night. Followed by eluting the TEV cleaved protein from the matrix and bind it to the CAM sepharose for 2 - 3 hours. CAM matrix was washed with [25 mM Tris pH 8.0., 200 mM KCl, 1 mM DTT, 3 mM $MgCl_2$, 1 mM Mg acetate, 1 mM imidazole, 2-5 mM $CaCl_2$ and 10 mM ß-mercaptoethanol] and finally eluted the RNP complexes with [25 mM Tris pH 8.0., 200 mM KCl, 20 mM EGTA, 3 mM $MgCl_2$, 1 mM Mg acetate, 1 mM imidazole, 3 mM $CaCl_2$ and 10 mM ß-mercaptoethanol] and TCA precipitated the RNP complexes. . Fractions were run by 12% polyacrylamide gel electrophoresis containing sodium dodecyl sulfate (SDS-PAGE). Aspects of the MS2TAP system might be functional as a band in the right size range of full-length MS2TAP fusion protein (39. 3 kDa) is observed prior to TEV protease **(Fig 3.16)**. After TEV protease treatment, where TEV protease removes the IgG portion leaving the CBP tag with the MS2 domain. The expected molecular size of the remaining portion of the fusion protein after TEV treatment is about ~28 kDa, which was observed in our test run of purification. Since the expression levels of both RNAs in transgenic mice are relatively low, we need large amounts of mouse brain extracts to obtain sufficient amounts of the proteins bound to the BC RNP complexes for sequence analysis.

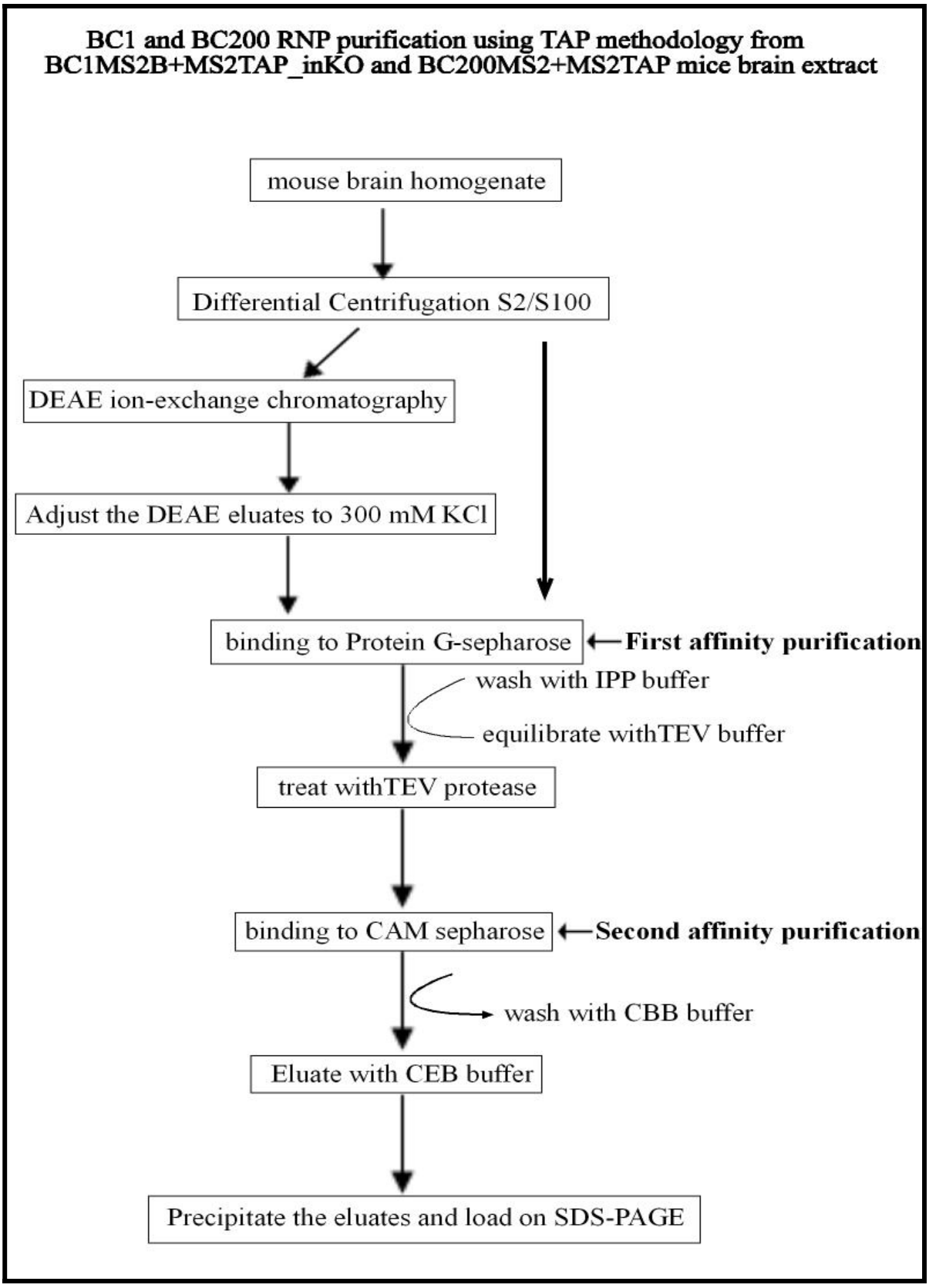


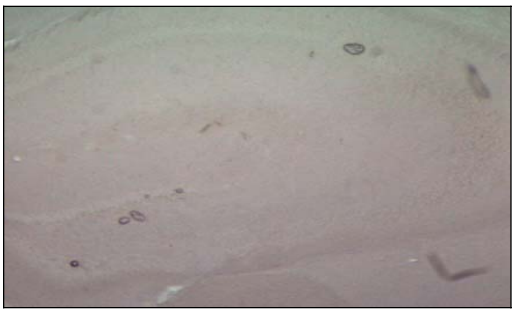
Fig 3.17 Schematic representation of Tandem Affinity Purification (TAP) of BC RNPs

3.15 Subcellular-localization of BC MS2 RNAs in BC1MS2B+MS2TAP in BC1 RNA$^{-/-}$ and BC200MS2+MS2TAP transgenic mice.

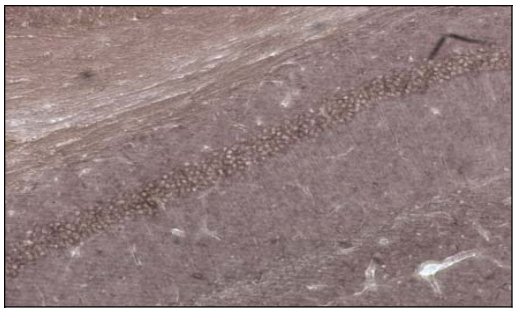
The rodent specific BC1 RNA and the anthropoid specific BC200 RNA are exclusively expressed in the nervous system. The distribution of both the RNAs showed a clear dendritic localization in the brain. We want to check that the MS2 tagged BC RNAs were transported to dendrites similar to native BC RNAs in mouse brain sections. *In situ* hybridization on coronal sections through the hippocampus of BC1MS2B+MS2TAP_in BC1 RNA$^{-/-}$ and BC200MS2+MS2TAP transgenic mice using sense (S) and antisense (AS) probes, were designed against the unique regions of BC1 RNA and BC200 RNA, revealed the tagged RNA localization was restricted to the cell bodies and were not detected in dendrites. The expression pattern in the transgenic mice [BC1MS2B+MS2TAP_in BC1 RNA$^{-/-}$ and BC200MS2+MS2TAP] revealing, the MS2 tagged BC RNAs were localized in the CA1, CA2 and CA3 pyramidal cells of the hippocampus. Both the BC RNAs were present in somatic cytoplasm of the dendrites. Dendritic localization of these RNAs was restricted to cytoplasm but not in the nucleus **(Fig 3.18 & 3.19)**. There could be two reasons for these observations: the expression level of the MS2 tagged BC RNAs is too low in transgenic mice to detect them in dendrites or tagging these RNAs abolishes transport of these RNAs into dendrites.

Though the RNA expression levels of these two transgenic mice (BC1MS2B+MS2TAP_in BC1 RNA$^{-/-}$ and BC200MS2+MS2TAP) were very weak compared to the wild type mice and BC200 transgenic mice, we still observed a similar expression pattern of the BC1MS2B and BC200MS2 RNA in the corresponding transgenic mice. Both transgenic mice (BC1MS2B+MS2TAP_in BC1 RNA$^{-/-}$ and BC200MS2+MS2TAP) on *in situ* hybridization using unique oligoribonucleotide probe designed against BC1 and BC200 RNA revealed that CA1, CA2 and CA3 showed similar localization pattern, compared to the BC1 and BC200 RNA. There is little or no expression in the granular cells of the dentate gyrus (DG) was observed in BC1MS2B+MS2TAP_BC1 RNA$^{-/-}$ and BC200MS2+MS2TAP, similar to the wild type mice and BC200 transgenic mice.

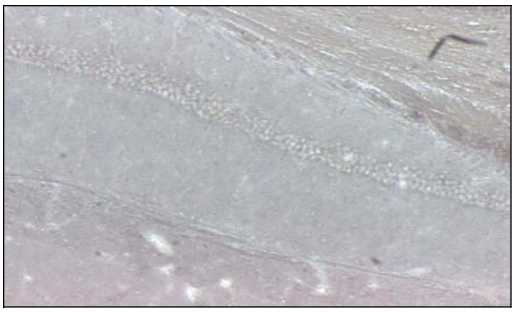
BC1MS2B_TAP_in BC1 RNA$^{-/-}$_AS_4X Hippo



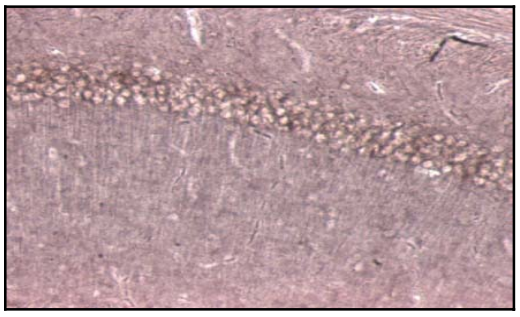
BC1MS2B_TAP_ in BC1 RNA$^{-/-}$_S_4X Hippo



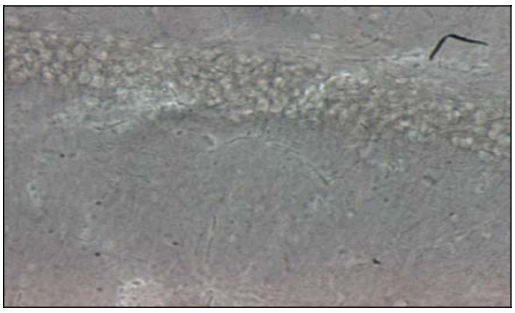
BC1MS2B_TAP_in BC1 RNA$^{-/-}$_AS_10X CA1



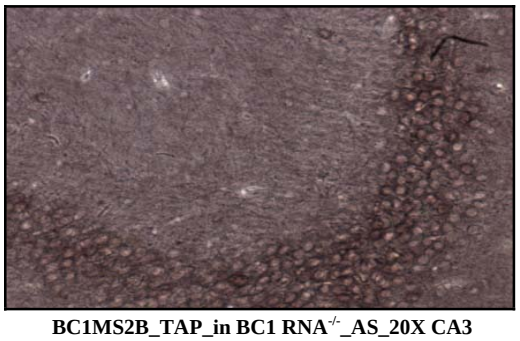
BC1MS2B_TAP_ in BC1 RNA$^{-/-}$_S_10X CA1



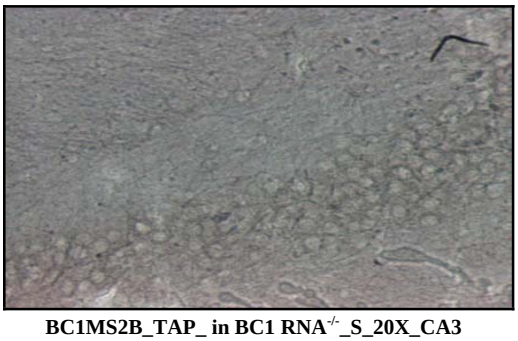
BC1MS2B_TAP_in BC1 RNA$^{-/-}$_AS_20X CA2



BC1MS2B_TAP_ in BC1 RNA$^{-/-}$_S_20X_CA2

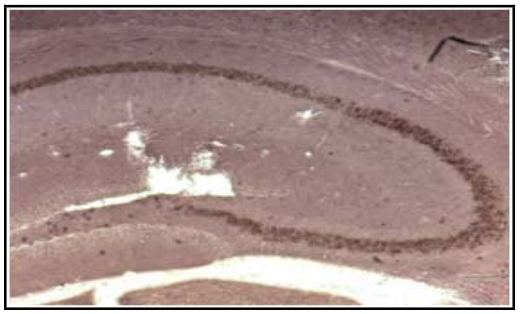


BC1MS2B_TAP_in BC1 RNA$^{-/-}$_AS_20X CA3



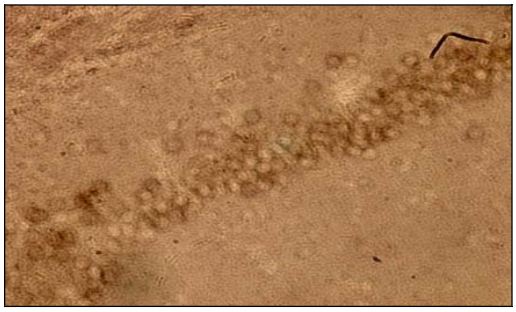
BC1MS2B_TAP_ in BC1 RNA$^{-/-}$_S_20X_CA3

Fig 3.18 BC1 RNA *in situ* hybridization in BC1MS2B+TAP_in endogenous BC1 RNA depleted transgenic (BC1 RNA$^{-/-}$) mice. Non-radioactive *in situ* hybridization with DIG labeled BC1 unique oligoribonucleotide probes on coronal sections of BC1MS2B+MS2TAP transgenic mouse brain. In the hippocampus, there is labeling in the cell bodies and dendrites of areas CA1, CA2 and CA3, while the dentate gyrus granule cells are not labeled. The higher magnification of CA1, CA2 and CA3 showed staining only in the proximal regions of dendrites not up to distal parts of the dendrites. The BC1 sense probe did not show any signal.

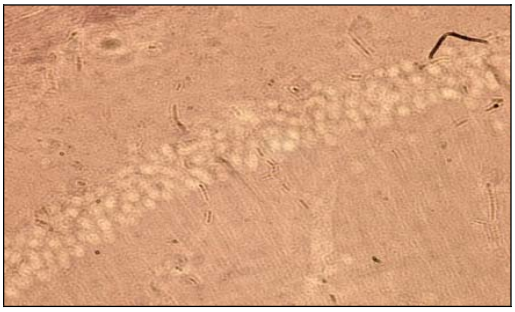


BC200TAP_AS_4X_Hippo

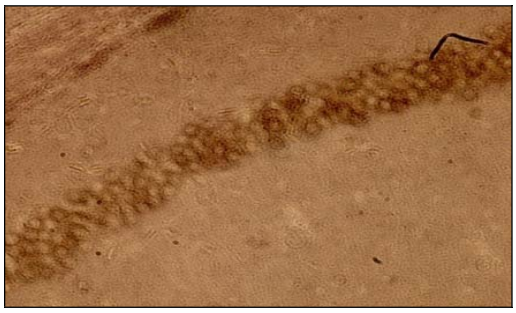
BC200TAP_S_4X_Hippo



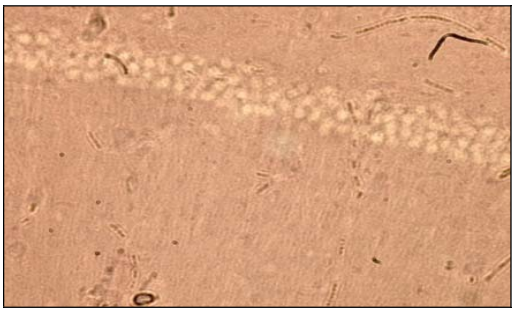
BC200TAP_AS_20X_CA1



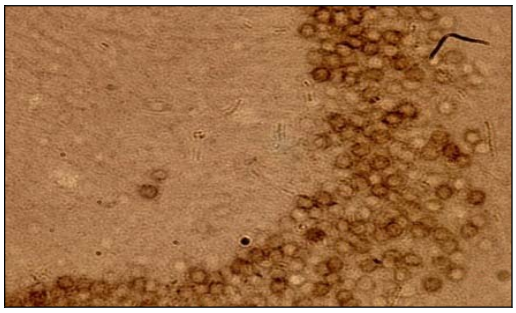
BC200TAP_S_20X_CA1



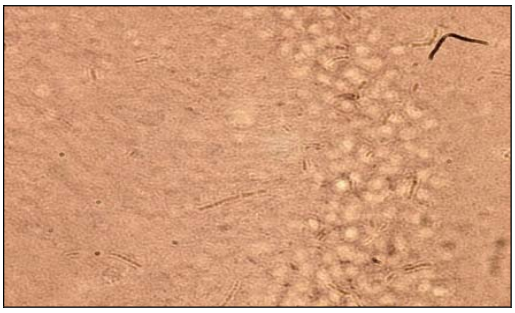
BC200TAP_AS_20X_CA2



BC200TAP_S_20X_CA2



BC200TAP_AS_20X_CA3



BC200TAP_S_20X_CA3

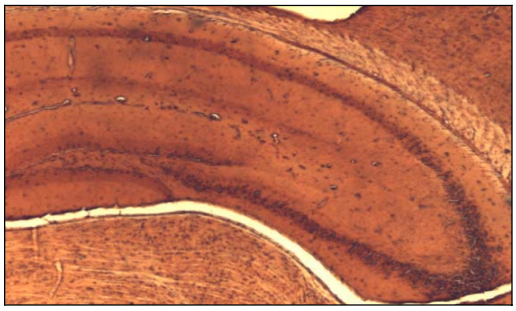
Fig 3.19 BC200 RNA *in situ* hybridization in BC200MS2+MS2TAP transgenic mice. Non-radioactive *in situ* hybridization with DIG labeled BC200 unique oligoribonucleotide probes on coronal sections of BC200MS2+MS2TAP transgenic mouse brain. In the hippocampus, there is intense labeling in the cell bodies and dendrites of areas CA1, CA2 and CA3, while the dentate gyrus granule cells are not labeled. The higher magnification of CA1, CA2 and CA3 showed staining only in the proximal regions of dendrites. The BC200 sense probe did not show any signal.

3.16 Subcellular-localization of MS2TAP fusion protein on coronal sections of transgenic mouse brain

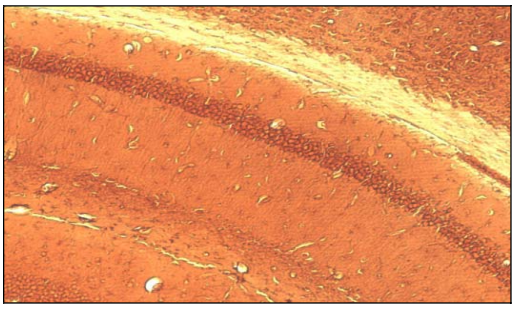
To determine the localization and expression pattern of the MS2TAP protein in BC1MS2B+MS2TAP_in BC1 $RNA^{-/-}$ and BC200MS2+MS2TAP transgenic mice, we performed immunocytochemical localization utilizing antibodies against MS2 and HA as MS2 domain and HA peptide are part of the MS2TAP fusion protein in transgenic mice. PABP was used as positive control in this experiment (Muddashetty et al., 2002). Immunocytochemistry was performed with DAB (diaminobenzidine) substrate on the mouse brain coronal sections. MS2, HA-immunoreactivity (MS2-IR and HA-IR) was found uniformly distributed in all neuronal cell populations of all brain areas examined, with exception of little difference in staining with HA-immunoreactivity, where hippocampus was not stained as much as with MS2 and PABP. MS2, HA and PABP immunoreactivity was found in somal cytoplasm. Intense immunoreactivity was observed in hippocampus of pyrimidal cells and in the cells of the dentate gyrus. Immunoreactivity is roughly equal in both transgenic mice (BC1MS2B+MS2TAP_in BC1 $RNA^{-/-}$ and BC200MS2+MS2TAP) with MS2 and HA antibodies and with PABP [as positive control]. In the experiment pre-immune sera of MS2 antibody and PABP antibody was used as a negative control (data not shown because we did not observe any immunoreactivity).

In wild type mice and BC1MS2B+MS2TAP_in BC1 $RNA^{-/-}$ transgenic mice, PABP and MS2 (both antibodies raised in our lab) had shown analogous localization pattern. Intense dendritic staining was observed in stratum radiatum of CA3 and significant but less staining in CA1 and CA2. In case of HA antibody, very intense staining had been observed in the CA3 region of hippocampus only and only weak or no staining in the CA1 and CA2 regions **(Fig 3.20)**. Dentate gyrus also shows intense staining in case of HA antibody with significant but less staining with PABP and MS2 antibodies. The reason for the differential staining is not known; it is possible that the HA peptide antibody might be more accessible to the CA3 layer as compared to PABP and MS2 antibodies.

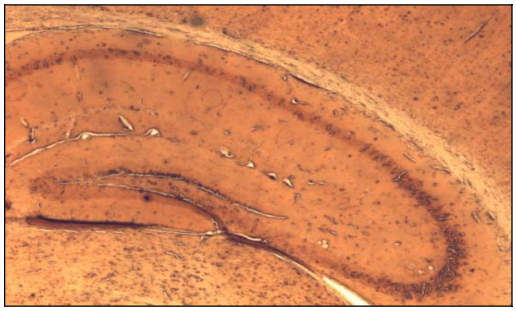
In BC200 and BC200MS2+MS2TAP transgenic mouse brain coronal sections with PABP, MS2 and HA antibodies have shown similar immunoreactivity (IR) **(Fig 3.21)** and intense staining with DAB (diaminobenzidine) substrate compared to wild type mice and BC1MS2B+MS2TAP_in BC1 $RNA^{-/-}$ transgenic mice.



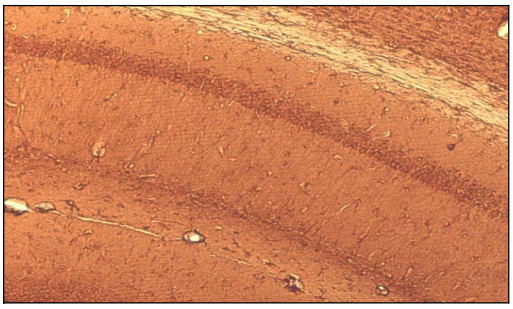
wild type_anti-PABP



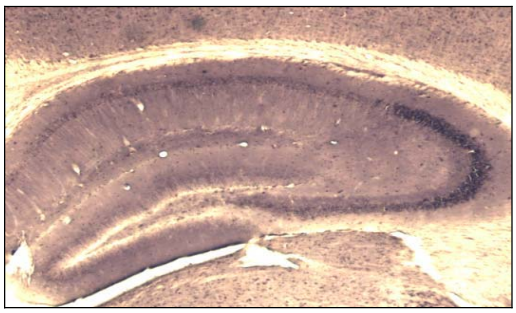
wild type_anti-PABP



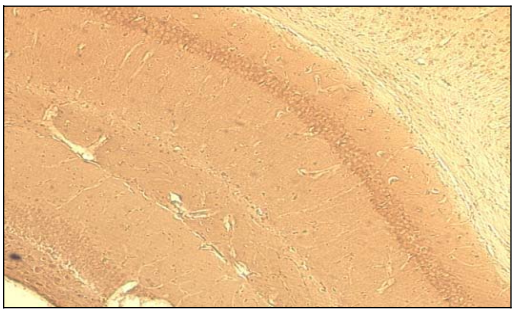
BC1MS2B_MS2TAP_in BC1 RNA$^{-/-}$_anti-PABP



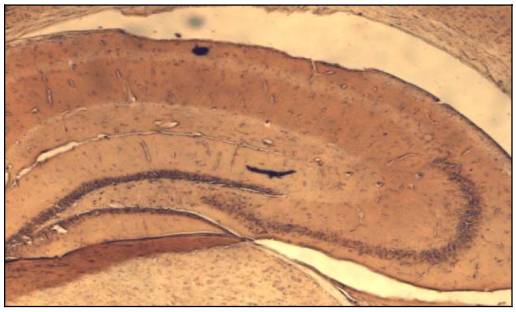
BC1MS2B_MS2TAP_in BC1 RNA$^{-/-}$_anti-PABP



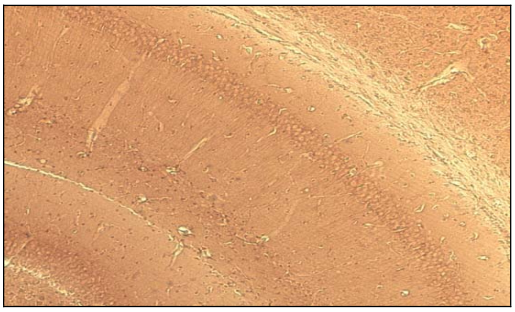
BC1MS2B_MS2TAP_in BC1 RNA$^{-/-}$_anti-MS2



BC1MS2B_MS2TAP_in BC1 RNA$^{-/-}$_anti-MS2

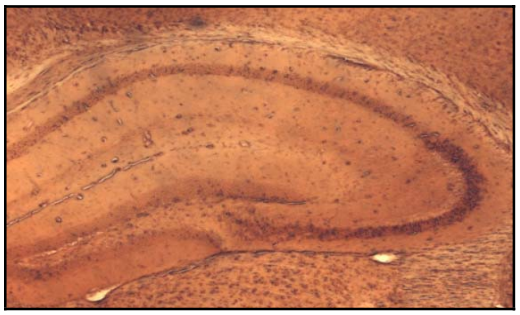


BC1MS2B_MS2TAP_in BC1 RNA$^{-/-}$_anti-HA

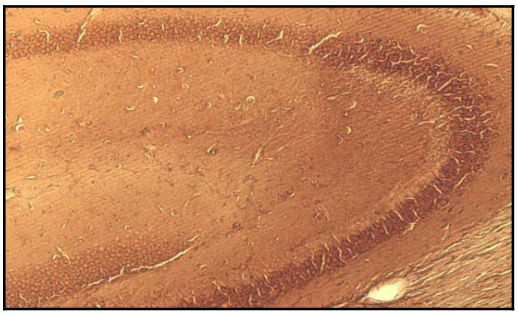


BC1MS2B_MS2TAP_in BC1 RNA$^{-/-}$_anti-HA

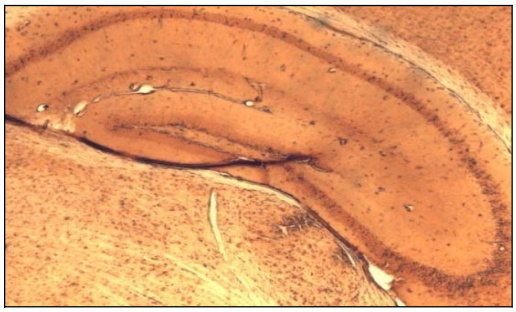
Fig 3.20 Immunocytochemistry of MS2TAP protein in BC1MS2B+MS2TAP_in endogenous BC1 RNA depleted (BC1 RNA$^{-/-}$) transgenic mice and wild type mice. Immunocytochemistry of MS2TAP protein localization and expression pattern on coronal sections of wild type mice and BC1MS2B+MS2TAP_in BC1 RNA$^{-/-}$ transgenic mice brain sections through the hippocampus region. Immunoreactivity (IR) was observed with DAB (diaminobenzidine) substrate in both cytoplasm and in dendrites with MS2, HA antibodies and PABP as a positive control. There is intense staining in all the regions of hippocampus, CA1, CA2 and CA3 also the cells in the dentate gyrus with MS2, HA and PABP antibodies.



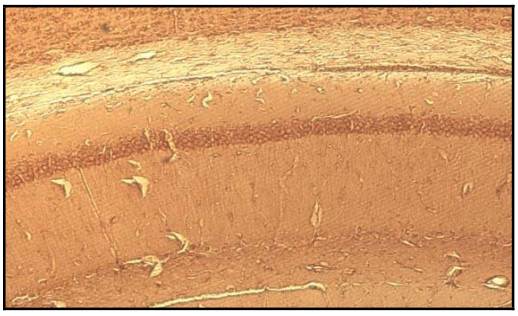
BC200_anti-PABP



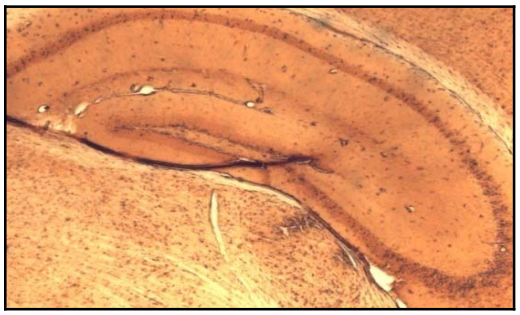
BC200_anti-PABP



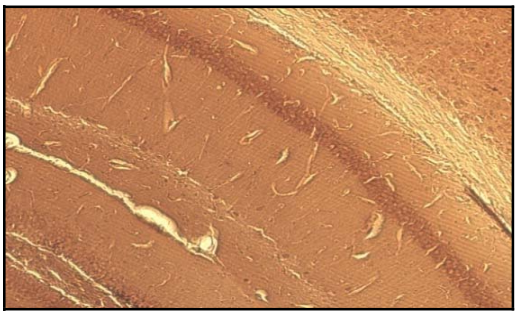
BC200MS2+MS2TAP_anti-PABP



BC200MS2+MS2TAP_anti-PABP

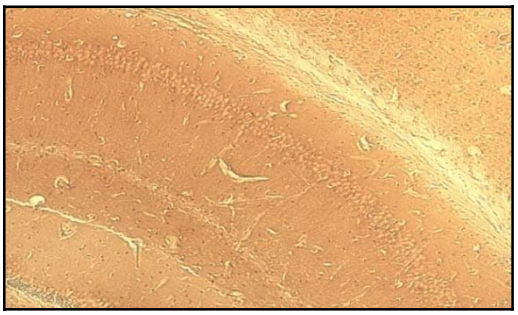


BC200MS2+MS2TAP_anti-MS2



BC200MS2+MS2TAP_anti-MS2

BC200MS2+MS2TAP_anti-HA



BC200MS2+MS2TAP_anti-HA

Fig 3.21 Immunocytochemistry of MS2TAP protein in BC200MS2+MS2TAP and BC200 transgenic mice. Immunocytochemistry of MS2TAP protein localization and expression pattern on coronal BC200 transgenic mouse and BC200MS2+MS2TAPtransgenic mouse brain sections through the hippocampus region. Immunoreactivity (IR) was observed with DAB (diaminobenzidine) substrate in both cytoplasm and in dendrites with MS2, HA antibodies and PABP as a positive control. There is intense staining in all the regions of hippocampus, CA1, CA2 and CA3 also the cells in the dentate gyrus with MS2, HA and PABP antibodies.

Hu proteins are the potential protein partners of BC RNP complexes

3.17 Hu proteins are the potential protein partners of BC RNP complexes

In recent years it has become apparent that post transcriptional mechanisms plays an important role in controlling a number of neural specific genes. These processes include translation, translational repression, transport and stabilization mechanism. ELAV/Hu proteins constitute a family of RNA-binding proteins that participate in the post-transcriptional control mechanisms of several neural genes. The Hu genes encode a large number of alternatively spliced transcripts to produce a series of related neuron-specific RNA-binding proteins named as HuA/R, HuB/HelN1 (Levine et al., 1993), HuC and HuD (Okano and Darnell, 1997). These are a family of highly similar RNA-binding proteins that are conserved among vertebrate species and share sequence similarity with the *Drosophila* genes *sxl* (sex lethal) and *elav* (embryonic lethal abnormal vision). The *ela*v gene is essential for the normal development of the embryonic nervous system. HuB is present in neurons and in gonads; HuC and HuD are restricted to neurons. These proteins contain three RNA recognition motifs (RRMs) and are known to bind AU- rich elements (AREs) in 3' untranslated regions (UTRs) of several mRNAs like GAP-43, c-myc, c-fos and neuroserpin etc (Barreau et al., 2005, Bolognani et al., 2004, Chagnovich et al., 1996, Cuadrado et al., 2002, Mobarak et al., 2000, Hinman and Lou, 2008). The mouse homolog of HuB (mHuB) has been reported to bind to a GAA-rich sequence (Abe et al., 1996). Hu proteins were originally identified as early markers of neuronal differentiation. Recent studies indicate that HuD is involved in learning and memory (Hinman and Lou, 2008). Using several complementary approaches, we have shown that HuB, HuC and HuD might interact with BC RNAs *in vitro*. Thus, Hu proteins could be some of the potential protein partners of BC RNP complexes.

3.18 Immunoprecipitation of BC RNAs with antibodies against Hu proteins

In order to determine the interaction of Hu proteins with BC RNP complexes, we have immunoprecipitated RNP complexes using Hu antibodies (gift from Dr. C. Durga Rao, Indian Institute of Sciences, Bangalore, India). Among the three Hu proteins, HuB had efficiently precipitated BC1 RNA but not BC200 RNA, whereas HuC and HuD antibodies had shown very weak precipitation of both the RNAs **(Fig 3.22)**. This weak precipitation may be due to the antibody titer or due to weak or no interaction with the BC RNAs. As pre-immune serum of HuB, HuC and HuD were not available, which will provide a better solution (picture) to imagine that weather the precipitation of BC RNAs found in HuB was due to the interaction with the HuB or an artifact. Further experiments were not carried out because, lack of antibodies.

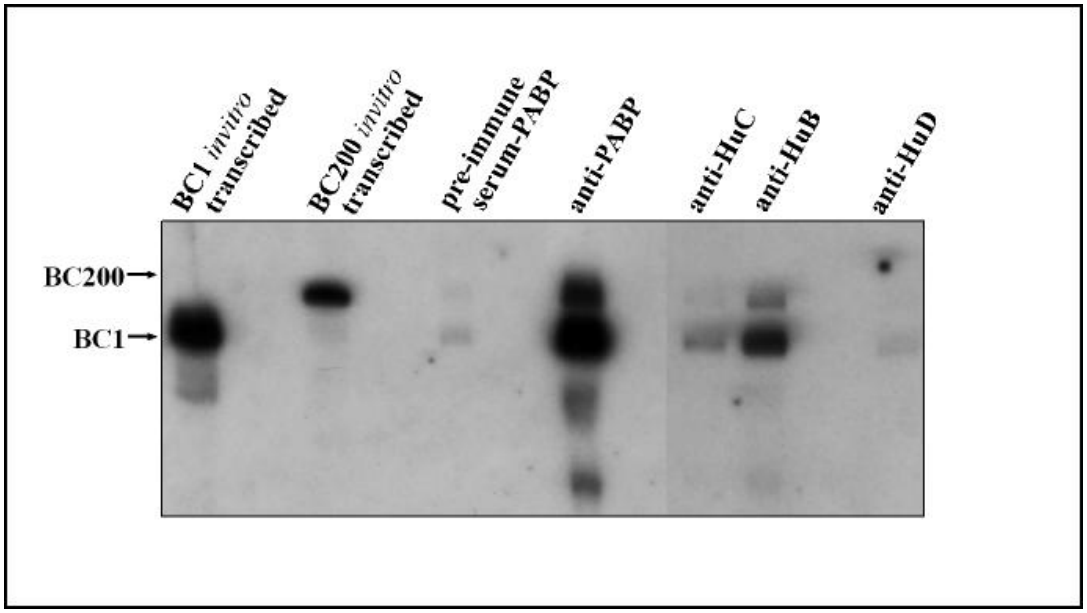


Fig 3.23 Immunoprecipitation (IP) of BC RNAs from BC200 transgenic mouse brain (S2 fraction) extracts with Hu antibodies. *In vitro* transcribed BC1 and BC200 RNAs were loaded as a control. Anti-PABP antibody was used as a positive control, where both the RNAs are efficiently precipitated. A pre-immune serum of PABP was used as a negative control. The extracted RNAs were separated on a 6% polyacrylamide/7 M urea gel and hybridized with oligonucleotide complimentary to the unique region of BC1 and BC200 RNAs (Table 1).

3.19 Localization and expression of Hu proteins in mouse brain

To understand the role role of Hu proteins in the brain, we performed immunocytochemical localization of Hu proteins using Hu antibodies (HuB, HuC and HuD) on mouse brain coronal sections of wild type, BC1MS2B+MS2TAP_in BC1 $RNA^{-/-}$ transgenic mice and BC200 transgenic mice and BC200MS2+MS2TAP transgenic mice. HuB, HuC and HuD-immunoreactivity (HuB-IR and HuC-IR and HuD-IR) was found uniformly distributed in all neuronal cell populations of all brain areas examined. A very intense immunoreactivity was observed in the hippocampus region of all mouse sections with HuB, HuC and HuD. HuB-IR, HuC-IR and HuD-IR immunoreactivity was found in cytoplasm. Intense immunoreactivity was observed in all regions of hippocampus, CA1, CA2 and CA3 and also the cells in the dentate gyrus. Immunoreactivity is approximately equal in all mice brain sections (wild type, BC1MS2B+MS2TAP_in BC1 $RNA^{-/-}$ and BC200 and BC200MS2+MS2TAP) with HuB, HuC and HuD antibodies. In this experiment pre-immune sera of HuB, HuC and HuD antibody was not available, hence we used MS2 pre-immune sera, as a negative control (data not shown because we did not observe any immunoreactivity).

As shown in **(Fig 3.24 & 3.25)** HuB, HuC and HuD had similar localization signal in the cells of CA1, CA2 and CA3 region and also in the cells of dentate gyrus. The signal was observed in the cell bodies and in the dendrtic fields. This exhibits that the dendritic compartments in the different regions are equipped with Hu proteins.

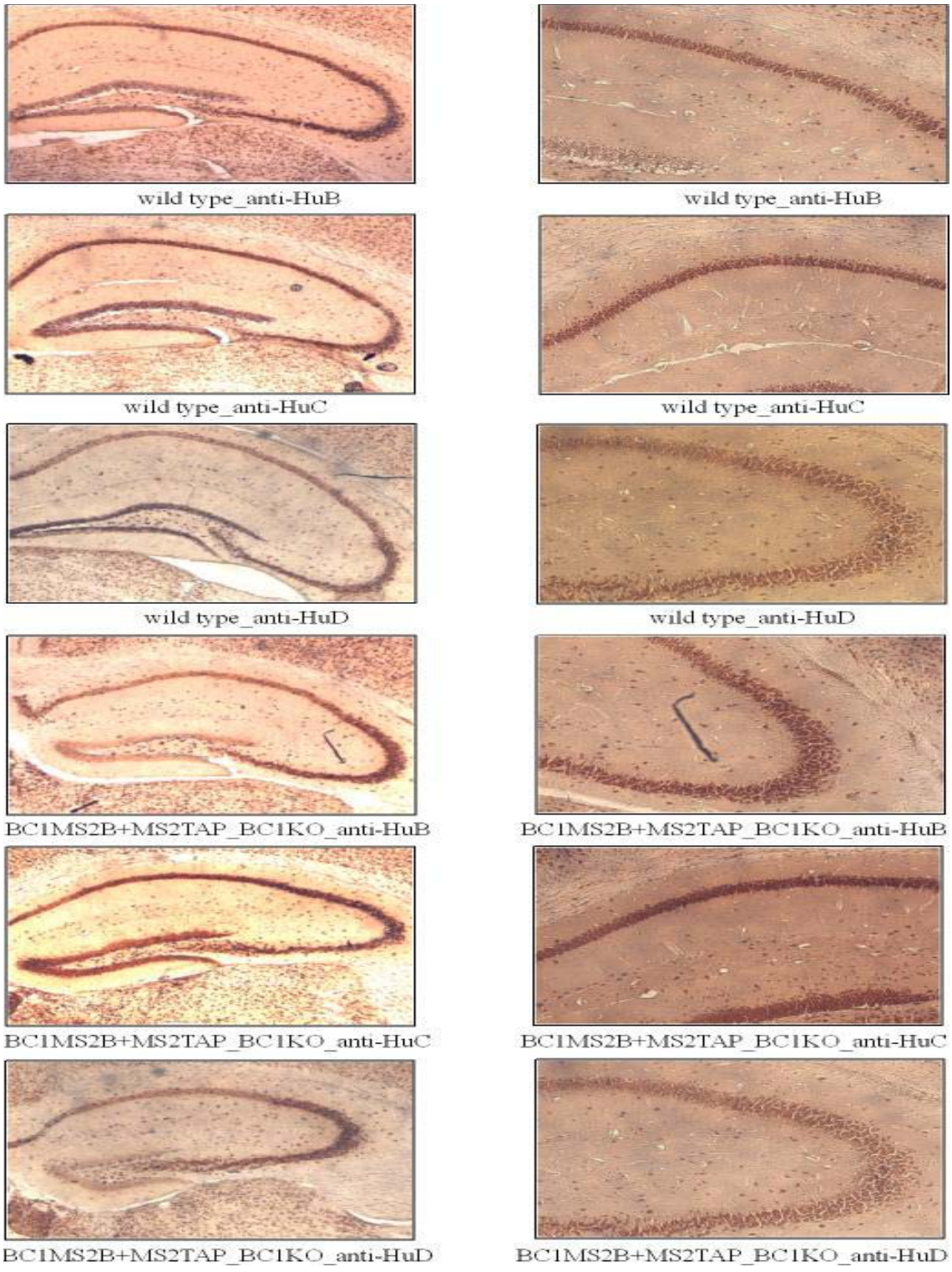


Fig: 3.24 Localization of Hu proteins (HuB, HuC and HuD) in wild type and BC1MS2B+MS2TAP_in BC1 RNA$^{-/-}$ transgenic mouse brain coronal sections of hippocampus. Immunoreactivity (IR) was observed with DAB (diaminobenzidine) substrate in both cytoplasm and in dendrites with HuB, HuC and HuD antibodies. There is intense staining in all regions of hippocampus, CA1, CA2, CA3 and the cells in dentate gyrus.

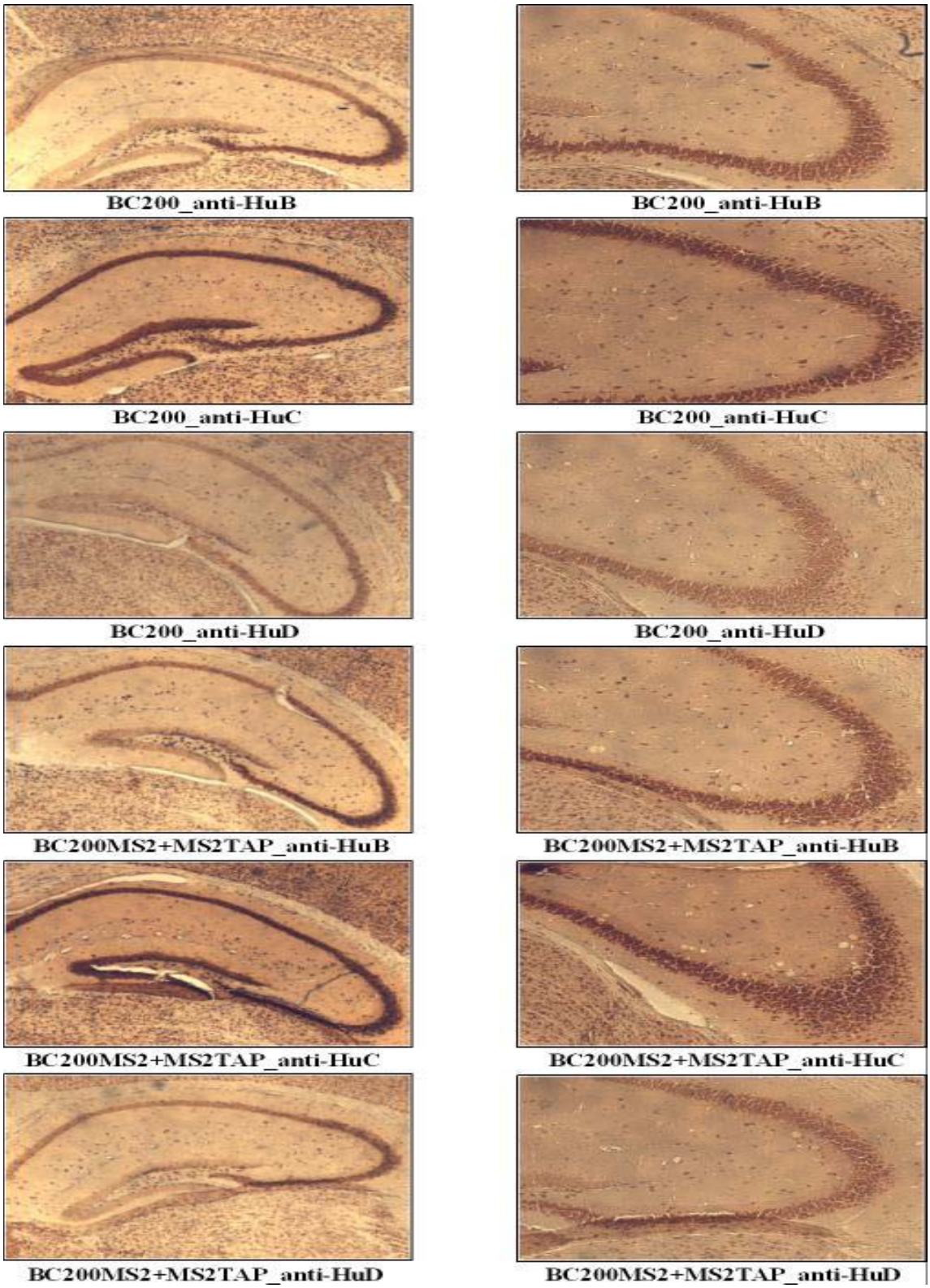


Fig: 3.25 Localization of Hu proteins (HuB, HuC and HuD) in BC200 and BC200MS2+MS2TAP transgenic mouse brain coronal sections of hippocampus. Immunoreactivity (IR) was observed with DAB (diaminobenzidine) substrate in both cytoplasm and in dendrites with HuB, HuC and HuD antibodies. There is intense staining in all regions of hippocampus, CA1, CA2, CA3 and the cells in dentate gyrus.

In order to show that Hu proteins might interact directly with BC RNAs and might be part of BC RNP complexes, we performed Electrophoretic Mobility Shift Assays (EMSAs) or gel- retardation assays. Radio labeled (^{32}P-labeled) BC RNAs (see Materials and Methods – section 2.2 a) were incubated with increasing concentration of full-length recombinanat Hu proteins (HuB, HuC and HuD) from 5 nM to 500 nM and fixed concentrations of radio labeled BC RNAs (0.01 picomole for BC1 RNA and 0.015 picomole for BC200 RNA) and run on 6% polyacrylamide gel [native conditions]. Our results show that Hu proteins (HuB, HuC and HuD) might interact as Hu proteins are known to bind A-rich regions and AU rich regions. It is tempting that Hu proteins might bind to BC RNAs (BC1 RNA and BC200 RNA) that harbor internal A-rich regions and AU rich regions. Both ^{32}P-labeled BC RNAs bind to the recombinant HuB, HuC and HuD proteins with different affinities (K50) was shown **(Fig 3.26).** HuD has high affinity (approximate K50 ~30 nM) with both BC RNAs, whereas HuC has intermediate affinity (approximate K50 ~250 nM) and HuB has low affinity (approximate K50 ~500 nM) [BC1 RNA showed further low affinity compared to BC200 RNA].

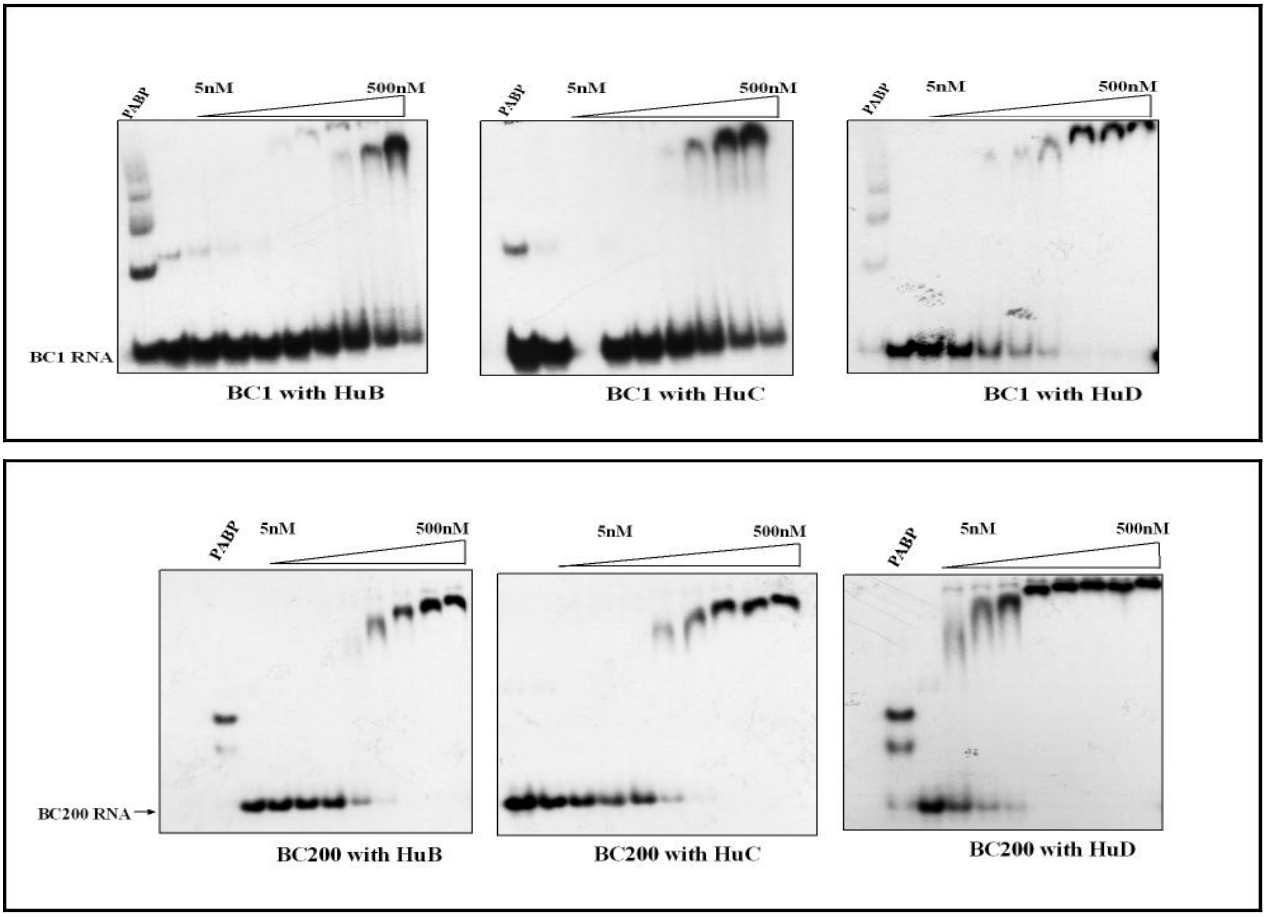


Fig 3.26 Electrophoretic mobility shift assays (EMSAs) showing binding of recombinant HuB, HuC, and HuD to BC RNAs. Binding experiments were performed using a fixed concentration of ^{32}P labeled RNA probe (10 nM) with a range of protein concentrations, indicated in nM. The first lane in each case represents the positive control with PABP protein (15 nM) except in BC200 with HuC. The second lane is free RNA (without any protein).

3.20 Hu expression and cross reactivity

There were reports that Hu genes encode a large number of alternatively spliced transcripts to produce a series of related neuron-specific RNA binding proteins, HuA/R, HuB, HuC and HuD. These proteins show high sequence similarity and high cross reactivity. We examined the reactivity of Hu antibodies with wild type and BC200 transgenic mouse brain extracts (S2 fraction), from Western blot analysis that Hu proteins may be part of BC RNP complex **(Fig 3.27)**. In order to determine weather these Hu antibodies were cross reacting with poly(A)-binding protein, we performed immunoblotting and confirmed that Hu antibodies were cross reacting among each other but not with poly(A)-binding protein (PABP). Further experiments are required to confirm that Hu proteins can be the potential protein binding partners of BC RNAs.

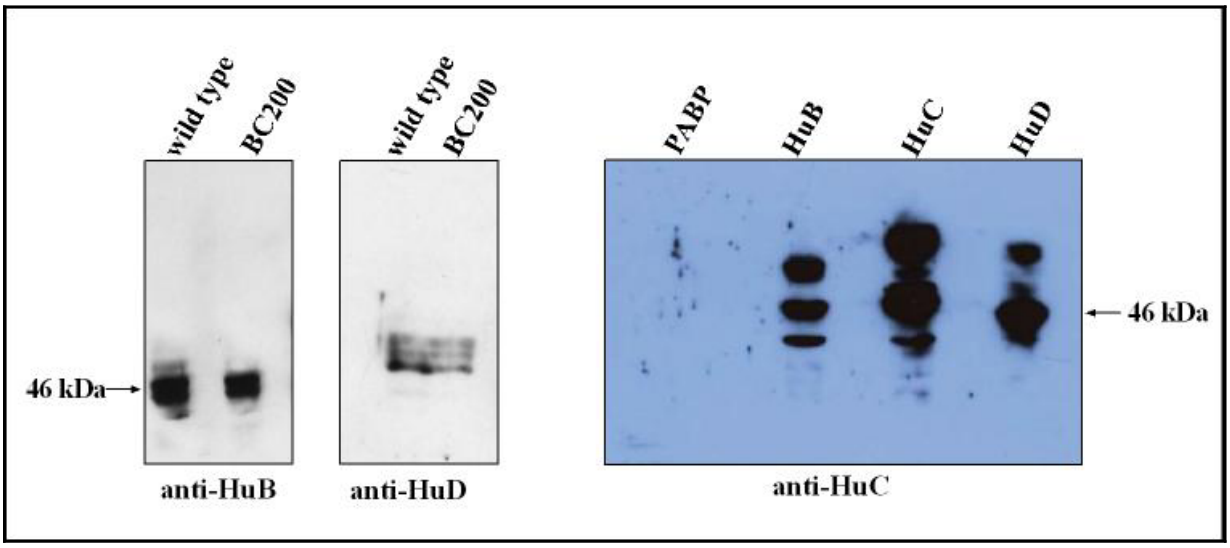


Fig 3.27 Western blot analysis of Hu protein expression in mouse brain extracts. Hu proteins (HuB, HuC and HuD) expressed in mouse brain extracts (S2 fraction) of wild type and BC200 transgenes. Mouse brain extracts (S2) were run on 10% polyacrylamide gel electrophoresis containing sodium dodecyl sulfate (SDS-PAGE), transferred on to polyvinylidene difluoride (PVDF) and probed with polyclonal Hu antibodies. Wild type and BC200 transgenic mouse brain extracts showed the presence of Hu proteins. As Hu proteins are splice variants, showing cross reactivity among the recombinant purified Hu proteins but not with the recombinant purified PABP.

Application of SINEs in rodent phylogeny

3.21 Conserved Presence/Absence Loci (CPAL) of SINEs

From the CPAL search out of 232 loci, 14 loci were previously reported (Farwick et al., 2006). We have selected two loci out of the remaining 218 loci, as this strategy is focusing only on the mouse-related clade. In addition, we further extended the species sampling to the previously published 6 loci (2, 5, 8, 13, 14 & 16) for PCR amplification and found important information concerning phylogenetic relationships among rodents.

3.22 Informative SINE-loci from alternative strategies

As the CPAL search was based on SINE elements in mouse, we further developed alternative strategies to resolve the rodent phylogeny applying the mouse based strategy that includes other clades (Ctenohystrica and Squirrel-related clade).

In this search, we have selected mouse intronic sequences and blasted them against available trace sequences of *Dipodomys ordii, Cavia porcellus and Spermophilus tridecemlineatus*. From these we selected 7 loci from *D.ordii*, 7 loci from *C.porcellus* and 3 loci from *S. tridecemlineatus* for PCR amplification. From the 17 loci, we found 21 phylogentically informative presence/absence markers.

3.23 SINE elements in the Mouse-related clade

From the CPAL search, we found a B3A element (marker 17), that group together all Muroidea. ID4 element (marker 18) was found in all three major clades, compromising the Mouse-related clade, Ctenohystrica and the Squirrel-related clades, supporting rodent monophyly (Cao et al., 1994, Farwick et al., 2006, Huchon et al., 2007, Martignetti and Brosius, 1993b). Previously analyzed markers 2, 5, 8, 13, 14 & 16 were further extended with newly available rodent species, which were not previously used in our sampling; they yielded some additional information. A B1 SINE (marker 2a, **Fig 3.28**) supports the grouping of Muridae, that include *Mus musculus, Rattus norvegicus, Gerbillus dasyurus, Psammomys obesus, Meriones tristami and Meriones unguiculatus.* Two SINEs B3 and a B1-related SINE (marker 8a & 5a, **Fig 3.28**) also supports the grouping Muridae.

An alternative strategy was developed for *Dipodomys* (mouse intronic sequences blasted against *Dipodomys ordii*), which yielded very interesting markers, that have not been shown previously. Four SINE elements were found, two IDL-Geo and two PB1-Geo, (markers 19 and 23 are IDL-Geo and markers 20 and 21 are PB1-Geo, **Fig 3.28**) which is characteristic of the superfamily *Geomyoidea*, present in *Dipodomys ordii*, *Dipodomys californicus*, *Dipodomys merriami* and absent in *Heteromys gaumeri*, and other rodent families in the mouse-related clade. These markers support the classification of *Heteromyidae*. Three retroposed SINEs, two PB1-Geo, and one B1F1 (markers 22, 24 and 25 respectively, **Fig 3.28**) clustered

Dipodomys ordii, Dipodomys californicus and Dipodomys merriami and Heteromys gaumeri and were absent in other rodents. Thus, these markers support the classification of Geomyoideae, which include Geomyidae and Heteromyidae families.

3.24 SINE elements in Ctenohystrica

Mouse intronic sequences blasted against *Cavia porcellus* trace sequences. From the blast, we found three orthologous ID4 elements (markers 28, 31 and 32c, **Fig 3.28**) in all rodents. Thus supporting the monophyly of all rodents and the corresponding SINE is absent in lagomorphs selected used as an outgroup. Two B1F SINEs (markers 26a and 27a respectively, **Fig 3.28**) were present in all investigated Ctenohystrica except *Ctenodactylus gundi*, *Laonastes aenigmamus* and other rodent clades including Mouse-related clade and Squirrel-related clade, grouping Caviomorpha and Phiomorpha supporting the grouping of Hystricognathi and rejecting the grouping of Laonastes and Hystricognaths. SINE B1 and ID4 (markers 29 and 30 respectively, **Fig 3.28**) were found in *Cavia porcellus, Dolichotis patagonium* and *Agouti taczanowskii*, supporting the grouping of families Caviidae and Agoutidae.Two further integrations, both ID2 SINEs (markers 27b and 32a, **Fig 3.28**) grouping *Proechimys cuvieri*, *Myocastor coypus* and *Capromys pillorie* and *Cavia porcellus, Dolichotis patagonium* respectively.

3.25 SINE elements in the Squirrel-related clade

In our search, ID-str (marker 33, **Fig 3.28**) was found in *Spermophilus tridecemlineatus* and *Aplodontia rufa*, but was not found in other rodents. A B1F SINE (marker 34, **Fig 3.28**) was found in all investigated Squirrel-related members, but not in other rodents thus support the monophyly of Squirrel-related clade. A ID2 SINE found in *Spermophilus tridecemlineatus, Marmot marmota* and *Sclurus vulgaris*.

3.26 ID4 SINE and rodent monophyly

All our search strategies except the *Dipodomys ordii* biased search, yielded at least one ID4 SINE from each search and a total of five ID4 SINE elements, that group all rodents (markers 18, 28, 31, 32 and 35, **Fig 3.28**) analyzed. The corresponding loci were empty in lagomorphs, supporting the monophyly of rodents.

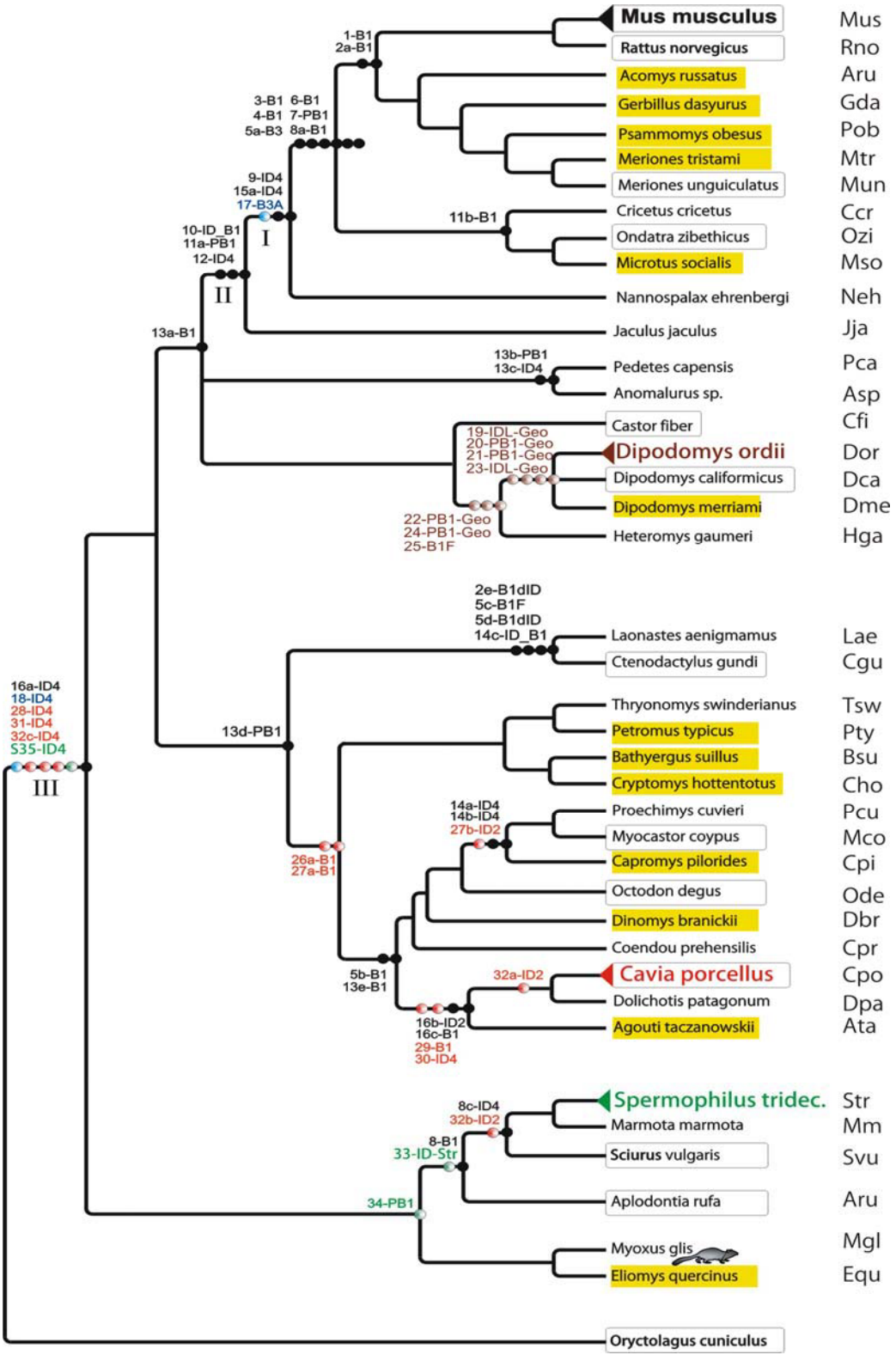


Fig 3.28 SINEs retroposed elements as molecular cladistic markers in rodent phylogeny. Roman numerals refer to previously discovered multilocus markers (I) Muroidea [B2 SINEs, (Serdobova and Kramerov, 1993)], (II) Myodonta [4.5S_H RNA, (Gogolevskaya et al., 2005)], (III) Rodents [BC1 RNA, (Martignetti and Brosius, 1993b)]. The black circles refer to previously published SINE markers (Farwick et al., 2006). SINE retroposed elements found in this study (gray, red and green circles are from strategy I, II and III respectively, see Materials and Methods).

Discussion

4.0. Impact of affinity tags on expression of BC RNPs

RNA-protein (RNP) complexes play fundamental roles in many biological processes. Classical example of RNP complex is ribosome, where a number of RNAs and proteins interact with each other in complicated manner to carry out an important cellular function, namely protein synthesis. Spliceosome complex is another prominent example of RNP complex, where small nuclear RNPs (snRNPs) have an essential role in the maturation of pre-mRNA into mRNA. SRP (Signal Recognition Particle) is another important complex playing a role in translational elongation arrest and protein secretion. The BC RNPs (brain-specific cytoplasmic RNAs/RNPs) namely BC1 and BC200 exclusively expressed in rodents and anthropoid primate nervous systems respectively (Cao et al., 2006, DeChiara and Brosius, 1987, Martignetti and Brosius, 1993b, Martignetti and Brosius, 1993a). They might function in protein translation in neuronal dendrites via translational repression (Kondrashov et al., 2005, Wang et al., 2005).

To date, close to a dozen proteins have been reported to be components of BC1 RNP complex, most of them are *in vitro* or interpretational artifacts (discussed in the introduction). General biochemical methods namely gel filtration, sucrose gradients, Mono-Q and DEAE affinity purification etc (Muddashetty thesis) have been used to enrich the BC1/BC200 RNP complexes. None of these methods could succeed in complete purification of BC RNPs. Therefore, we have generated the basis for a generic RNA tag based affinity purification for these BC RNPs to identify the proteins bound to these complexes.

Protein affinity tags namely histidine tag, glutathione-S-transferase (**GST**) tag, FLAG tag etc., are fused to the protein of interest for purification and detection (localization) purposes. Unlike protein affinity tags, RNA affinity tags were not frequently used to purify the RNP complexes assembled endogenously. But recently RNA affinity tags have been started coming into the picture, where RNA-protein interactions like MS2/MS2 coat protein, λ phage N anti-terminator protein, PP7 coat protein (Lim and Peabody, 2002, Peabody, 1990, Tars et al., 2000) are the most prominent examples of this kind. Based on the MS2 coat protein, we generated a TAP tag fusion protein that should bind to MS2 RNA stem tagged BC RNPs for purification and localization.

In general, the tag should be placed at the 3' end or 5' end of the RNA for either purification or detection. The folding problem is simply a matter of inserting the tag in such a way that both the tag and RNA of interest should fold correctly. Furthermore, the tagged RNA's expression level should not be too low when introduced into the system (*in vivo*, *ex vivo* or *in vitro*). So keeping this information in our mind, we had tagged BC1 and BC200 RNAs with two MS2 aptamers at the 3' end and the corresponding MS2 coat protein fused to the TAP tag.

4.1 Characterization of tagged BC RNAs in transgenic mice

The three main factors (I) folding (II) steric hindrance and (III) keeping the tag (not losing the tag from BC RNAs) to the RNAs from processing to study either localization or purification were considered while tagging MS2 motifs to the BC RNAs. There are reports that minimum two MS2 stem motifs are required to bind efficiently to MS2 coat protein (Ko and Gunderson, 2002). Therefore, in our study two MS2 stem motifs were used to tag BC RNAs or the MS2 stems were placed internal to BC1 RNA. The folding of BC RNAs with MS2 stems are shown in **(Fig 3.0 & 3.1)**. BC1 RNA with two MS2 stems showed no apparent interference with folding of the predicted BC1RNA structures with MS2 stem motifs. While mfold application showed a different folding of BC200 RNA with two MS2 stem motifs (data not shown).

The MS2 coat protein is fused to TAP tag (Rigaut et al., 1999) for two step affinity purification. **Figures 3.0, 3.9 & 4.0**. shows the chimeric RNAs (BC1 with two MS2 stems and BC200 with two MS2 stems) and the MS2 TAP protein complexes (BC1MS2/MS2TAP and BC200MS2/MS2TAP) respectively.

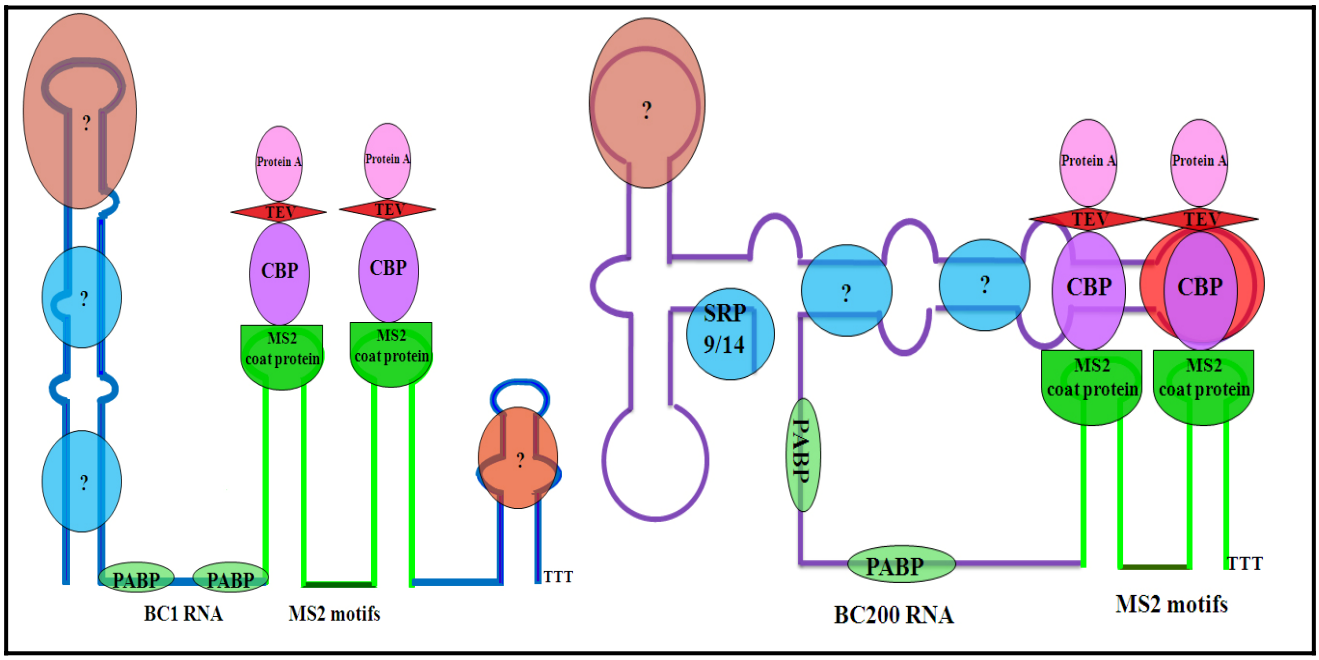


Fig 4.0 Predicted model showing the BC RNAs tagged to MS2 motifs and the fusion protein MS2TAP complexed with BC RNPs. The MS2 RNA motifs [green] internal to BC1 RNA [blue] and at the 3' end of BC200 RNA [purple] are shown. MS2 coat protein [green half oval] fused to TAP tag contains CBP tag [calmodulin binding peptide - purple] and protein A [ProtA – magenta] separated by TEV protease site [tobacco etch virus -red] are shown. The MS2TAP fusion protein bound to the MS2 RNA motif to facilitate affinity purification of the respective BC MS2/MS2TAP RNP complexes. Model also exhibit PABP (poly(A)-binding protein) bound to both BC RNAs and in addition to SRP9/14 (signal recognition particle) bound to BC200 RNA and the other protein partners bound to the BC RNAs at different places is in question mark.

We have cloned the chimeric RNAs, BC1MS2 and BC200MS2 under the native promoter of the BC1 and BC200 RNA genes, respectively. The MS2TAP protein cloned under the β-actin promoter, which was previously used in our lab and showed high expression in transgenic mice (Khanam et al., 2007a). For our convenience both the RNA expressing and the protein expressing cassettes initially were cloned into a single plasmid (discussed in Materials & Methods, section 1.13. and Results 1.5).The combined construct was used in NIH 3T3 transfections, both RNA and protein were expressed at relatively high levels **(Fig 3.3 & 3.4)**. Therefore, we generated transgenic mice for the *in vivo* purification and the localization. We failed to detect BC1MS2 RNA expression in transgenic mouse brain extract in any of the BC1MS2 [MS2 at 3' end] transgenic mice generated **(Fig 3.5)**. But MS2TAP protein was expressed in all transgenes analyzed from the same transgenic mice **(Fig 3.6)**. This may be due to several reasons (I) processing of BC1MS2 RNA, which can not be ruled out, even though we have not observed any processed RNA in our transgenic mice, but from the NIH 3T3 stably transfected cells, we observed processing as removal of the MS2 stems from the 3' end **(Fig 3.8)**. (II) locus of integration (III) promoter hindrance and transcription termination. These reasons will be discussed later. To answer these questions we had generated BC200MS2 transgenic mice. A weak expression of full length BC200 RNA with MS2 tag and BC200 RNA without MS2 tag was observed, revealing that the MS2 tag was being removed in case of BC200MS2 RNA and that this could have happened analogously for BC1MS2.

In order to achieve the expression of BC1MS2 RNA in transgenic mice, we placed the tag internal to BC1 RNA. There are two reasons behind changing the MS2 tag position into the center (I) 3' processing of the transcript, has been obsereved in case of BC1MS2 stably transfected NIH 3T3 cells **(Fig 3.8)**, where MS2 tag has been partially removed from the BC1MS2 transcript and also other parallel evidences like BC200MS2 transcript also showed a clear processing event and BC200 U6 tag (unpublished, Khanam, thesis), where a U6 RNA tag was removed from the 3' end of BC200 RNA. (II) The RNA polymerase recycling and transcription efficiency [MS2 stems placed at 3' portion of BC1 RNA, we were unable to detect the BC1MS2 RNA expression] will be hindered, if we externally add any structured sequences (Dumpelmann et al., 2003, Goodier and Maraia, 1998). For efficient transcription termination and re-initiation of transcription for multiple rounds by proficient recycling of the RNA polymerase (pol III), merely depends on the terminator sequence. RNA polymerase III transcripts need three or four U residues surrounded by GC rich sequences at the 3' end. The disturbance in these sequences leads to low/no expression due to inefficient recycling of the RNA polymerase III polymerase (Dumpelmann et al., 2003, Goodier and Maraia, 1998). Another factor may also regulate efficient transcription, the folded structure of the transcript depends on the spatial proximity of the 3' and 5' end of the RNAs. Considering all these factors stated above, we have placed MS2 tag internally of BC1 RNA at two different locations. By using RNA folding algorithm mfold by M. Zucker **(Fig 3.1)**, MS2 tag was placed close to the ID stem designated as BC1MS2A **(Fig 3.10)**. MS2 tag

placed further 3' - distal to the A_{22} sequence, designated BC1MS2B **(Fig 3.10)**. We had generated transgenic mice from both the constructs and anticipated BC1MS2A/B RNA expression.

In BC1MS2A transgenic mice, unexpectedly, BC1MS2A RNA expression was not observed in any of the founders. The reasons behind no expression of BC1MS2A RNA in BC1MS2A transgenic mice is not clear. Interestingly, we found low levels of expression of BC1MS2B RNA in three individual founders of BC1MS2B transgenic mice **(Fig 3.11)**; one of the offsprings from one of the founders showed higher BC1MS2B RNA expression. This may be due to the position of integration in the mouse genome.

4.2 Factors influencing the expression of BC RNAs in MS2 tagged mice

In general, transgene expression is influenced by many factors, some are controlled by the elements present in the gene itself that is to be integrated others by the locus of integration or the numbers of integrated copies (usually in tandem). In BC1MS2 and BC200MS2, we have used the native promoters, because BC1 and BC200 transgenes had shown high expression by using their native promoters. From another important study (Khanam et al., 2007b, Ludwig et al., 2005) on transcription analysis and the role of promoter, we learned that removal of upstream promoter elements lead to transcription repression or tissue non-specificity. Therefore we have not shortend the 5' flanking sequences in both constructs. Second reason might be the tagged RNAs are not efficiently transcribed [need RNA polymerase III polymerase recycling] (Dumpelmann et al., 2003) and may be processed as in case of BC200MS2 transgene. Processing the tag was unable to predict and difficult to control, this should be studied differently because this needs understanding of the structural elements behind the processing, which is not the aim of this work.

Transgene expression was also influenced by other important factors, which are host dependent and difficult to control. We had observed in case of BC1MS2B transgene, where BC1MS2B RNA expression was observed only in some transgenic mice but not all. This drives us to think of locus of integration and position of integration. In transgenic mice, integration of the interested gene is random and unable to control the site of integration and the copy (copies). There are reports about the multiple copies or high copy number integration (Davis and MacDonald, 1988, Garrick et al., 1998, Williams et al., 2008) of transgene in to the genomic locus leads to no expression of the desired gene. This had been observed in case BC1MS2B transgene with multiple copies **(Fig 3.12)**.

(a) Position-Effect Variegation (PEV)

Transgenes integrated into the genome randomly, is not possible to control the place of integration, and varies from individual to individual. Stably integrated transgenes are prone to low expression due to position effects that are caused by the site of chromosomal integration (Weiler and Wakimoto, 1995, Garrick et al., 1998). Postion of integration might play a major role in controlling the BC1MS2 RNA expression (we had not observed any expression when MS2 was placed at the very 3' end of BC1 RNA), same for

BC1MS2A RNA expression (MS2 stems closer to the ID domain) in transgenes. An other serious problem, we had encountered in establishing BC1MS2B transgenic mice was the transgenes frequently integrate as multiple copies in a concatameric array, leading to no expression (Williams et al., 2008). Position of integration might not be ruled out the low expression in both the MS2 tagged BC RNAs. Nevertheless, we have observed specific BC1MS2B or BC200MS2 RNA expression only in brain. Incase of MS2TAP transgenic mice, we observed protein expression in all offsprings of all founders of the MS2TAP transgenic mice. We are assuming that the established transgenic mice (BC1MS2B+MS2TAP_ BC1 RNA$^{-/-}$) might be a low copy number transgene (single copy), result in high fusion protein or RNA expression (Williams et al., 2008).

(b) Tissue specificity of BC MS2 transgenic mouse line

In order to determine the MS2 tagged BC RNAs expression in various tissues in the transgenic mouse line, we have performed the Northern blot. The total RNA from different tissues of all transgenic mouse founders were analyzed and confirmed, that MS2 tagged BC RNAs expressed only in brain not in any other tissues. This supports our previous findings that BC200 and G22 transgenic mouse models also showed expression of BC200 RNA or G22 RNA only in brain (Khanam et al., 2007b, Ludwig et al., 2005, Skryabin et al., 1998). It has been shown that upstream sequence elements in BC1 and BC200/G22 locus are responsible for tissue specific expression (Martignetti and Brosius, 1995). BC1 RNA gene 1.4 kb (*Sac*I-*Bam*HI) and BC200 RNA gene 3.2 kb (*Eco*RI-*Hind*III) fragments contains sufficient information to direct tissue specific expression. Hence, we have used the same fragment length, when we cloned BC1MS2 and BC200MS2 RNA genes. BC200MS2 RNA expression observed only in brain not in other tissues **(Fig 3.7)** from BC200MS2 RNA transgenic mouse and BC1MS2B RNA expression also observed only in brain from BC1MS2 RNA transgenic mouse (data not shown).

(c) Establishing BC1MS2B+MS2TAP_in BC1 RNA$^{-/-}$ and BC200MS2+MS2TAP transgenic mouse lines

Over a period of difficulties, we could finally identify the transgenic mice, which expresses BC1MS2B RNA (BC1MS2B transgenic mouse) even though expression levels are much lower compared to wild type mice for BC1 RNP complex purification with few mice. In case of BC200MS2+MS2TAP, expression levels are also weak compared to BC200 transgene (Khanam et al., 2007b), which made it difficult to obtain conclusive results in the first trial runs in purification RNP complexes. For purification of these complexes, we need lot of mice to obtain the required results. Thus, we need to establish transgenic mice, which is homozygous to BC1MS2B RNA+MS2TAP and BC200MS2+MS2TAP transgene. So that the resulting transgenic mice may have high probability of RNA/RNP expression. Moreover making the mouse line homozygous to transgenes has several advantages. First, all animals born from homozygous parents will

be homozygous to transgenes and those animals need not to be checked for the transgene integration and their expression. Second, expression level of the transgene is usually doubled in case of homozygous animals, this could be helpful incase of dendritic transport of BC1 and BC200 RNA. In case of BC1MS2B+MS2TAP transgenic mice, we crossed with BC1 RNA depleted mice to obtain BC1MS2B+MS2TAP_in BC1 RNA$^{-/-}$ transgenic mice, where we can use unique oligoribonucleotide probe against BC1 RNA for localization of BC1MS2B RNA and also eleminate competition of native BC1 RNA/RNP.

4.3 RNA localization and dendritic transport

In general RNA localization contributes to cell polarity and synaptic plasticity. Several RNAs have been reported to be localized in dendrites, including microtubule-associated protein (MAP2), (Garner et al., 1988, Kleiman et al., 1990) the α-subunit of calcium/calmodulin-dependent protein kinase II (α-CAMKII), (Miyashiro et al., 1994) Arc protein, (Lyford et al., 1995) and BC1 RNA and many more. Several factors influence the localization of these RNAs namely *trans* acting and *cis* acting factors. It has been shown that a *cis* acting targeting element is required for BC1 RNA localization into the dendrites (Muslimov et al., 1997). There are many evidences that RNAs are transported up to the distal parts of dendrites, where local translation takes place. (polyribosomes are found in synaptodendrosome preparations). It has been reported that, BC RNAs complex with proteins to form a ribonucleoprotein particle and transported to dendrites and their function is not known. Muslimov et al., (Muslimov et al., 2006, Muslimov et al., 2002) had shown the chimeric BC1 RNA (5' BC1 chimeric *bcd* RNA) has been delivered to the distal dendritic domains by microinjecting into sympathetic neurons in culture. They further demonstrated that 74nt 5' BC1 stem-loop is sufficient to direct distal dendritic delivery of a normally nondendritic RNA (bcd RNA of *Drosophila melanogaster*). However natural *in vivo* system contradicting the microinjection studies, where 5' ID or the full length BC1 RNA was not able to deliver the EGFP message to the dendrites, when inserted into the 3' UTR of the reporter mRNA. However the 3' UTR of α-CamKII mRNA, which is normally found in neuronal dendrites, could efficiently target the reporter mRNA to dendrites *in vivo* (Khanam et al., 2007a). Nevertheless, the mechanism behind these BC RNAs transport *in vivo*/ *ex vitro* (microinjected cells) still not very clear, even though Muslimov et al., had shown that 5' ID region is responsible for transport by dissecting most of the nucleotides and precisely pinpointing the U22 and basal stem is responsible for dendritic transport. It is still an open question how these RNAs are transported or localized into the real/natural *in vivo* system (transgenic mice). For this purpose, we have tagged both the RNAs with MS2 aptamer to study the transport besides purification. Furthermore, for future studies at the electron microscope level we need to have antibodies available (e.g., MS2) to detect a protein bound to the BC RNAs using the immunogold labeling technique (Hoyer et al., 1979, de Harven et al., 1984).

In situ hybridization studies on coronal sections of MS2 tagged BC RNAs revealed, thus far, that the RNAs are localized in the cell body and are not transported up to the dendrites. Both MS2 tagged BC RNAs clearly showed expression in CA1, CA2 and CA3 pyramidal cells of hippocampus **(Fig 3.18 & 3.19)**. We have not observed dendritic localization of both MS2 tagged BC RNAs. Expression pattern of both tagged BC RNAs was analogous to the BC1 RNA and BC200 RNA expression of wild type mice and BC200 transgenic mouse model respectively, except the dendritic localization and low expression **(Fig 3.18 & 3.19)**. Our results from BC200 MS2 transgenic mice, showed that MS2 tag has been processed in BC200MS2 transgenic mice, we have observed processing of about ~10% - 20% of the full-length RNA as judged by Northern blot analysis **(Fig 3.7)**. BC200MS2 tagged RNA is not transported up to neuronal dendrites, it remains in the cell body, which was identified by using BC200 unique riboprobe **(Fig 3.19)**. In case of BC1MS2 (MS2 at 3' end of BC1 RNA), BC1MS2 RNA was unable to detect in any of the transgenic mice. However, MS2 tag when placed internal to the BC1 RNA had shown a weak BC1MS2B RNA expression. In BC1MS2B transgenic mice, where processing of the MS2 tag can not be ruled out, even though, we were unable to detect the processed form of BC1MS2B RNA also showed similar localization **(Fig 3.18)** pattern as BC200MS2. This analogous expression pattern in both MS2 tagged BC RNAs, revealed that the tagged BC RNAs are unable to travel up to the dendrites. The reasons behind the tagged RNA that are unable to transport up to the dendrites are not clearly understood. However, our initial studies on the transgenic mice of MS2 tagged BC RNAs had promised to reveal the fates of these MS2 tagged BC RNAs dendritic transport, which was the previous shown in our laboratory, where 5' portion of BC1 RNA or full length BC1 RNA was unable to deliver the message, in natural *in vivo* systems (Khanam et al., 2007a).

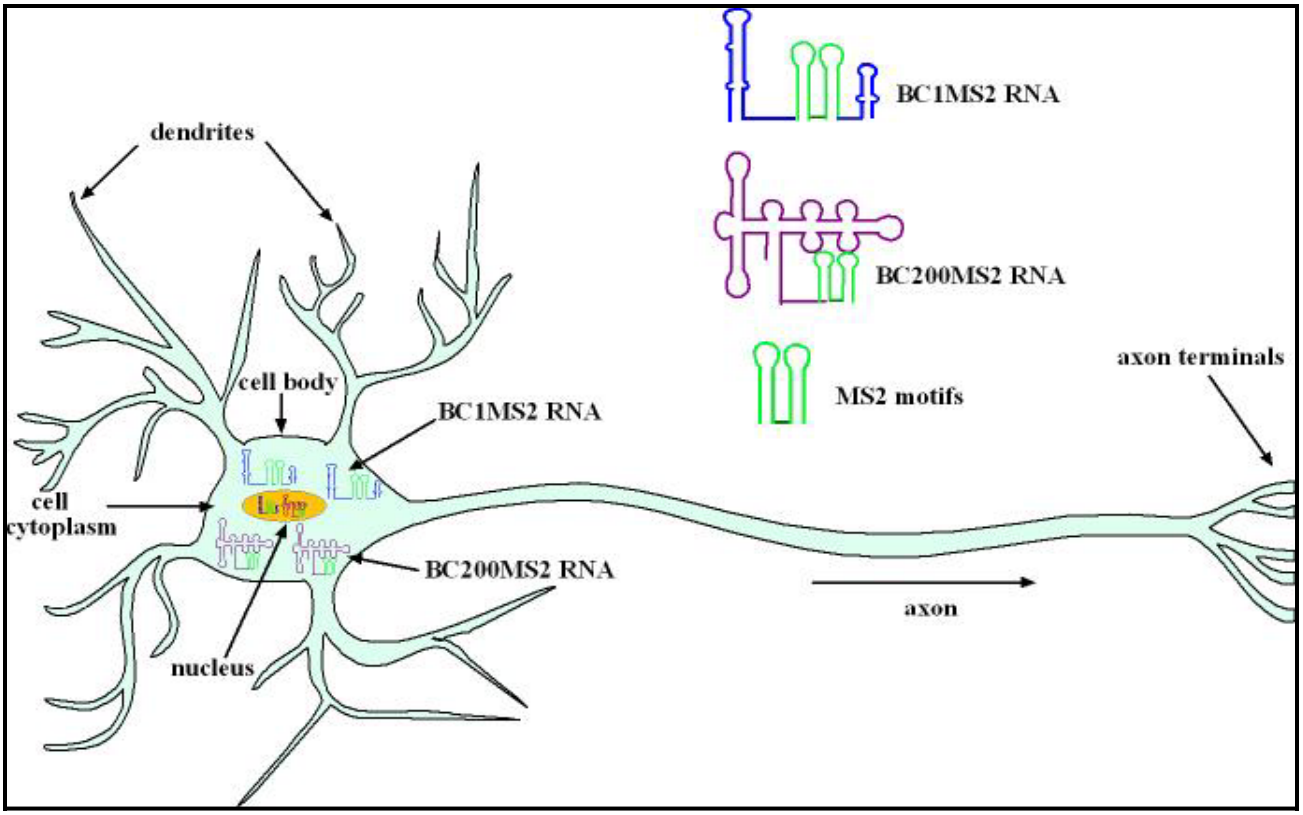


Fig. 4.1 A model representing the *in vivo* fate of BC RNAs with MS2 tag transcripts in neurons of the corresponding transgenic mice (BC1MS2B+MS2TAP_BC1 RNA$^{-/-}$ and BC200MS2+MS2TAP). The blue with green color shaped structures represents BC1MS2B RNA and purple with green color shaped structures represent BC200MS2 RNA. Between the two MS2 stems a short spacer (dark green) was placed for proper folding with BC RNAs. MS2 placed internal to BC1 RNA showed no processing of MS2 tag from Northern, but the RNA remained restricted to the cell body **(Fig 3.18)**. MS2 placed at 3' end of BC200 RNA showed ~10 - 20% processing as judged by Northern blot analysis (**Fig 3.7**) the full length unprocessed BC200MS2 RNA also remained restricted to the cell body **(Fig 3.19)**.

4.4 Reconstruction of Rodent phylogeny-evidence of SINEs

To date SINE insertion analysis is the most reliable strategy to resolve the phylogenetic tree reconstructions, provided several independent markers support the monophyly of each phylogenetic group statistically (Waddell et al., 2001). Using SINE insertion analysis generated an ambiguity-free tree reconstruction of primate phylogeny (Schmitz et al., 2001, Schmitz et al., 2005). The distribution of rodent SINE families and subfamilies (Gogolevskaya et al., 2005, Martignetti and Brosius, 1993b, Serdobova and Kramerov, 1998) has contributed to our understanding of rodent phylogeny. However, no significant phylogenetic information was available based on the analysis of SINE insertions. Recently (Farwick et al., 2006) conducted an exhaustive, automated CPAL search to generate phylogenetic informative SINE markers in rodents, uncovered 31 SINE elements from 16 genomic loci in 22 rodent species, covering representatives of all major rodent clades. The CPAL search requires a well annotated genome, like mouse

and an extensive sequence information from a distinct reference species, like human. From alternative strategies including search against Dipodomys, Ctenohystrica and Squirrel-realted clades for finding more SINE presence/absence patterns, we uncovered diverse phylogentically informative markers. From the 17 genomic loci from three search stratagies, we found 21 markers from 40 rodent species covering all three rodent clades.

Retroposed SINEs offer the most reliable marker system for extrapolating the evolutionary history of any phylogenetic tree reconstruction. Their integration in the genome is random and the probability of same element integration independently at orthologous positions in different species (orthologous) is very low. We have observed this situation in two of our previously published markers (locus 8 and 14), nevertheless both elements are from different sources, and are homoplastic (Farwick et al., 2006), and will not cause erroneous interpretation of relationships. Previous reports mostly based on nuclear (IRBP) and mitochondrial genes (12S rRNA) (DeBry and Sagel, 2001, Murphy et al., 2001) have their own limitations. The presence of similar molecules in different species do not represent common ancestry as compared to SINEs at orthologous position.

4.4.1 SINEs in Mouse-related clade

The mouse-related clade has been divided into three suborders Myomorpha (Myodonta), Anomaluromorpha, and Castoriomorpha (Castoridae + Geomyoidae) (Adkins et al., 2003, Huchon et al., 2002, Murphy et al., 2001). Farwick et al 2006 showed retroposed SINE markers to support the grouping of Myodonta and monophyly of this lineage. From our study a B3A SINE revealed the monophyly of Myodonta, comprises Muroidea and Dipodidae. Several SINEs (Markers 19, 20, 21, 22, 23, 24) supporting the grouping of Geomyoidae and its monophyly, comprise *Dipodomys* and *Heteromys* species. Thus far, we had not found yet any SINE, which group Castorimorpha.

4.4.2 SINEs in Ctenohystrica

There are several phylogenetic relationships within the Ctenohystrica comprising Hystricognathi and Ctenodactylidae. The monophyly of Hystricognathi is supported by morphological, paleontological and recent molecular data (Farwick et al., 2006, Huchon et al., 2007). Two SINE markers (26a and 27a) PB1 and B1F respectively support monophyly of Hystricognathi. Two additional SINEs B1 and ID4 (29 and 30) group Caviomorpha.

4.4.3 SINEs in Squirrel-related clade

Squirrel-related clade was supported by mitochondrial DNA data (Montgelard et al., 2002) clustering Sciuridae + Aplodontidae. From alternative SINE search strategy against Squirrel-related clade, we uncovered a new SINE element named ID-Str (Marker 33, See **Fig 3.8**), which group Sciuridae and

Aplodontidae. In addition, we found a B1F SINE (Marker 34, See **Fig 3.8**), which group Sciuridae, Aplodontidae and Gliridae.

4.4.4 Rodent Monophyly - SINE evidence

Morphological classifications showed the order Rodentia a monophyletic group, even though this has been challenged by different molecular studies (Graur et al., 1991). Nevertheless the hypothesis of rodent paraphyly was disproven immediately (Cao et al., 1994, Martignetti and Brosius, 1993b). Present study revealed five ID4 SINEs (Markers 18, 28, 31, 32, 35) supporting the morphological studies i.e. monophyly of rodents.

Table 3 and 4 summarizes the SINE presence/absence information retrieved from CPAL and other alternative search strategies respectively. The monophyly of rodents supported by ID4 element (Marker 18, 28, 31, 32, and 35) is present in all analyzed rodents clades. The monophyly of rodents has previously supported by the presence of BC1 RNA in all rodents (Martignetti and Brosius, 1993b). Previous results support the monophyly of Mouse-related and Ctenohystrica (markers 13a and 13d) (Farwick et al., 2006). Present results revealed the support of hystricognathi (markers 26 and 27) also support the monophyly of Squirrel-related clade (marker 34). From this study, we have solved another intriguing problem that is grouping of Sciuridae (Squirrel-related clade) which has never been shown previously via retroposon evidence. Unfortunately, we have not found any SINE element so far, which resolve the relationship between the three major clades – Mouse-related clade; Ctenohystrica; and Squirrel-related clade.

Table 3. Detailed presence/absence patterns (CPAL-*Mus musculus* screening)

			1	2a	2b	2c	2d	2e	3	4	5a	5b	5c	5d	6	7a	7b	8a	8b	8c	9	10	11a	11b	12	13a	13b	13c	13d	13e	13f	14a	14b	14c	15a	15b	16a	16b	16c	16d	16e	**17 B3A**	**18 ID4**
Mouse-related clade																																											
Mus musculus	mouse	GD	+	+	-	-	-	-	+	+	+	-	-	-	+	+	▲	+	-	-	+	+	+	-	+	+	-	-	-	-	-	-	-	-	+	-	+	-	-	+	+	+	+
Rattus norvegicus	rat	GD	+	+	-	-	-	-	+	+	+	-	-	-	+	▲	▲	+	-	-	+	+	+	-	+	+	-	-	-	-	-	-	-	-	+	-	+	-	-	+	+	+	+
Acomys russatus	spiny mice		?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	+	?	+	-	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	+	+
Gerbillus dasyurus	gerbil		?	+	▲	-	-	-	?	?	?	?	?	?	?	+	+	?	?	?	+	?	?	?	?	?	?	?	?	?	-	-	-	?	?	?	?	?	?	?	+	+
Psammomys obesus	fat sand rat		?	+	▲	-	-	-	?	?	?	?	?	?	?	?	?	?	?	?	+	?	+	-	?	?	?	?	?	?	-	-	-	?	?	?	?	?	?	?	+	+
Meriones tristrami	jirds		?	+	▲	-	-	-	?	?	?	?	?	?	?	?	?	?	?	?	+	?	+	-	?	?	?	?	?	?	-	-	-	?	?	?	?	?	?	?	?	+
Meriones unguiculatus	mongolian gerbil		+	+	▲	▲	-	-	+	+	+	-	-	-	+	+	+	+	-	-	+	+	+	-	+	+	-	-	-	-	-	-	-	-	+	-	?	?	?	?	?	+	+
Cricetus cricetus	hamster		-	-	-	-	-	-	+	+	+	-	-	-	+	+	+	+	-	-	+	+	+	+	+	+	-	-	-	-	-	?	?	?	+	-	+	-	-	+	+	?	+
Ondatra zibethicus	muskrat		-	-	-	-	-	-	+	+	+	-	-	-	+	+	+	+	-	-	+	+	+	+	?	+	-	-	-	-	-	-	-	-	?	?	?	?	?	?	?	+	+
Microtus socialis	social vole		?	-	-	-	-	-	?	?	?	?	?	?	?	+	+	?	?	?	+	?	+	+	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	+	+
Nannospalax ehrenbergi	mole rat		?	?	?	?	?	?	-	-	-	-	-	-	-	-	+	-	-	-	+	+	+	-	+	+	-	-	-	-	-	-	-	-	+	-	+	-	-	+	+	+	+
Jaculus jaculus	jerboa		-	-	-	-	-	-	?	-	-	-	-	-	?	?	?	-	-	-	-	+	+	-	+	+	-	-	-	-	-	-	-	-	-	-	?	?	?	?	?	-	+
Pedetes capensis	springhare		-	-	-	-	-	-	-	-	?	?	?	?	?	?	?	-	-	-	-	-	-	-	-	+	+	+	-	-	-	-	-	-	-	-	?	?	?	?	?	?	?
Anomalurus sp.	flying squirrel		?	-	-	-	-	-	-	-	-	-	-	-	?	-	+	-	-	-	?	-	-	-	-	+	+	+	-	-	-	-	-	-	?	?	+	-	-	-	-	-	+
Castor fiber	beaver		?	?	?	?	?	?	-	?	?	?	?	?	-	-	+	?	?	?	-	-	-	-	-	+	-	-	-	-	-	?	?	?	?	?	?	?	?	?	-	+
Dipodomys ordii	ord's kangaroo rat	GD	-	?	?	?	?	?	-	-	?	?	?	?	-	-	+	-	-	-	-	-	-	-	-	+	-	-	-	-	-	?	?	?	-	-	?	?	?	?	?	-	+
Dipodomys californicus	agile kangaroo rat		?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	+
Dipodomys merriami	merriams's kanagaro rat		?	-	-	-	-	-	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	+
Heteromys gaumeri	spiny pocket mouse		?	-	-	-	-	-	?	-	?	?	?	?	-	?	?	?	?	?	?	?	?	?	-	+	-	-	-	-	-	?	?	?	?	?	?	?	?	?	?	?	+
Ctenohystrica																																											
Laonastes aenigmamus	laotian rock rat		-	-	-	-	-	+	?	?	-	▲	+	+	-	-	+	?	?	?	-	-	?	?	-	-	-	-	+	-	-	-	-	+	-	-	?	?	?	?	?	?	?
Ctenodactylus gundi	gundi		?	-	-	-	-	+	-	?	-	-	+	+	-	?	?	?	?	?	?	-	?	?	-	-	-	-	+	-	-	-	-	+	-	-	?	?	?	?	?	-	?
Thryonomys swinderianus	greater can rat		?	-	-	-	-	-	-	-	-	-	-	-	?	?	?	-	-	-	?	?	?	?	-	-	-	▲	+	▲	-	?	?	?	-	-	+	-	-	-	-	?	+
Petromys typicus	dassie rat		?	-	-	-	-	-	?	?	-	-	-	-	?	-	+	?	?	?	?	?	-	-	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	+
Bathyergus suillus	nacked mole rat		?	-	-	-	-	-	?	?	?	?	?	?	?	-	+	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	+	-	-	-	-	?	?
Cryptomys hottentotus	hottentot mole rat		?	-	-	-	-	-	?	?	?	?	?	?	?	-	+	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	+	-	-	-	-	?	?
Proechimys cuvieri	cuvier's spiny rat		-	-	+	+	+	-	-	-	-	+	-	-	?	?	?	?	?	?	?	-	-	-	?	-	-	-	+	+	-	+	+	-	-	-	+	-	-	-	-	-	+
Myocastor coypus	nutria		?	-	+	+	+	-	-	-	?	?	?	?	?	?	?	?	?	?	?	-	-	-	-	-	-	-	+	+	-	+	+	-	-	-	+	-	-	-	-	-	+
Capromys pilorides	desmarest's hutia		?	-	+	+	+	-	?	?	?	?	?	?	?	-	+	?	?	?	?	?	?	?	?	?	?	?	?	?	?	+	+	-	?	?	+	-	-	-	-	?	+
Octodon degus	degu		?	-	+	▲	+	-	?	?	?	?	?	?	?	?	?	?	?	?	?	?	-	-	?	-	-	-	+	+	-	-	-	-	?	?	+	-	-	-	-	?	?
Dinomys branickii	pacarana		?	-	▲	-	+	-	?	?	?	?	?	?	?	-	+	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
Coendou prehensilis	brazilian porcupine		?	-	-	-	-	-	?	?	-	+	-	-	?	?	?	?	?	?	?	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	+	-	-	-	-	?	?
Cavia porcellus	guinea pig	GD	-	-	-	-	-	-	-	-	-	+	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	+	+	+	-	-	-	-	+	+	+	-	-	-	+
Dolichotis patagonum	patagonian mara		?	-	-	-	-	-	-	-	-	+	-	-	?	?	?	-	-	-	?	-	?	?	-	-	-	-	+	+	+	-	-	-	-	+	+	+	+	-	-	?	?
Agouti taczanowskii	mountain paca		?	-	-	-	-	-	?	?	?	?	?	?	?	-	+	?	?	?	?	?	?	?	?	?	?	?	?	?	?	-	-	-	?	?	+	+	+	-	-	?	?
Squirrel-related clade																																											
Spermophilus tridecemlineatus	13-lined ground squirrel	GD	-	-	-	-	-	-	-	-	?	?	?	?	-	-	+	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	?	?	+	-	-	-	-	-	+
Marmota marmota	european marmot		?	?	?	?	?	?	?	-	?	?	?	?	?	?	?	-	+	+	?	-	?	?	-	?	?	?	?	?	?	?	?	-	-	+	-	-	-	-	?	?
Sciurus vulgaris	european red squirrel		?	-	-	-	-	-	?	-	?	?	?	?	-	?	?	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	+
Aplodontia rufa	mountain beaver		?	-	-	-	-	-	?	?	?	?	?	?	?	?	?	-	+	-	?	?	?	?	?	?	?	?	?	?	?	-	-	-	?	?	?	?	?	?	?	?	?
Myoxus glis	fat dormouse		?	?	?	?	?	?	-	-	?	?	?	?	?	?	?	-	-	-	-	?	?	?	-	-	-	-	-	-	-	?	?	?	-	-	?	?	?	?	?	?	?
Eliomys quercinus	garden dormouse		?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
Outgroup																																											
Oryctolagus cuniculus	rabbit	GD	-	-	-	-	-	-	-	?	-	-	-	-	?	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Presence (+) and absence (-) patterns. GD denotes sequences supported by available genome data entries. The outgroup for locus 5 is *Lepus europaeus*.

Table 4. Detailed presence/absence patterns (CPAL-*Mus musculus* screening)

			19 IDL Geo	20 PB1 Geo	21	22 PB1 Geo	23 PB1 Geo	24 IDL Geo	25 PB1 Geo	26a B1F	26b B1	27a B1F	27b B1	28 ID2	29 ID4	30 B1	31 ID4	32a ID4	32b ID2	32c ID2	ID4	33	34 ID Str	35 PB1
Mouse-related clade																								
Mus musculus	Mouse	GD	-	-	-	-	-	-	-	-	-	-	-	+	-	-	▲	-	-	+		-	-	+
Rattus norvegicus	Rat	GD	-	-	-	-	-	-	-	-	-	-	-	+	-	-	▲	?	?	?		-	-	?
Acomys russatus	spiny mouse		?	?	?	-	?	?	?	?	?	?	?	+	?	?	?	-	-	+		?	?	+
Gerbillus dasyurus	Gerbil		?	-	-	-	?	?	?	?	?	?	?	+	?	?	?	-	-	+		?	?	+
Psammomys obesus	fat sand rat		?	-	-	-	?	?	?	?	?	?	?	+	?	?	?	-	-	+		?	?	+
Meriones tristrami	Jirds		?	-	-	-	?	-	?	-	-	?	?	?	?	?	?	?	?	?		?	?	+
Meriones unguiculatus	mongolian gerbil		-	-	-	-	-	?	?	?	?	-	-	?	?	-	?	-	-	+		?	?	+
Cricetus cricetus	Hamster		?	?	?	-	?	-	-	?	?	?	?	+	?	?	?	-	-	+		?	?	+
Ondatra zibethicus	Muskrat		-	?	-	-	?	?	-	?	?	?	?	?	?	?	?	-	-	+		?	-	?
Microtus socialis	social vole		?	-	-	-	?	?	?	?	?	?	?	+	?	?	?	-	-	+		?	?	+
Nannospalax ehrenbergi	mole rat		?	-	-	-	?	-	-	?	?	?	?	+	?	?	?	?	?	?		?	?	+
Jaculus jaculus	Jerboa		?	-	-	-	-	?	-	?	?	?	?	?	?	?	?	-	-	+		?	?	?
Pedetes capensis	Springhare		?	?	-	-	?	-	?	?	?	?	?	+	?	?	?	-	-	+		?	?	?
Anomalurus sp.	flying squirrel		?	?	-	-	-	-	-	?	?	?	?	+	?	?	?	-	-	+		?	?	+
Castor fiber	Beaver		-	?	-	-	?	-	-	?	?	-	-	+	?	-	?	-	-	+		?	?	+
Dipodomys ordii	ord's kangaroo rat	GD	+	+	+	+	+	+	+	-	-	-	-	+	-	?	+	-	-	+		-	-	+
Dipodomys californicus	agile kangaroo rat		+	+	+	+	+	?	+	?	?	?	?	?	?	?	?	?	?	?		?	?	?
Dipodomys merriami	merriams's kanagaro rat		+	+	+	+	+	+	+	?	?	?	?	?	?	?	?	?	?	?		?	?	?
Heteromys gaumeri	spiny pocket mouse		-	-	-	+	-	+	+	?	?	?	?	+	?	?	?	?	?	?		?	?	?
Ctenohystrica																								
Laonastes aenigmamus	laotian rock rat		?	?	?	?	?	?	?	-	-	-	-	?	?	?	?	?	?	?		?	-	+
Ctenodactylus gundi	Gundi		?	?	?	?	?	?	?	-	-	-	-	+	?	?	+	-	-	+		?	-	+
Thryonomys swinderianus	greater can rat		?	?	?	?	?	?	?	?	?	+	-	+	-	-	+	-	-	+		?	?	+
Petromys typicus	dassie rat		?	?	?	?	?	?	?	?	?	+	-	+	-	-	+	-	-	+		?	?	+
Bathyergus suillus	nacked mole rat		?	?	?	?	?	?	?	+	+	+	-	?	?	?	?	?	?	?		?	?	?
Cryptomys hottentotus	hottentot mole rat		?	?	?	?	?	?	?	+	+	+	-	?	?	-	+	-	-	+		?	?	+
Proechimys cuvieri	cuvier's spiny rat		?	?	?	?	?	?	?	+	-	+	+	?	-	-	+	-	-	+		?	?	+
Myocastor coypus	Nutria		?	?	?	?	?	?	?	?	?	+	+	+	-	-	+	-	-	+		?	?	+
Capromys pilorides	desmarest's hutia		?	?	?	?	?	?	?	+	-	+	+	+	-	-	+	-	-	+		?	?	+
Octodon degus	Degu		?	?	?	?	-	?	?	+	-	+	-	+	-	-	+	-	-	+		?	?	+
Dinomys branickii	Pacarana		?	?	?	?	?	?	?	+	-	+	-	?	-	-	+	?	?	?		?	?	?
Coendou prehensilis	brazilian porcupine		?	?	?	?	?	?	?	+	-	+	-	?	-	-	+	?	?	?		?	?	?
Cavia porcellus	guinea pig	GD	-	-	-	-	-	-	-	+	-	+	-	+	+	+	+	+	-	+		-	-	+
Dolichotis patagonum	patagonian mara		?	?	?	-	?	?	?	+	-	+	-	?	+	+	+	+	-	+		?	?	+
Agouti taczanowskii	mountain paca		?	?	?	?	?	?	?	+	-	+	-	+	+	+	+	-	-	+		?	?	?
Squirrel-related clade																								
Spermophilus tridecemlineatus	13-lined ground squirrel	GD	?	-	-	-	-	-	?	?	?	-	-	+	?	-	+	-	+	+		+	+	+
Marmota marmota	european marmot		?	?	?	?	?	?	?	?	?	?	?	+	?	?	?	?	?	?		?	+	?
Sciurus vulgaris	european red squirrel		?	?	?	-	-	?	-	?	?	-	-	+	?	-	?	-	+	+		?	+	+
Aplodontia rufa	mountain beaver		?	?	?	?	?	?	?	-	-	?	?	+	?	?	?	-	-	+		+	+	+
Myoxus glis	fat dormouse		?	?	?	?	?	?	?	?	?	?	?	+	?	?	?	-	-	+		-	+	+
Eliomys quercinus	garden dormouse		?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	-	-	+		?	+	+
Outgroup																								
Oryctolagus cuniculus	Rabbit	GD	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	?	?	?		-	-	-

Presence (+) and absence (-) patterns. GD denotes sequences supported by available genome data entries. The outgroup for locus 5 is *Lepus europaeus.*

Summary and Outlook

The MS2-tagged BC1 and BC200 RNA/RNP transgenic mouse models will be very important tools to study tagged RNA dendritic transport (dendritic delivery), as well as the composition and function of RNP particles, initially via their purification. The homozygous transgenic mouse models (BC1MS2B+MS2TAP_BC1 RNA$^{-/-}$ and BC200MS2+MS2TAP) will provide the basis to purify neuronal BC RNPs. Then, proteomics (mass spectrometry) can be used to exhaustively characterize the molecular compositions of these complexes. In analogy to PABP and its involvement in translation (Kondrashov et al., 2005, Muddashetty et al., 2002, Wang et al., 2005, Wang et al., 2002), possible known functions of the newly uncovered proteins bound to the BC RNAs will aid us in making further progress towards understanding the functional significance of BC RNPs.

The second role of MS2-tagged BC RNAs is to observe the tagged RNPs in dendritic transport using antibodies against tagged proteins (e.g., the BC1MS2 RNA and BC200MS2 RNA in combination with MS2TAP fusion proteins with the aid of immunoelectron microscopy). In this work, I have established two important transgenic mouse lines, which will be vital for carrying out studies on BC RNP composition. This approach could also be extended to the purification of any RNA/RNP via tagging. Mice generated with an MS2 domain tagged to any RNA could be crossed with our mouse line expressing the TAP tag fusion protein, and thus make it possible to characterize any RNP assembled *in vivo*.

The present analysis of rodent orthologous SINE data contributes to the solution of rodent phylogeny. SINE insertions are highly stable and provide a molecular fossil record over the course of evolution. Even though previous work showed that SINE insertions and the automated computational strategy (CPAL) provide a reliable source of markers to reconstruct the rodent phylogenetic tree, the most debated branches within the Squirrel-related clade and within the superfamily *Geomyoidea* were not yet successfully addressed with these analyses. Here we developed alternative screening strategies to resolve the highly debated Squirrel-related clade, and also confirmed the monophyly of rodents from SINE insertion analyses. The unresolved interrelationships among the three major clades require further investigation. Our approach shows that its application is capable of resolving the remaining questions in the phylogenetic relationships among rodents, the mammalian order with the most species (about 2277 species with 30-33 families).

References

ABE, R., YAMAMOTO, K. & SAKAMOTO, H. (1996) Target specificity of neuronal RNA-binding protein, Mel-N1: direct binding to the 3' untranslated region of its own mRNA. *Nucleic Acids Res,* 24**,** 2011-6.

ADKINS, R. M., GELKE, E. L., ROWE, D. & HONEYCUTT, R. L. (2001) Molecular phylogeny and divergence time estimates for major rodent groups: evidence from multiple genes. *Mol Biol Evol,* 18**,** 777-91.

ADKINS, R. M., WALTON, A. H. & HONEYCUTT, R. L. (2003) Higher-level systematics of rodents and divergence time estimates based on two congruent nuclear genes. *Mol Phylogenet Evol,* 26**,** 409-20.

BACHLER, M., SCHROEDER, R. & VON AHSEN, U. (1999) StreptoTag: a novel method for the isolation of RNA-binding proteins. *RNA,* 5**,** 1509-16.

BARREAU, C., PAILLARD, L. & OSBORNE, H. B. (2005) AU-rich elements and associated factors: are there unifying principles? *Nucleic Acids Res,* 33**,** 7138-50.

BASILE, V., VICENTE, A., MARTIGNETTI, J. A., SKRYABIN, B. V., BROSIUS, J. & KENNEDY, J. L. (1998) Assignment of the human BC200 RNA gene (BCYRN1) to chromosome 2p16 by radiation hybrid mapping. *Cytogenet Cell Genet,* 82**,** 271-2.

BERTRAND, E., CHARTRAND, P., SCHAEFER, M., SHENOY, S. M., SINGER, R. H. & LONG, R. M. (1998) Localization of ASH1 mRNA particles in living yeast. *Mol Cell,* 2**,** 437-45.

BLICHENBERG, A., SCHWANKE, B., REHBEIN, M., GARNER, C. C., RICHTER, D. & KINDLER, S. (1999) Identification of a cis-acting dendritic targeting element in MAP2 mRNAs. *J Neurosci,* 19**,** 8818-29.

BLOMMEL, P. G. & FOX, B. G. (2007) A combined approach to improving large-scale production of tobacco etch virus protease. *Protein Expr Purif,* 55**,** 53-68.

BOLOGNANI, F., MERHEGE, M. A., TWISS, J. & PERRONE-BIZZOZERO, N. I. (2004) Dendritic localization of the RNA-binding protein HuD in hippocampal neurons: association with polysomes and upregulation during contextual learning. *Neurosci Lett,* 371**,** 152-7.

BORODULINA, O. R. & KRAMEROV, D. A. (2001) Short interspersed elements (SINEs) from insectivores. Two classes of mammalian SINEs distinguished by A-rich tail structure. *Mamm Genome,* 12**,** 779-86.

BOVIA, F. & STRUB, K. (1996) The signal recognition particle and related small cytoplasmic ribonucleoprotein particles. *J Cell Sci,* 109 (Pt 11)**,** 2601-8.

BOVIA, F., WOLFF, N., RYSER, S. & STRUB, K. (1997) The SRP9/14 subunit of the human signal recognition particle binds to a variety of Alu-like RNAs and with higher affinity than its mouse homolog. *Nucleic Acids Res,* 25**,** 318-26.

BROSIUS, J. (1999) RNAs from all categories generate retrosequences that may be exapted as novel genes or regulatory elements. *Gene,* 238**,** 115-34.

BROSIUS, J. & TIEDGE, H. (1995) Reverse transcriptase: mediator of genomic plasticity. *Virus Genes,* 11**,** 163-79.

BUI, N. & STRUB, K. (1999) New insights into signal recognition and elongation arrest activities of the signal recognition particle. *Biol Chem,* 380**,** 135-45.

CAO, X., YEO, G., MUOTRI, A. R., KUWABARA, T. & GAGE, F. H. (2006) Noncoding RNAs in the Mammalian Central Nervous System. *Annu Rev Neurosci*.

CAO, Y., ADACHI, J., YANO, T. & HASEGAWA, M. (1994) Phylogenetic place of guinea pigs: no support of the rodent-polyphyly hypothesis from maximum-likelihood analyses of multiple protein sequences. *Mol Biol Evol,* 11**,** 593-604.

CHAGNOVICH, D., FAYOS, B. E. & COHN, S. L. (1996) Differential activity of ELAV-like RNA-binding proteins in human neuroblastoma. *J Biol Chem,* 271**,** 33587-91.

CHAO, J. A., PATSKOVSKY, Y., ALMO, S. C. & SINGER, R. H. (2008) Structural basis for the coevolution of a viral RNA-protein complex. *Nat Struct Mol Biol,* 15**,** 103-5.

CHEN, W., BOCKER, W., BROSIUS, J. & TIEDGE, H. (1997a) Expression of neural BC200 RNA in human tumours. *J Pathol,* 183**,** 345-51.

CHEN, W., HEIERHORST, J., BROSIUS, J. & TIEDGE, H. (1997b) Expression of neural BC1 RNA: induction in murine tumours. *Eur J Cancer,* 33**,** 288-92.

CHENG, J. G., TIEDGE, H. & BROSIUS, J. (1996) Identification and characterization of BC1 RNP particles. *DNA Cell Biol,* 15**,** 549-59.

CHENG, J. G., TIEDGE, H. & BROSIUS, J. (1997) Expression of dendritic BC200 RNA, component of a 11.4S ribonucleoprotein particle, is conserved in humans and simians. *Neurosci Lett,* 224**,** 206-10.

CHICUREL, M. E., TERRIAN, D. M. & POTTER, H. (1993) mRNA at the synapse: analysis of a synaptosomal preparation enriched in hippocampal dendritic spines. *J Neurosci,* 13**,** 4054-63.

CHURAKOV, G. A. & NOVIKOV, S. N. (2000) Heterogeneity and differential expression of MUP proteins as a genetic basis of the physiological activity of androgen-dependent pheromones. *Dokl Biol Sci,* 375**,** 549-52.

CUADRADO, A., NAVARRO-YUBERO, C., FURNEAUX, H., KINTER, J., SONDEREGGER, P. & MUNOZ, A. (2002) HuD binds to three AU-rich sequences in the 3'-UTR of neuroserpin mRNA and promotes the accumulation of neuroserpin mRNA and protein. *Nucleic Acids Res,* 30**,** 2202-11.

D'ERCHIA, A. M., GISSI, C., PESOLE, G., SACCONE, C. & ARNASON, U. (1996) The guinea-pig is not a rodent. *Nature,* 381**,** 597-600.

DANIELS, G. R. & DEININGER, P. L. (1985) Repeat sequence families derived from mammalian tRNA genes. *Nature,* 317**,** 819-22.

DAVIS, B. P. & MACDONALD, R. J. (1988) Limited transcription of rat elastase I transgene repeats in transgenic mice. *Genes Dev,* 2**,** 13-22.

DE HARVEN, E., LEUNG, R. & CHRISTENSEN, H. (1984) A novel approach for scanning electron microscopy of colloidal gold-labeled cell surfaces. *J Cell Biol,* 99**,** 53-7.

DEBRY, R. W. (2003) Identifying conflicting signal in a multigene analysis reveals a highly resolved tree: the phylogeny of Rodentia (Mammalia). *Syst Biol,* 52**,** 604-17.

DEBRY, R. W. & SAGEL, R. M. (2001) Phylogeny of rodentia (Mammalia) inferred from the nuclear-encoded gene IRBP. *Mol Phylogenet Evol,* 19**,** 290-301.

DECHIARA, T. M. & BROSIUS, J. (1987) Neural BC1 RNA: cDNA clones reveal nonrepetitive sequence content. *Proc Natl Acad Sci U S A,* 84**,** 2624-8.

DEININGER, P. L., BATZER, M. A., HUTCHISON, C. A., 3RD & EDGELL, M. H. (1992) Master genes in mammalian repetitive DNA amplification. *Trends Genet,* 8**,** 307-11.

DREYFUSS, G., CHOI, Y. D. & ADAM, S. A. (1984) Characterization of heterogeneous nuclear RNA-protein complexes in vivo with monoclonal antibodies. *Mol Cell Biol,* 4**,** 1104-14.

DUMPELMANN, E., MITTENDORF, H. & BENECKE, B. J. (2003) Efficient transcription of the EBER2 gene depends on the structural integrity of the RNA. *RNA,* 9**,** 432-42.

FARWICK, A., JORDAN, U., FUELLEN, G., HUCHON, D., CATZEFLIS, F., BROSIUS, J. & SCHMITZ, J. (2006) Automated scanning for phylogenetically informative transposed elements in rodents. *Syst Biol,* 55**,** 936-48.

GARNER, C. C., TUCKER, R. P. & MATUS, A. (1988) Selective localization of messenger RNA for cytoskeletal protein MAP2 in dendrites. *Nature,* 336**,** 674-7.

GARRICK, D., FIERING, S., MARTIN, D. I. & WHITELAW, E. (1998) Repeat-induced gene silencing in mammals. *Nat Genet,* 18**,** 56-9.

GIBBS, R. A., WEINSTOCK, G. M., METZKER, M. L., MUZNY, D. M., SODERGREN, E. J., SCHERER, S., SCOTT, G., STEFFEN, D., WORLEY, K. C., BURCH, P. E., OKWUONU, G., HINES, S., LEWIS, L., DERAMO, C., DELGADO, O., DUGAN-ROCHA, S., MINER, G., MORGAN, M., HAWES, A., GILL, R., CELERA, HOLT, R. A., ADAMS, M. D., AMANATIDES, P. G., BADEN-TILLSON, H., BARNSTEAD, M., CHIN, S., EVANS, C. A., FERRIERA, S., FOSLER, C., GLODEK, A., GU, Z., JENNINGS, D., KRAFT, C. L., NGUYEN, T., PFANNKOCH, C. M., SITTER, C., SUTTON, G. G., VENTER, J. C., WOODAGE, T., SMITH, D., LEE, H. M., GUSTAFSON, E., CAHILL, P., KANA, A., DOUCETTE-STAMM, L., WEINSTOCK, K., FECHTEL, K., WEISS, R. B., DUNN, D. M., GREEN, E. D., BLAKESLEY, R. W., BOUFFARD, G. G., DE JONG, P. J., OSOEGAWA, K., ZHU, B., MARRA, M., SCHEIN, J., BOSDET, I., FJELL, C., JONES, S., KRZYWINSKI, M., MATHEWSON, C., SIDDIQUI, A., WYE, N., MCPHERSON, J., ZHAO, S., FRASER, C. M., SHETTY, J., SHATSMAN, S., GEER, K., CHEN,

Y., ABRAMZON, S., NIERMAN, W. C., HAVLAK, P. H., CHEN, R., DURBIN, K. J., EGAN, A., REN, Y., SONG, X. Z., LI, B., LIU, Y., QIN, X., CAWLEY, S., COONEY, A. J., D'SOUZA, L. M., MARTIN, K., WU, J. Q., GONZALEZ-GARAY, M. L., JACKSON, A. R., KALAFUS, K. J., MCLEOD, M. P., MILOSAVLJEVIC, A., VIRK, D., VOLKOV, A., WHEELER, D. A., ZHANG, Z., BAILEY, J. A., EICHLER, E. E., TUZUN, E., et al. (2004) Genome sequence of the Brown Norway rat yields insights into mammalian evolution. *Nature,* 428**,** 493-521.

GOGOLEVSKAYA, I. K., KOVAL, A. P. & KRAMEROV, D. A. (2005) Evolutionary history of 4.5SH RNA. *Mol Biol Evol,* 22**,** 1546-54.

GOGOLEVSKY, K. P., VASSETZKY, N. S. & KRAMEROV, D. A. (2008) Bov-B-mobilized SINEs in vertebrate genomes. *Gene,* 407**,** 75-85.

GOODIER, J. L. & MARAIA, R. J. (1998) Terminator-specific recycling of a B1-Alu transcription complex by RNA polymerase III is mediated by the RNA terminus-binding protein La. *J Biol Chem,* 273**,** 26110-6.

GRAUR, D., HIDE, W. A. & LI, W. H. (1991) Is the guinea-pig a rodent? *Nature,* 351**,** 649-52.

HARTENBERGER, J.-L. (1998) *The Cenozoic radiation of mammals : some comments on the shape and tempo of a major evolutionary event,* Milano, Società italiana di scienze naturali : Museo civico di storia naturale di Milano.

HARTMUTH, K., VORNLOCHER, H. P. & LUHRMANN, R. (2004) Tobramycin affinity tag purification of spliceosomes. *Methods Mol Biol,* 257**,** 47-64.

HASLER, J. & STRUB, K. (2006) Alu RNP and Alu RNA regulate translation initiation in vitro. *Nucleic Acids Res,* 34**,** 2374-85.

HINMAN, M. N. & LOU, H. (2008) Diverse molecular functions of Hu proteins. *Cell Mol Life Sci,* 65**,** 3168-81.

HOCHULI, E. (1990) Purification of recombinant proteins with metal chelate adsorbent. *Genet Eng (N Y),* 12**,** 87-98.

HOGAN, B., COSTANTINI, F. & LACY, E. (1986) *Manipulating the mouse embryo : a laboratory manual,* Cold Spring Harbor, N.Y., Cold Spring Harbor Laboratory.

HOGG, J. R. & COLLINS, K. (2007) RNA-based affinity purification reveals 7SK RNPs with distinct composition and regulation. *RNA,* 13**,** 868-80.

HORN, W. T., CONVERY, M. A., STONEHOUSE, N. J., ADAMS, C. J., LILJAS, L., PHILLIPS, S. E. & STOCKLEY, P. G. (2004) The crystal structure of a high affinity RNA stem-loop complexed with the bacteriophage MS2 capsid: further challenges in the modeling of ligand-RNA interactions. *RNA,* 10**,** 1776-82.

HOYER, L. C., LEE, J. C. & BUCANA, C. (1979) Scanning immunoelectron microscopy for the identification and mapping of two or more antigens on cell surfaces. *Scan Electron Microsc***,** 629-36.

HUCHON, D., CHEVRET, P., JORDAN, U., KILPATRICK, C. W., RANWEZ, V., JENKINS, P. D., BROSIUS, J. & SCHMITZ, J. (2007) Multiple molecular evidences for a living mammalian fossil. *Proc Natl Acad Sci U S A,* 104**,** 7495-9.

HUCHON, D., MADSEN, O., SIBBALD, M. J., AMENT, K., STANHOPE, M. J., CATZEFLIS, F., DE JONG, W. W. & DOUZERY, E. J. (2002) Rodent phylogeny and a timescale for the evolution of Glires: evidence from an extensive taxon sampling using three nuclear genes. *Mol Biol Evol,* 19**,** 1053-65.

IACOANGELI, A., LIN, Y., MORLEY, E. J., MUSLIMOV, I. A., BIANCHI, R., REILLY, J., WEEDON, J., DIALLO, R., BOCKER, W. & TIEDGE, H. (2004) BC200 RNA in invasive and preinvasive breast cancer. *Carcinogenesis,* 25**,** 2125-33.

IACOANGELI, A., ROZHDESTVENSKY, T. S., DOLZHANSKAYA, N., TOURNIER, B., SCHUTT, J., BROSIUS, J., DENMAN, R. B., KHANDJIAN, E. W., KINDLER, S. & TIEDGE, H. (2008a) On BC1 RNA and the fragile X mental retardation protein. *Proc Natl Acad Sci U S A,* 105**,** 734-9.

IACOANGELI, A., ROZHDESTVENSKY, T. S., DOLZHANSKAYA, N., TOURNIER, B., SCHUTT, J., BROSIUS, J., DENMAN, R. B., KHANDJIAN, E. W., KINDLER, S. & TIEDGE, H. (2008b) Reply to Bagni: On BC1 RNA and the fragile X mental retardation protein. *Proc Natl Acad Sci U S A,* 105**,** E29.

JOHANSSON, H. E., DERTINGER, D., LECUYER, K. A., BEHLEN, L. S., GREEF, C. H. & UHLENBECK, O. C. (1998) A thermodynamic analysis of the sequence-specific binding of RNA by bacteriophage MS2 coat protein. *Proc Natl Acad Sci U S A,* 95**,** 9244-9.

KAGANMAN, I. (2008) Another handle for RNA. *Nature Methods,* 5**,** 126-126.

KHANAM, T., MUDDASHETTY, R. S., KAHVEJIAN, A., SONENBERG, N. & BROSIUS, J. (2006) Poly(A)-binding protein binds to A-rich sequences via RNA-binding domains 1+2 and 3+4. *RNA Biol,* 3**,** 170-7.

KHANAM, T., RAABE, C. A., KIEFMANN, M., HANDEL, S., SKRYABIN, B. V. & BROSIUS, J. (2007a) Can ID repetitive elements serve as cis-acting dendritic targeting elements? An in vivo study. *PLoS ONE,* 2**,** e961.

KHANAM, T., ROZHDESTVENSKY, T. S., BUNDMAN, M., GALIVETI, C. R., HANDEL, S., SUKONINA, V., JORDAN, U., BROSIUS, J. & SKRYABIN, B. V. (2007b) Two primate-specific small non-protein-coding RNAs in transgenic mice: neuronal expression, subcellular localization and binding partners. *Nucleic Acids Res,* 35**,** 529-39.

KIM, J., MARTIGNETTI, J. A., SHEN, M. R., BROSIUS, J. & DEININGER, P. (1994) Rodent BC1 RNA gene as a master gene for ID element amplification. *Proc Natl Acad Sci U S A,* 91**,** 3607-11.

KLEIMAN, R., BANKER, G. & STEWARD, O. (1990) Differential subcellular localization of particular mRNAs in hippocampal neurons in culture. *Neuron,* 5**,** 821-30.

KO, B. & GUNDERSON, S. I. (2002) Identification of new poly(A) polymerase-inhibitory proteins capable of regulating pre-mRNA polyadenylation. *J Mol Biol,* 318**,** 1189-206.

KONDRASHOV, A. V., KIEFMANN, M., EBNET, K., KHANAM, T., MUDDASHETTY, R. S. & BROSIUS, J. (2005) Inhibitory effect of naked neural BC1 RNA or BC200 RNA on eukaryotic in vitro translation systems is reversed by poly(A)-binding protein (PABP). *J Mol Biol,* 353**,** 88-103.

KOZAK, M. (1987) An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res,* 15**,** 8125-48.

KRAMEROV, D. A. (1999) [The evidence of close phylogenetic relationship between Gliridae and Sciuridae families based on the study of short retroposon B1-dID]. *Dokl Akad Nauk,* 364**,** 277-80.

KRAMEROV, D. A. & VASSETZKY, N. S. (2001) Structure and origin of a novel dimeric retroposon B1-diD. *J Mol Evol,* 52**,** 137-43.

KRAMEROV, D. A. & VASSETZKY, N. S. (2005) Short retroposons in eukaryotic genomes. *Int Rev Cytol,* 247**,** 165-221.

KRAYEV, A. S., MARKUSHEVA, T. V., KRAMEROV, D. A., RYSKOV, A. P., SKRYABIN, K. G., BAYEV, A. A. & GEORGIEV, G. P. (1982) Ubiquitous transposon-like repeats B1 and B2 of the mouse genome: B2 sequencing. *Nucleic Acids Res,* 10**,** 7461-75.

KREMERSKOTHEN, J., NETTERMANN, M., OP DE BEKKE, A., BACHMANN, M. & BROSIUS, J. (1998a) Identification of human autoantigen La/SS-B as BC1/BC200 RNA-binding protein. *DNA Cell Biol,* 17**,** 751-9.

KREMERSKOTHEN, J., ZOPF, D., WALTER, P., CHENG, J. G., NETTERMANN, M., NIEWERTH, U., MARAIA, R. J. & BROSIUS, J. (1998b) Heterodimer SRP9/14 is an integral part of the neural BC200 RNP in primate brain. *Neurosci Lett,* 245**,** 123-6.

KRIEGS, J. O., CHURAKOV, G., JURKA, J., BROSIUS, J. & SCHMITZ, J. (2007) Evolutionary history of 7SL RNA-derived SINEs in Supraprimates. *Trends Genet,* 23**,** 158-61.

KRIEGS, J. O., CHURAKOV, G., KIEFMANN, M., JORDAN, U., BROSIUS, J. & SCHMITZ, J. (2006) Retroposed elements as archives for the evolutionary history of placental mammals. *PLoS Biol,* 4**,** e91.

KURYSHEV, V. Y., SKRYABIN, B. V., KREMERSKOTHEN, J., JURKA, J. & BROSIUS, J. (2001) Birth of a gene: locus of neuronal BC200 snmRNA in three prosimians and human BC200 pseudogenes as archives of change in the Anthropoidea lineage. *J Mol Biol,* 309**,** 1049-66.

LAIRD, P. W., ZIJDERVELD, A., LINDERS, K., RUDNICKI, M. A., JAENISCH, R. & BERNS, A. (1991) Simplified mammalian DNA isolation procedure. *Nucleic Acids Res,* 19**,** 4293.

LANDER, E. S., LINTON, L. M., BIRREN, B., NUSBAUM, C., ZODY, M. C., BALDWIN, J., DEVON, K., DEWAR, K., DOYLE, M., FITZHUGH, W., FUNKE, R., GAGE, D., HARRIS, K., HEAFORD, A., HOWLAND, J., KANN, L., LEHOCZKY, J., LEVINE, R., MCEWAN, P., MCKERNAN, K.,

MELDRIM, J., MESIROV, J. P., MIRANDA, C., MORRIS, W., NAYLOR, J., RAYMOND, C., ROSETTI, M., SANTOS, R., SHERIDAN, A., SOUGNEZ, C., STANGE-THOMANN, N., STOJANOVIC, N., SUBRAMANIAN, A., WYMAN, D., ROGERS, J., SULSTON, J., AINSCOUGH, R., BECK, S., BENTLEY, D., BURTON, J., CLEE, C., CARTER, N., COULSON, A., DEADMAN, R., DELOUKAS, P., DUNHAM, A., DUNHAM, I., DURBIN, R., FRENCH, L., GRAFHAM, D., GREGORY, S., HUBBARD, T., HUMPHRAY, S., HUNT, A., JONES, M., LLOYD, C., MCMURRAY, A., MATTHEWS, L., MERCER, S., MILNE, S., MULLIKIN, J. C., MUNGALL, A., PLUMB, R., ROSS, M., SHOWNKEEN, R., SIMS, S., WATERSTON, R. H., WILSON, R. K., HILLIER, L. W., MCPHERSON, J. D., MARRA, M. A., MARDIS, E. R., FULTON, L. A., CHINWALLA, A. T., PEPIN, K. H., GISH, W. R., CHISSOE, S. L., WENDL, M. C., DELEHAUNTY, K. D., MINER, T. L., DELEHAUNTY, A., KRAMER, J. B., COOK, L. L., FULTON, R. S., JOHNSON, D. L., MINX, P. J., CLIFTON, S. W., HAWKINS, T., BRANSCOMB, E., PREDKI, P., RICHARDSON, P., WENNING, S., SLEZAK, T., DOGGETT, N., CHENG, J. F., OLSEN, A., LUCAS, S., ELKIN, C., UBERBACHER, E., FRAZIER, M., et al. (2001) Initial sequencing and analysis of the human genome. *Nature,* 409**,** 860-921.

LEE, I. Y., WESTAWAY, D., SMIT, A. F., WANG, K., SETO, J., CHEN, L., ACHARYA, C., ANKENER, M., BASKIN, D., COOPER, C., YAO, H., PRUSINER, S. B. & HOOD, L. E. (1998) Complete genomic sequence and analysis of the prion protein gene region from three mammalian species. *Genome Res,* 8**,** 1022-37.

LEVINE, T. D., GAO, F., KING, P. H., ANDREWS, L. G. & KEENE, J. D. (1993) Hel-N1: an autoimmune RNA-binding protein with specificity for 3' uridylate-rich untranslated regions of growth factor mRNAs. *Mol Cell Biol,* 13**,** 3494-504.

LIM, F. & PEABODY, D. S. (2002) RNA recognition site of PP7 coat protein. *Nucleic Acids Res,* 30**,** 4138-44.

LIN, Y., BROSIUS, J. & TIEDGE, H. (2001) Neuronal BC1 RNA: co-expression with growth-associated protein-43 messenger RNA. *Neuroscience,* 103**,** 465-79.

LUDWIG, A., ROZHDESTVENSKY, T. S., KURYSHEV, V. Y., SCHMITZ, J. & BROSIUS, J. (2005) An unusual primate locus that attracted two independent Alu insertions and facilitates their transcription. *J Mol Biol,* 350**,** 200-14.

LYFORD, G. L., YAMAGATA, K., KAUFMANN, W. E., BARNES, C. A., SANDERS, L. K., COPELAND, N. G., GILBERT, D. J., JENKINS, N. A., LANAHAN, A. A. & WORLEY, P. F. (1995) Arc, a growth factor and activity-regulated gene, encodes a novel cytoskeleton-associated protein that is enriched in neuronal dendrites. *Neuron,* 14**,** 433-45.

MANIATIS, T., FRITSCH, E. F. & SAMBROOK, J. (1982) *Molecular cloning : a laboratory manual,* Cold Spring Harbor, N.Y., Cold Spring Harbor Laboratory.

MARTIGNETTI, J. A. & BROSIUS, J. (1993a) BC200 RNA: a neural RNA polymerase III product encoded by a monomeric Alu element. *Proc Natl Acad Sci U S A,* 90**,** 11563-7.

MARTIGNETTI, J. A. & BROSIUS, J. (1993b) Neural BC1 RNA as an evolutionary marker: guinea pig remains a rodent. *Proc Natl Acad Sci U S A,* 90**,** 9698-702.

MARTIGNETTI, J. A. & BROSIUS, J. (1995) BC1 RNA: transcriptional analysis of a neural cell-specific RNA polymerase III transcript. *Mol Cell Biol,* 15**,** 1642-50.

MCKINNON, R. D., DANIELSON, P., BROW, M. A., BLOOM, F. E. & SUTCLIFFE, J. G. (1987) Expression of small cytoplasmic transcripts of the rat identifier element in vivo and in cultured cells. *Mol Cell Biol,* 7**,** 2148-54.

MIYASHIRO, K., DICHTER, M. & EBERWINE, J. (1994) On the nature and differential distribution of mRNAs in hippocampal neurites: implications for neuronal functioning. *Proc Natl Acad Sci U S A,* 91**,** 10800-4.

MOBARAK, C. D., ANDERSON, K. D., MORIN, M., BECKEL-MITCHENER, A., ROGERS, S. L., FURNEAUX, H., KING, P. & PERRONE-BIZZOZERO, N. I. (2000) The RNA-binding protein HuD is required for GAP-43 mRNA stability, GAP-43 gene expression, and PKC-dependent neurite outgrowth in PC12 cells. *Mol Biol Cell,* 11**,** 3191-203.

MONTGELARD, C., BENTZ, S., TIRARD, C., VERNEAU, O. & CATZEFLIS, F. M. (2002) Molecular systematics of sciurognathi (rodentia): the mitochondrial cytochrome b and 12S rRNA genes support the Anomaluroidea (Pedetidae and Anomaluridae). *Mol Phylogenet Evol,* 22**,** 220-33.

MONTGELARD, C., FORTY, E., ARNAL, V. & MATTHEE, C. A. (2008) Suprafamilial relationships among Rodentia and the phylogenetic effect of removing fast-evolving nucleotides in mitochondrial, exon and intron fragments. *BMC Evol Biol,* 8**,** 321.

MUDDASHETTY, R., KHANAM, T., KONDRASHOV, A., BUNDMAN, M., IACOANGELI, A., KREMERSKOTHEN, J., DUNING, K., BARNEKOW, A., HUTTENHOFER, A., TIEDGE, H. & BROSIUS, J. (2002) Poly(A)-binding protein is associated with neuronal BC1 and BC200 ribonucleoprotein particles. *J Mol Biol,* 321**,** 433-45.

MULLIN, C., DUNING, K., BARNEKOW, A., RICHTER, D., KREMERSKOTHEN, J. & MOHR, E. (2004) Interaction of rat poly(A)-binding protein with poly(A)- and non-poly(A) sequences is preferentially mediated by RNA recognition motifs 3+4. *FEBS Lett,* 576**,** 437-41.

MURPHY, W. J., EIZIRIK, E., JOHNSON, W. E., ZHANG, Y. P., RYDER, O. A. & O'BRIEN, S. J. (2001) Molecular phylogenetics and the origins of placental mammals. *Nature,* 409**,** 614-8.

MUS, E., HOF, P. R. & TIEDGE, H. (2007) Dendritic BC200 RNA in aging and in Alzheimer's disease. *Proc Natl Acad Sci U S A,* 104**,** 10679-84.

MUSLIMOV, I. A., IACOANGELI, A., BROSIUS, J. & TIEDGE, H. (2006) Spatial codes in dendritic BC1 RNA. *J Cell Biol,* 175**,** 427-39.

MUSLIMOV, I. A., SANTI, E., HOMEL, P., PERINI, S., HIGGINS, D. & TIEDGE, H. (1997) RNA transport in dendrites: a cis-acting targeting element is contained within neuronal BC1 RNA. *J Neurosci,* 17**,** 4722-33.

MUSLIMOV, I. A., TITMUS, M., KOENIG, E. & TIEDGE, H. (2002) Transport of Neuronal BC1 RNA in Mauthner Axons. *J Neurosci,* 22**,** 4293-301.

NAPOLI, I., MERCALDO, V., BOYL, P. P., ELEUTERI, B., ZALFA, F., DE RUBEIS, S., DI MARINO, D., MOHR, E., MASSIMI, M., FALCONI, M., WITKE, W., COSTA-MATTIOLI, M., SONENBERG, N., ACHSEL, T. & BAGNI, C. (2008) The fragile X syndrome protein represses activity-dependent translation through CYFIP1, a new 4E-BP. *Cell,* 134**,** 1042-54.

OHSHIMA, K. & OKADA, N. (1994) Generality of the tRNA origin of short interspersed repetitive elements (SINEs). Characterization of three different tRNA-derived retroposons in the octopus. *J Mol Biol,* 243**,** 25-37.

OHSHIMA, K. & OKADA, N. (2005) SINEs and LINEs: symbionts of eukaryotic genomes with a common tail. *Cytogenet Genome Res,* 110**,** 475-90.

OKADA, N. (1991) SINEs. *Curr Opin Genet Dev,* 1**,** 498-504.

OKANO, H. J. & DARNELL, R. B. (1997) A hierarchy of Hu RNA binding proteins in developing and adult neurons. *J Neurosci,* 17**,** 3024-37.

PEABODY, D. S. (1990) Translational repression by bacteriophage MS2 coat protein expressed from a plasmid. A system for genetic analysis of a protein-RNA interaction. *J Biol Chem,* 265**,** 5684-9.

QUENTIN, Y. (1992) Origin of the Alu family: a family of Alu-like monomers gave birth to the left and the right arms of the Alu elements. *Nucleic Acids Res,* 20**,** 3397-401.

RAO, A. & STEWARD, O. (1993) Evaluation of RNAs present in synaptodendrosomes: dendritic, glial, and neuronal cell body contribution. *J Neurochem,* 61**,** 835-44.

REYES, A., GISSI, C., PESOLE, G., CATZEFLIS, F. M. & SACCONE, C. (2000) Where do rodents fit? Evidence from the complete mitochondrial genome of Sciurus vulgaris. *Mol Biol Evol,* 17**,** 979-83.

REYES, A., PESOLE, G. & SACCONE, C. (1998) Complete mitochondrial DNA sequence of the fat dormouse, Glis glis: further evidence of rodent paraphyly. *Mol Biol Evol,* 15**,** 499-505.

RIGAUT, G., SHEVCHENKO, A., RUTZ, B., WILM, M., MANN, M. & SERAPHIN, B. (1999) A generic protein purification method for protein complex characterization and proteome exploration. *Nat Biotechnol,* 17**,** 1030-2.

ROGOZIN, I. B., MAYOROV, V. I., LAVRENTIEVA, M. V., MILANESI, L. & ADKISON, L. R. (2000) Prediction and phylogenetic analysis of mammalian short interspersed elements (SINEs). *Brief Bioinform,* 1**,** 260-74.

ROZHDESTVENSKY, T. S., KOPYLOV, A. M., BROSIUS, J. & HUTTENHOFER, A. (2001) Neuronal BC1 RNA structure: evolutionary conversion of a tRNA(Ala) domain into an extended stem-loop structure. *RNA,* 7**,** 722-30.

SALEM, A. H., RAY, D. A., XING, J., CALLINAN, P. A., MYERS, J. S., HEDGES, D. J., GARBER, R. K., WITHERSPOON, D. J., JORDE, L. B. & BATZER, M. A. (2003) Alu elements and hominid phylogenetics. *Proc Natl Acad Sci U S A,* 100**,** 12787-91.

SCHMITZ, J., OHME, M. & ZISCHLER, H. (2001) SINE insertions in cladistic analyses and the phylogenetic affiliations of Tarsius bancanus to other primates. *Genetics,* 157**,** 777-84.

SCHMITZ, J., ROOS, C. & ZISCHLER, H. (2005) Primate phylogeny: molecular evidence from retroposons. *Cytogenet Genome Res,* 108**,** 26-37.

SERDOBOVA, I. M. & KRAMEROV, D. A. (1993) [Use of the short retroposon B2 in the study of phylogenetic relationship in rodents]. *Genetika,* 29**,** 1969-81.

SERDOBOVA, I. M. & KRAMEROV, D. A. (1998) Short retroposons of the B2 superfamily: evolution and application for the study of rodent phylogeny. *J Mol Evol,* 46**,** 202-14.

SHEDLOCK, A. M. & OKADA, N. (2000) SINE insertions: powerful tools for molecular systematics. *Bioessays,* 22**,** 148-60.

SHEDLOCK, A. M., TAKAHASHI, K. & OKADA, N. (2004) SINEs of speciation: tracking lineages with retroposons. *Trends Ecol Evol,* 19**,** 545-53.

SHEN, M. R., BROSIUS, J. & DEININGER, P. L. (1997) BC1 RNA, the transcript from a master gene for ID element amplification, is able to prime its own reverse transcription. *Nucleic Acids Res,* 25**,** 1641-8.

SINGER, M. F. (1982) SINEs and LINEs: highly repeated short and long interspersed sequences in mammalian genomes. *Cell,* 28**,** 433-4.

SINGER, M. F. (1990) SINE and LINE nomenclature. *Trends Genet,* 6**,** 204.

SKRYABIN, B. V., KREMERSKOTHEN, J., VASSILACOPOULOU, D., DISOTELL, T. R., KAPITONOV, V. V., JURKA, J. & BROSIUS, J. (1998) The BC200 RNA gene and its neural expression are conserved in Anthropoidea (Primates). *J Mol Evol,* 47**,** 677-85.

SKRYABIN, B. V., SUKONINA, V., JORDAN, U., LEWEJOHANN, L., SACHSER, N., MUSLIMOV, I., TIEDGE, H. & BROSIUS, J. (2003) Neuronal untranslated BC1 RNA: targeted gene elimination in mice. *Mol Cell Biol,* 23**,** 6435-41.

SRISAWAT, C. & ENGELKE, D. R. (2001) Streptavidin aptamers: affinity tags for the study of RNAs and ribonucleoproteins. *RNA,* 7**,** 632-41.

SRISAWAT, C., GOLDSTEIN, I. J. & ENGELKE, D. R. (2001) Sephadex-binding RNA ligands: rapid affinity purification of RNA from complex RNA mixtures. *Nucleic Acids Res,* 29**,** E4.

TARS, K., FRIDBORG, K., BUNDULE, M. & LILJAS, L. (2000) The three-dimensional structure of bacteriophage PP7 from Pseudomonas aeruginosa at 3.7-A resolution. *Virology,* 272**,** 331-7.

TAYLOR, B. A., NAVIN, A., SKRYABIN, B. V. & BROSIUS, J. (1997) Localization of the mouse gene (Bc1) encoding neural BC1 RNA near the fibroblast growth factor 3 locus (Fgf3) on distal chromosome 7. *Genomics,* 44**,** 153-4.

THOMPSON, J. D., GIBSON, T. J., PLEWNIAK, F., JEANMOUGIN, F. & HIGGINS, D. G. (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res,* 25**,** 4876-82.

TIEDGE, H. & BROSIUS, J. (1996) Translational machinery in dendrites of hippocampal neurons in culture. *J Neurosci,* 16**,** 7171-81.

TIEDGE, H., CHEN, W. & BROSIUS, J. (1993) Primary structure, neural-specific expression, and dendritic location of human BC200 RNA. *J Neurosci,* 13**,** 2382-90.

TIEDGE, H., FREMEAU, R. T., JR., WEINSTOCK, P. H., ARANCIO, O. & BROSIUS, J. (1991) Dendritic location of neural BC1 RNA. *Proc Natl Acad Sci U S A,* 88**,** 2093-7.

ULLU, E., MURPHY, S. & MELLI, M. (1982) Human 7SL RNA consists of a 140 nucleotide middle-repetitive sequence inserted in an alu sequence. *Cell,* 29**,** 195-202.

ULLU, E. & TSCHUDI, C. (1984) Alu sequences are processed 7SL RNA genes. *Nature,* 312**,** 171-2.

VAN DE LAGEMAAT, L. N., GAGNIER, L., MEDSTRAND, P. & MAGER, D. L. (2005) Genomic deletions and precise removal of transposable elements mediated by short identical DNA segments in primates. *Genome Res,* 15**,** 1243-9.

VAN DEN WORM, S. H., STONEHOUSE, N. J., VALEGARD, K., MURRAY, J. B., WALTON, C., FRIDBORG, K., STOCKLEY, P. G. & LILJAS, L. (1998) Crystal structures of MS2 coat protein mutants in complex with wild-type RNA operator fragments. *Nucleic Acids Res,* 26**,** 1345-51.

VASSETZKY, N. S., TEN, O. A. & KRAMEROV, D. A. (2003) B1 and related SINEs in mammalian genomes. *Gene,* 319**,** 149-60.

VENTER, J. C., ADAMS, M. D., MYERS, E. W., LI, P. W., MURAL, R. J., SUTTON, G. G., SMITH, H. O., YANDELL, M., EVANS, C. A., HOLT, R. A., GOCAYNE, J. D., AMANATIDES, P., BALLEW, R. M., HUSON, D. H., WORTMAN, J. R., ZHANG, Q., KODIRA, C. D., ZHENG, X. H., CHEN, L., SKUPSKI, M., SUBRAMANIAN, G., THOMAS, P. D., ZHANG, J., GABOR MIKLOS, G. L., NELSON, C., BRODER, S., CLARK, A. G., NADEAU, J., MCKUSICK, V. A., ZINDER, N., LEVINE, A. J., ROBERTS, R. J., SIMON, M., SLAYMAN, C., HUNKAPILLER, M., BOLANOS, R., DELCHER, A., DEW, I., FASULO, D., FLANIGAN, M., FLOREA, L., HALPERN, A., HANNENHALLI, S., KRAVITZ, S., LEVY, S., MOBARRY, C., REINERT, K., REMINGTON, K., ABU-THREIDEH, J., BEASLEY, E., BIDDICK, K., BONAZZI, V., BRANDON, R., CARGILL, M., CHANDRAMOULISWARAN, I., CHARLAB, R., CHATURVEDI, K., DENG, Z., DI FRANCESCO, V., DUNN, P., EILBECK, K., EVANGELISTA, C., GABRIELIAN, A. E., GAN, W., GE, W., GONG, F., GU, Z., GUAN, P., HEIMAN, T. J., HIGGINS, M. E., JI, R. R., KE, Z., KETCHUM, K. A., LAI, Z., LEI, Y., LI, Z., LI, J., LIANG, Y., LIN, X., LU, F., MERKULOV, G. V., MILSHINA, N., MOORE, H. M., NAIK, A. K., NARAYAN, V. A., NEELAM, B., NUSSKERN, D., RUSCH, D. B., SALZBERG, S., SHAO, W., SHUE, B., SUN, J., WANG, Z., WANG, A., WANG, X., WANG, J., WEI, M., WIDES, R., XIAO, C., YAN, C., et al. (2001) The sequence of the human genome. *Science,* 291**,** 1304-51.

WADDELL, P. J., KISHINO, H. & OTA, R. (2001) A phylogenetic foundation for comparative mammalian genomics. *Genome Inform,* 12**,** 141-54.

WANG, H., IACOANGELI, A., LIN, D., WILLIAMS, K., DENMAN, R. B., HELLEN, C. U. & TIEDGE, H. (2005) Dendritic BC1 RNA in translational control mechanisms. *J Cell Biol,* 171**,** 811-21.

WANG, H., IACOANGELI, A., POPP, S., MUSLIMOV, I. A., IMATAKA, H., SONENBERG, N., LOMAKIN, I. B. & TIEDGE, H. (2002) Dendritic BC1 RNA: functional role in regulation of translation initiation. *J Neurosci,* 22**,** 10232-41.

WATERSTON, R. H., LINDBLAD-TOH, K., BIRNEY, E., ROGERS, J., ABRIL, J. F., AGARWAL, P., AGARWALA, R., AINSCOUGH, R., ALEXANDERSSON, M., AN, P., ANTONARAKIS, S. E., ATTWOOD, J., BAERTSCH, R., BAILEY, J., BARLOW, K., BECK, S., BERRY, E., BIRREN, B., BLOOM, T., BORK, P., BOTCHERBY, M., BRAY, N., BRENT, M. R., BROWN, D. G., BROWN, S. D., BULT, C., BURTON, J., BUTLER, J., CAMPBELL, R. D., CARNINCI, P., CAWLEY, S., CHIAROMONTE, F., CHINWALLA, A. T., CHURCH, D. M., CLAMP, M., CLEE, C., COLLINS, F. S., COOK, L. L., COPLEY, R. R., COULSON, A., COURONNE, O., CUFF, J., CURWEN, V., CUTTS, T., DALY, M., DAVID, R., DAVIES, J., DELEHAUNTY, K. D., DERI, J., DERMITZAKIS, E. T., DEWEY, C., DICKENS, N. J., DIEKHANS, M., DODGE, S., DUBCHAK, I., DUNN, D. M., EDDY, S. R., ELNITSKI, L., EMES, R. D., ESWARA, P., EYRAS, E., FELSENFELD, A., FEWELL, G. A., FLICEK, P., FOLEY, K., FRANKEL, W. N., FULTON, L. A., FULTON, R. S., FUREY, T. S., GAGE, D., GIBBS, R. A., GLUSMAN, G., GNERRE, S., GOLDMAN, N., GOODSTADT, L., GRAFHAM, D., GRAVES, T. A., GREEN, E. D., GREGORY, S., GUIGO, R., GUYER, M., HARDISON, R. C., HAUSSLER, D., HAYASHIZAKI, Y., HILLIER, L. W., HINRICHS, A., HLAVINA, W., HOLZER, T., HSU, F., HUA, A., HUBBARD, T., HUNT, A., JACKSON, I., JAFFE, D. B., JOHNSON, L. S., JONES, M., JONES, T. A., JOY, A., KAMAL, M., KARLSSON, E. K., et al. (2002) Initial sequencing and comparative analysis of the mouse genome. *Nature,* 420**,** 520-62.

WATSON, J. B. & SUTCLIFFE, J. G. (1987) Primate brain-specific cytoplasmic transcript of the Alu repeat family. *Mol Cell Biol,* 7**,** 3324-7.

WEILER, K. S. & WAKIMOTO, B. T. (1995) Heterochromatin and gene expression in Drosophila. *Annu Rev Genet,* 29**,** 577-605.

WEST, N., ROY-ENGEL, A. M., IMATAKA, H., SONENBERG, N. & DEININGER, P. (2002) Shared protein components of SINE RNPs. *J Mol Biol,* 321**,** 423-32.

WILLIAMS, A., HARKER, N., KTISTAKI, E., VEIGA-FERNANDES, H., RODERICK, K., TOLAINI, M., NORTON, T., WILLIAMS, K. & KIOUSSIS, D. (2008) Position effect variegation and imprinting of transgenes in lymphocytes. *Nucleic Acids Res,* 36**,** 2320-9.

ZALFA, F., GIORGI, M., PRIMERANO, B., MORO, A., DI PENTA, A., REIS, S., OOSTRA, B. & BAGNI, C. (2003) The fragile X syndrome protein FMRP associates with BC1 RNA and regulates the translation of specific mRNAs at synapses. *Cell,* 112**,** 317-27.

ZUKER, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res,* 31**,** 3406-15.

Appendix

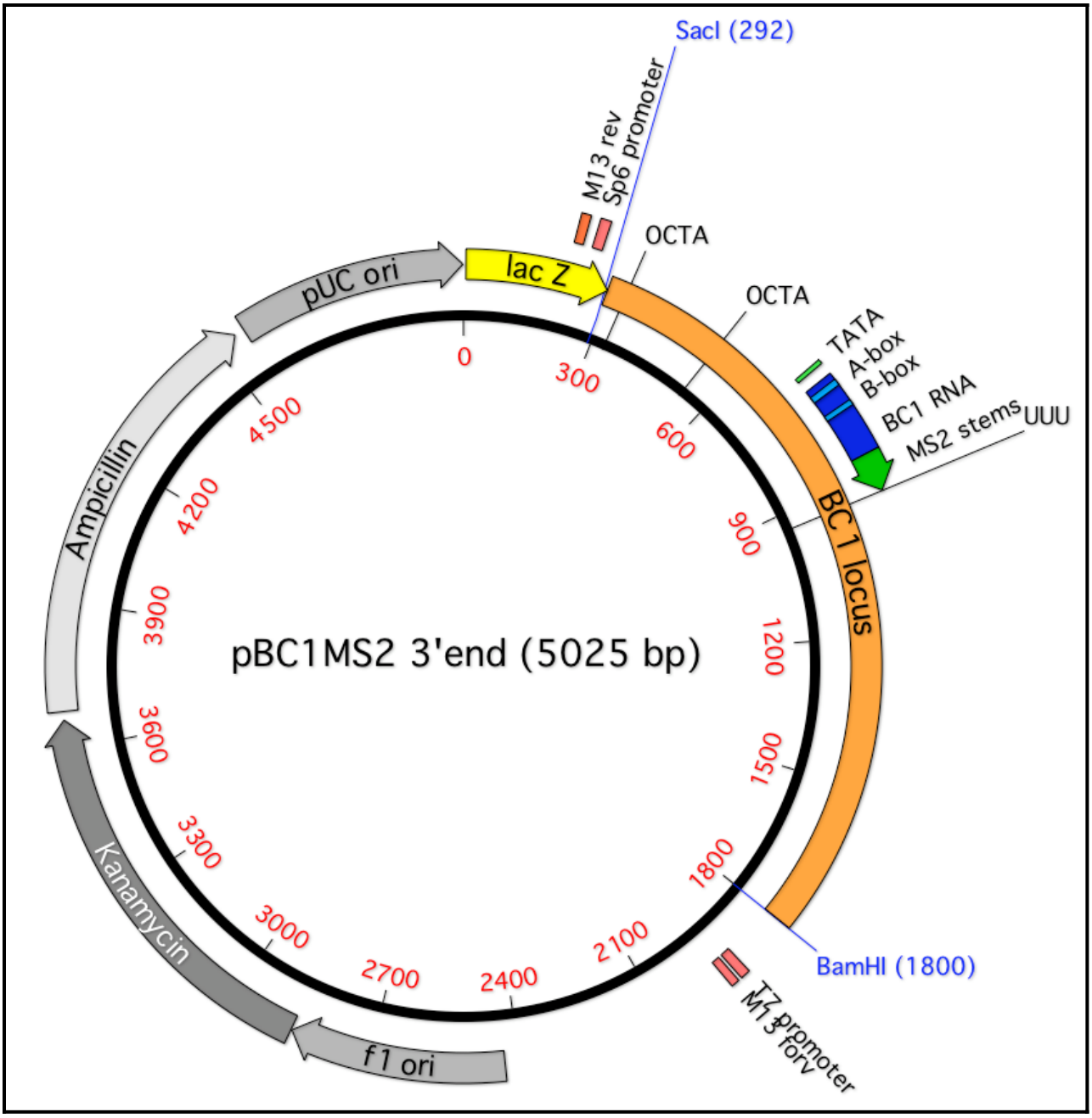


Figure 1 pBC1MS2 plasmid map.
For generating BC1MS2 transgenic mice, the rat BC1 RNA gene was used. The two MS2 RNA motifs [green] were introduced at the very 3' end of the BC1 RNA coding region [blue] proximal to the RNA polymerase III terminator (UUU). The A-box element and B'-box element in the BC1 RNA coding region are indicated [turquoise]. In the upstream region two putative octamer-binding sequences (OCTA) and a TATA-box are shown. MS2 mutagenesis was carried out by PCR based amplification (see text, Materials and Methods, section 1.1). PCR amplified rat BC1MS2 (MS2 stems at the very 3' end of BC1 RNA) fragment was cloned in pCRTOPOII vector (Invitrogen). BC1MS2 transgenic mice were generated using the 1.5 kb *Sac*I - *Bam*HI fragment released from the vector.

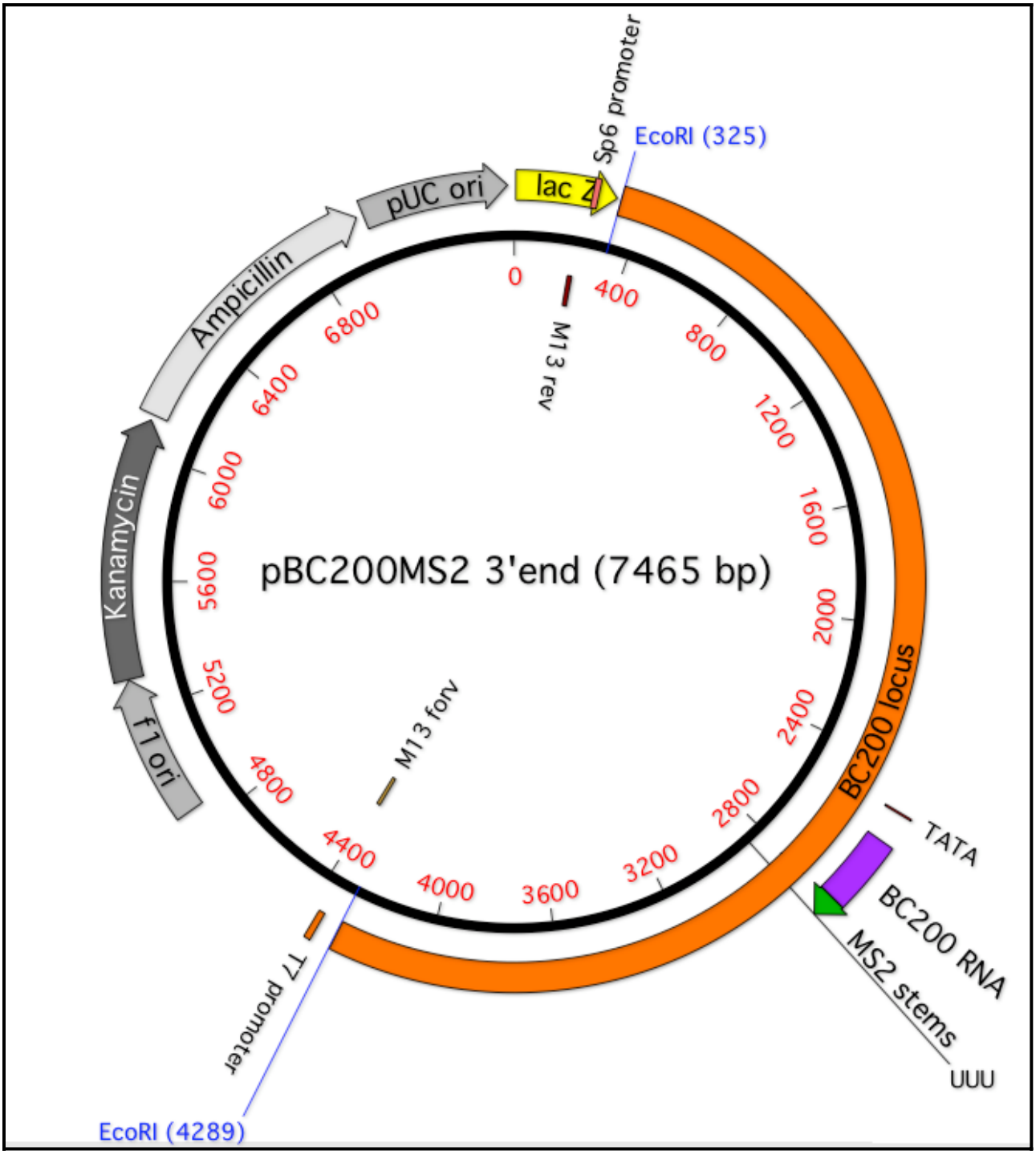


Figure 2 pBC200MS2 plasmid map.
For generating BC200MS2 transgenic mice, the human BC200 RNA locus was used. The two MS2 RNA motifs [green] were introduced at the very 3' end of the BC200 RNA coding region [purple] proximal to RNA polymerase III terminator (UUU). The BC200 RNA locus contains control regions such as a TATA-box. MS2 mutagenesis was carried out by PCR based amplification (see text, Materials and Methods, section 1.1). PCR amplified human BC200MS2 fragment was cloned in pCRTOPOII vector (Invitrogen). BC200MS2 transgenic mice were generated by using the 3.9 kb *Eco*RI - *Eco*RI fragment released from the pBC200MS2 vector.

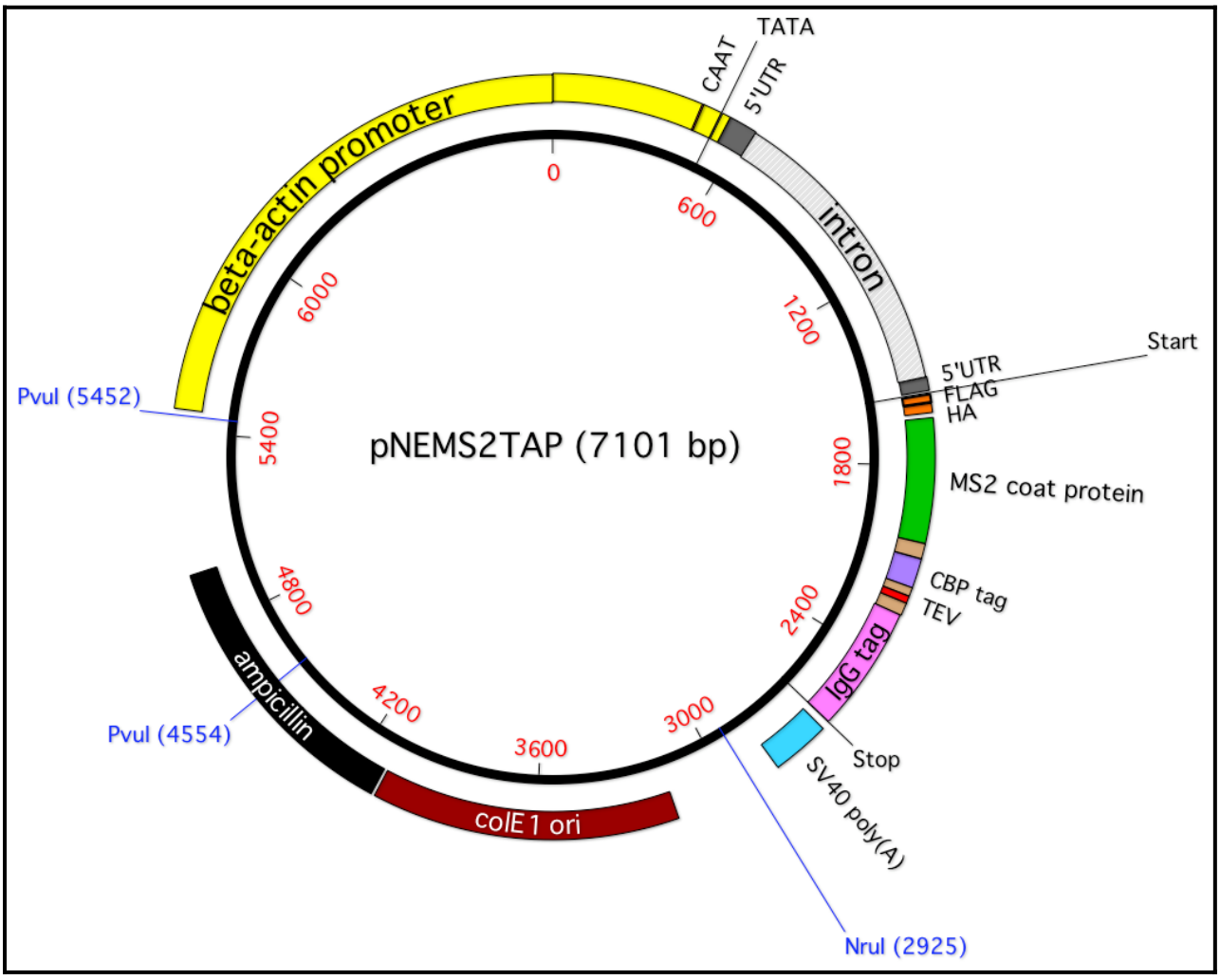


Figure 3 pNEMS2TAP plasmid map.
For generating MS2TAP protein expressing mice, the MS2TAP fragment was cloned into the pNECKu vector (Blichenberg et al., 1999). The vector features the chicken β-actin promoter [yellow] and the 5' UTR [gray] interrupted by a ~1 kb intron [striped area] of the chicken β-actin gene. The MS2TAP fusion protein consists of the MS2 coat protein [green] fused to the CBP tag [purple] and protein A tag [magenta] separated by a TEV protease cleavage site [red]. HA and FLAG peptide sequences [orange] were fused to the N-terminus of MS2TAP fusion protein. The MS2TAP fusion is followed by a polyadenylation signal from a fragment from SV40 virus [turquoise].The fragment encoding the MS2 coat protein was obtained from plasmid pG14-MS2-GFP (Bertrand et al., 1998); the fragment encoding the TAP protein from pTAP (Rigaut et al., 1999). (see text for PCR amplification details, Materials and Methods, section 1.2). MS2TAP transgenic mice were generated by using the ~4.6 kb *Pvu*I - *Nru*I fragment released from pNEMS2TAP.

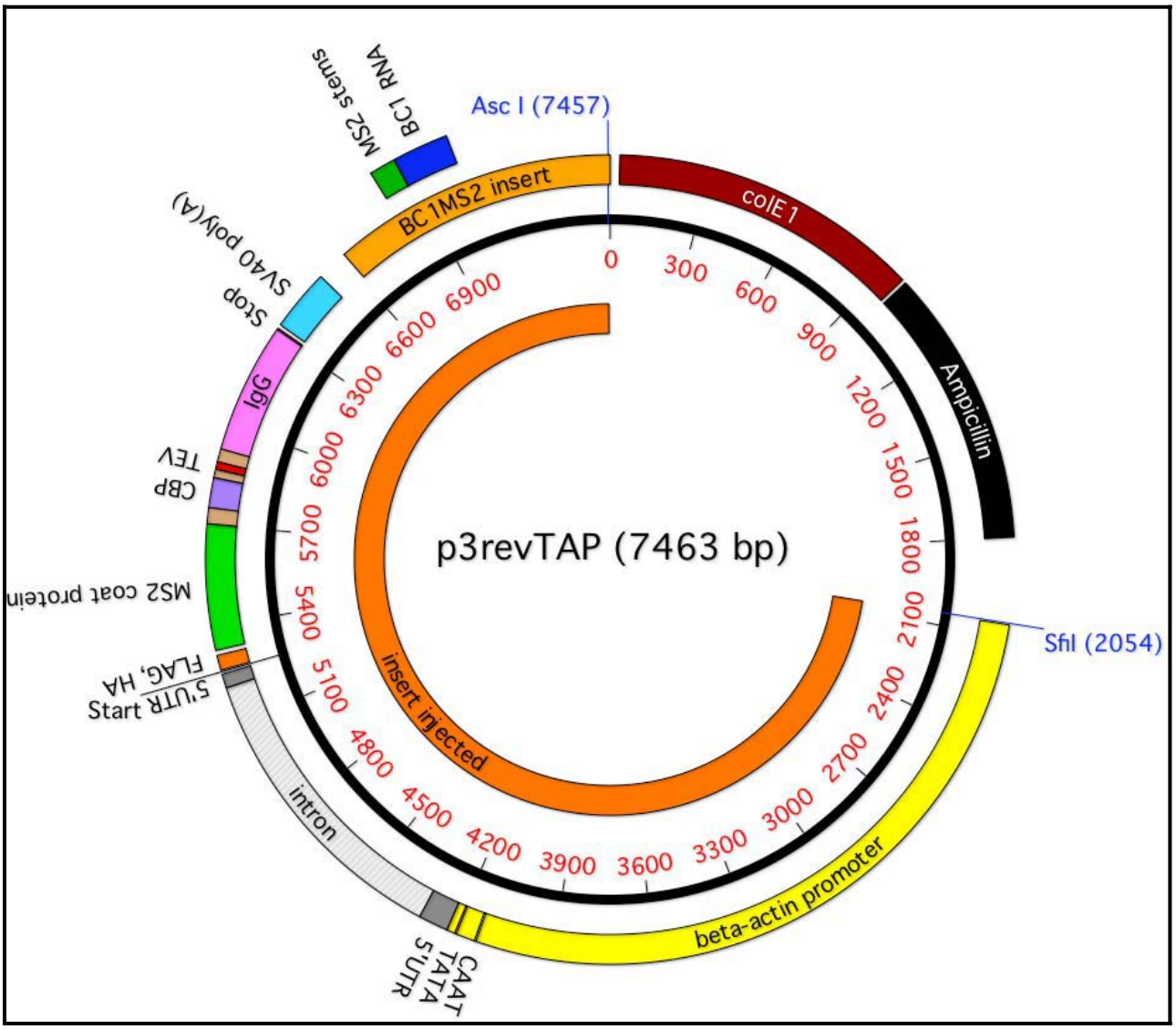


Figure 4 p3revTAP plasmid map.
Both the RNA expressing cassette (BC1MS2 RNA – MS2 at the very 3' end of BC1 RNA) and the protein expressing cassette (MS2TAP fusion protein) were cloned into a single vector. The plasmid features BC1MS2 RNA (blue and green) fragment and the chicken β-actin promoter [yellow] including the 5' UTR [gray] interrupted by a ~1 kb intron [striped area]. The MS2TAP fusion protein features the MS2 coat protein [green] fused to the TAP module with the CBP tag [purple] and protein A tag [magenta] separated by a TEV protease cleavage site [red] followed by a polyadenylation signal from a fragment from SV40 virus containing poly(A) additional signal [turquoise]. HA and FLAG peptide sequences [orange] were fused to the N-terminus of MS2TAP fusion protein. The BC1MS2 fragment (*Sac*I - *Apa*I) contained 423 bp upstream of the BC1MS2 RNA gene and about 200 bp downstream of the RNA polymerase III terminator of the RNA gene was obtained from pBC1MS2 (Appendix, **Fig 1)** and cloned into pBluescript II KS+. The MS2TAP fragment was isolated by *Pvu*I - *Nru*I digestion from pNEMS2TAP plasmid (Appendix, **Fig 3**) and subcloned into pBluescript II KS+, which harbored BC1MS2 (with the MS2 domains located at the very 3'end of BC1 RNA). The plasmid with MS2TAP inserted in the opposite orientation with respect to BC1MS2 [head to head] was designated the plasmid as 3 revTAP. The insert carrying the RNA and protein coding genes (inner circle) were released with *Sfi*I and *Asc*I to generate the 3 revTAP transgenic mice.

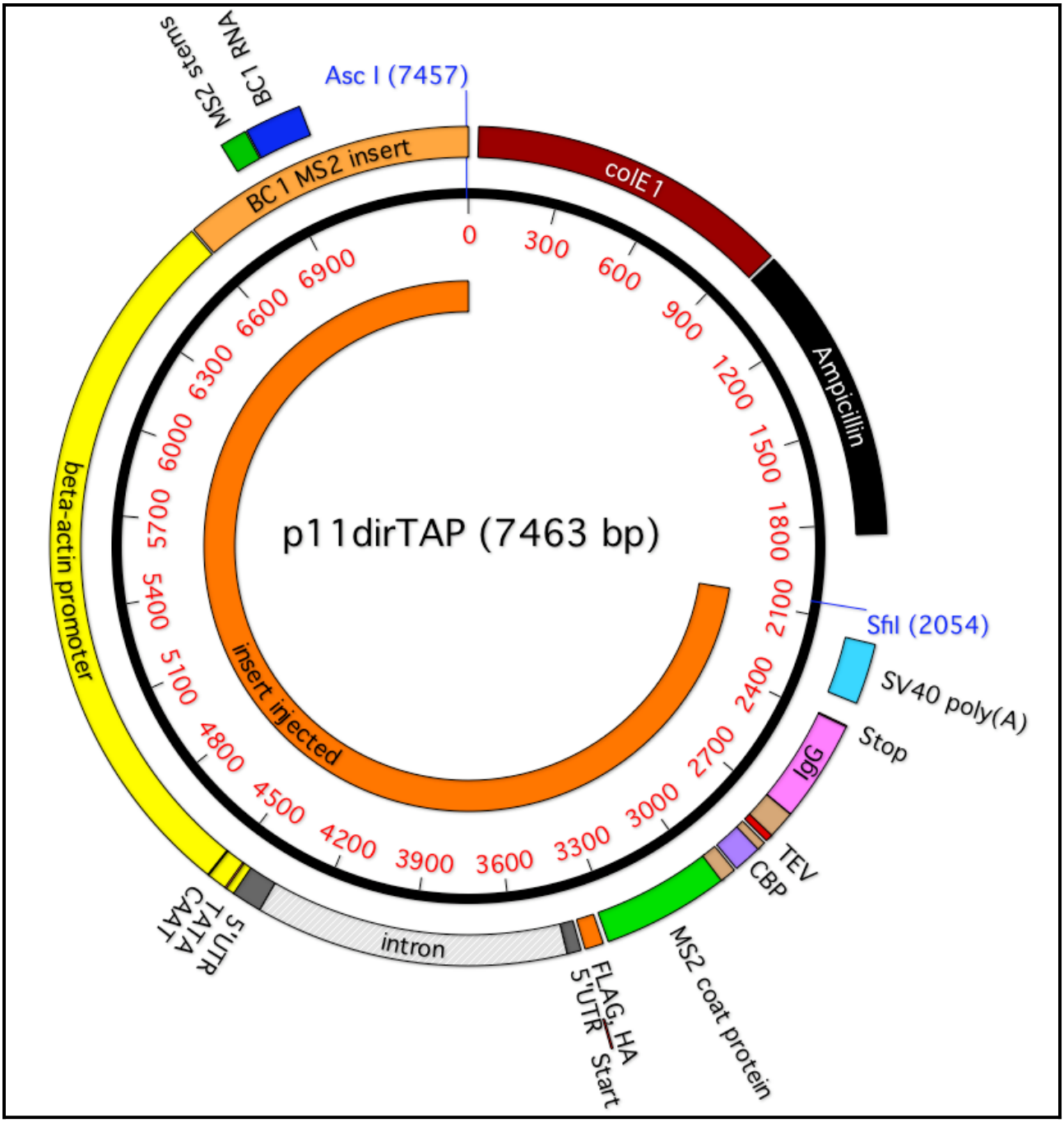


Figure 5 p11dirTAP plasmid map.
Same as Appendix, **Fig. 4**, except that the cassette with MS2TAP inserted in the opposite orientation with respect to BC1MS2 [head to tail]. The vector was designated as 11 dirTAP. The insert carrying the RNA and protein coding genes (inner circle) were released with *Sfi*I and *Asc*I to generate the 11 dirTAP transgenic mice.

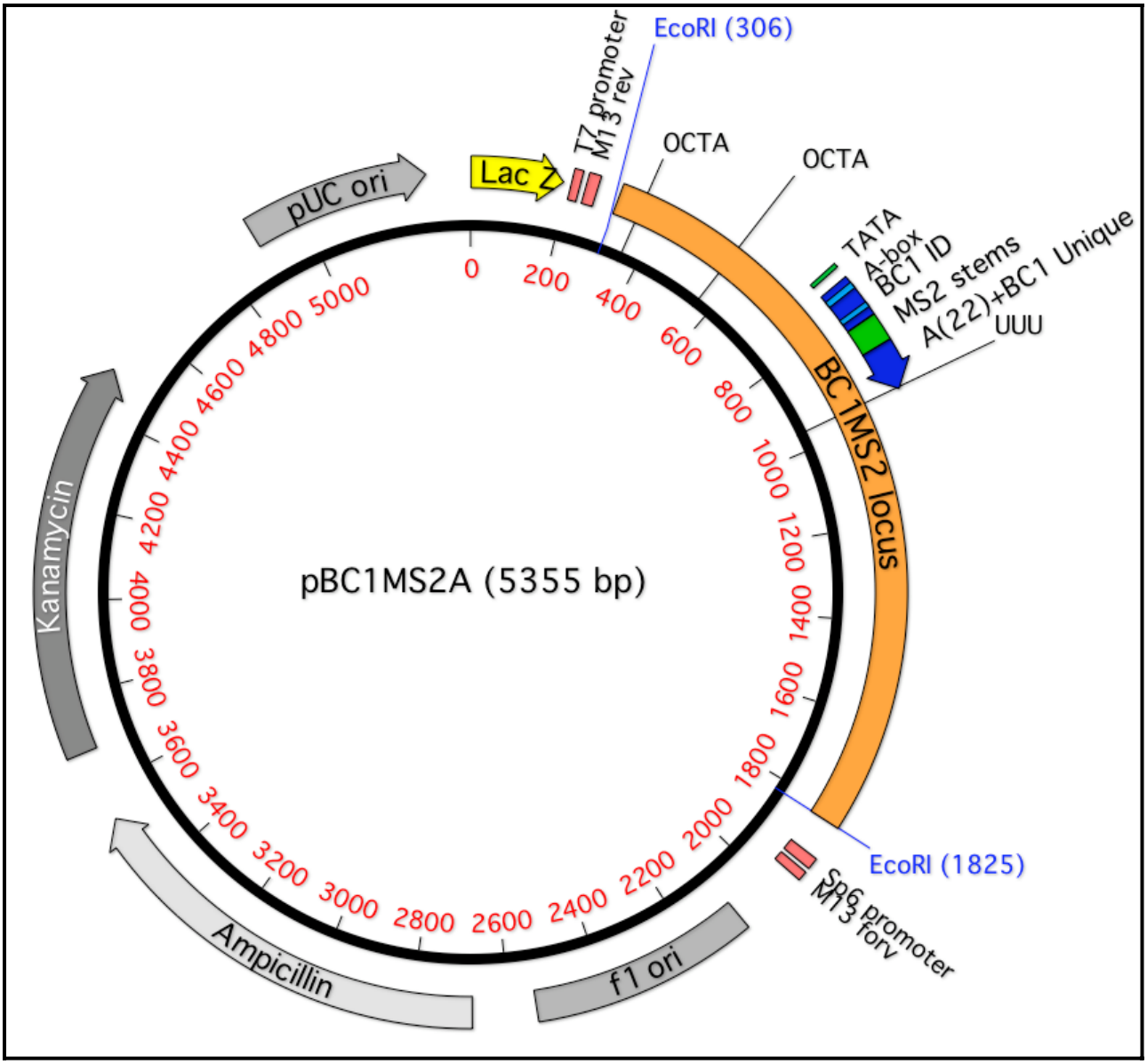


Figure 6 pBC1MS2A plasmid map.
In this construct, the two MS2 RNA motifs [green] were placed closer to the ID domain internal within the BC1 RNA coding region [blue] followed by A_{22} and the BC1 RNA unique region. The RNA polymerase III transcription termination signal from wild type locus was placed downstream to the MS2 RNA motifs. The A-box element and B'-box element in the BC1 RNA coding region are indicated [turquoise]. In the upstream region two putative octamer-binding sequences (OCTA) and a TATA-box are shown. MS2 mutagenesis was carried out by PCR based amplification (see text, Mareials and Methods, section 1.4). PCR amplified rat BC1MS2A fragment was cloned in pDrive cloning vector (Qiagen, Hilden). BC1MS2A transgenic mice were generated by inserting the 1.5 kb *Eco*RI fragment released from vector.

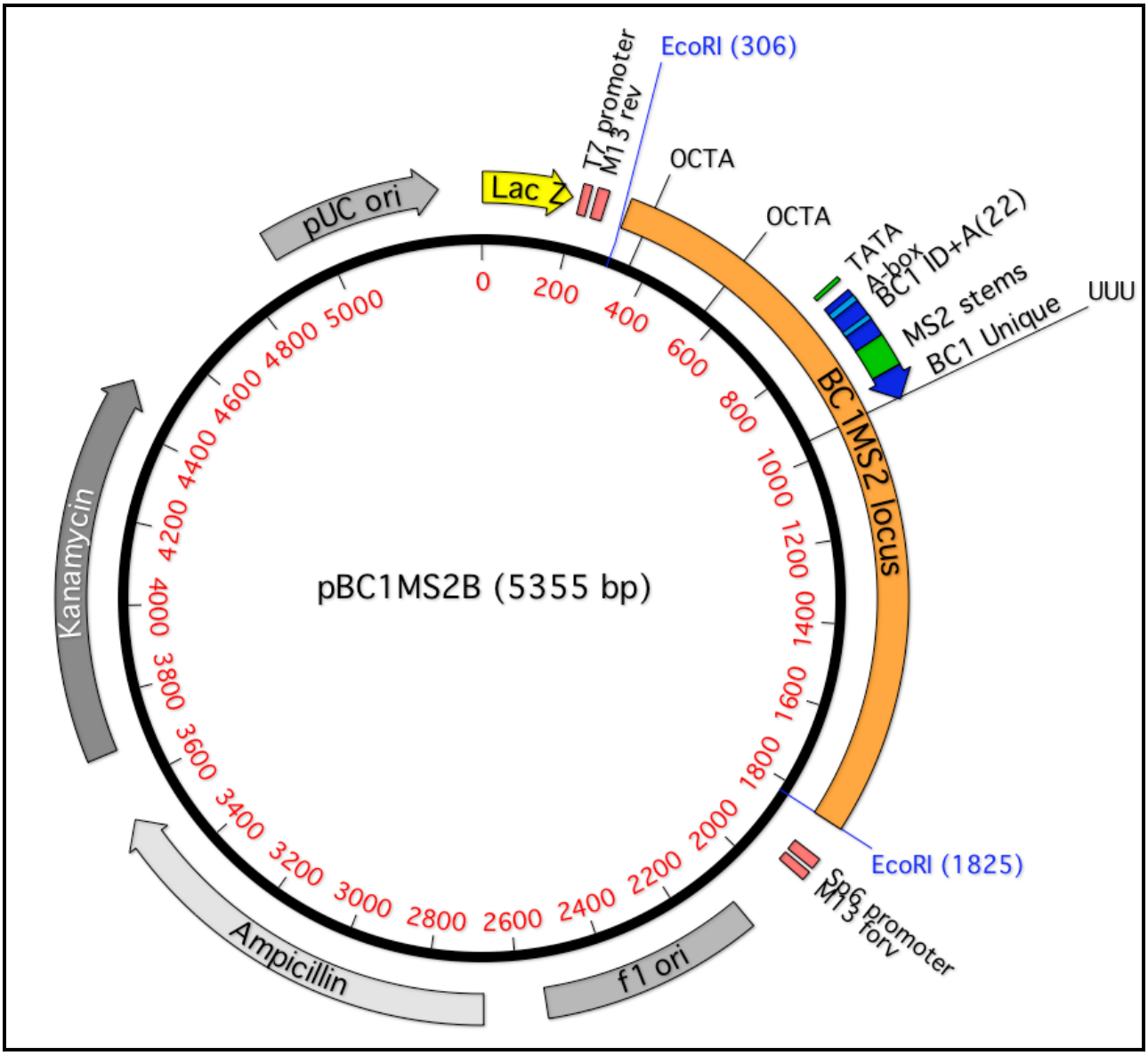


Figure 7 pBC1MS2B plasmid map.
In this construct, the two MS2 RNA motifs [green] were placed closer to the 3' unique domain within the BC1 RNA coding region [blue] downstream of the A_{22} sequence. The RNA polymerase III transcription termination signal from wild type locus was placed downstream to the MS2 RNA motifs. The A-box element and B'-box element in the BC1 RNA coding region are indicated [turquoise]. In the upstream region two putative octamer-binding sequences (OCTA) and a TATA-box are shown. MS2 mutagenesis was carried out by PCR based amplification (see text, Mareials and Methods, section 1.4). PCR amplified rat BC1MS2B fragment was cloned in pDrive cloning vector (Qiagen, Hilden). BC1MS2B transgenic mice were generated by inserting the 1.5 kb *Eco*RI fragment released from vector.

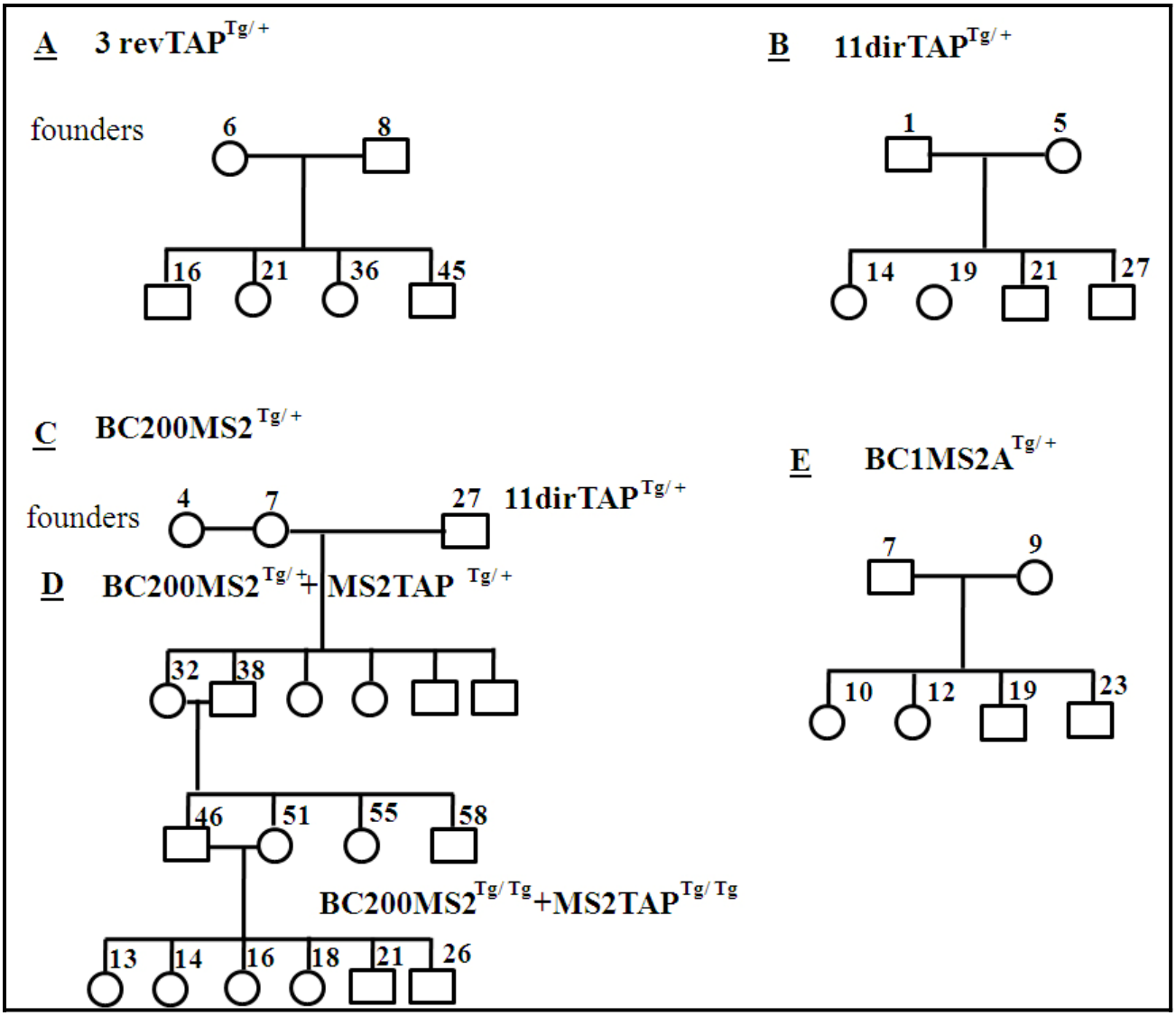


Fig 8 Pedigree of 3 revTAP, 11 dirTAP, BC200MS2, BC200MS2+MS2TAP and BC1MS2A transgenic mice. 3 revTAP and 11 dirTAP transgenic mice were generated by injecting fragment harbor RNA and protein coding genes (Appendix, **Fig 4 and 5**). Offsprings of the 3 revTAP and 11 dirTAP was analysed by Northern blot (BC1MS2 RNA – not detected, 3 revTAP mice # 21, 45 and 11 dirTAP mice # 19 and 25, see **Fig 3.5**) and Western blot (MS2TAP fusion protein detected, 11 dirTAP mouse # 21). BC200MS2 transgene was generated by injecting BC200MS2 gene fragment (Appendix, **Fig 2**). Founders of BC200MS2 RNA expressing transgene was crossed with 11 dirTAP transgene (expresses MS2TAP fusion protein). The resulted transgenic mice (BC200MS2+MS2TAP mice # 11, 13 and 21, see **Fig 3.5**) showed BC200MS2 RNA expression. BC1MS2A transgenic mice were generated by injecting BC1MS2A gene (Appendix, **Fig 6**). Offsprings were analyzed by Northern blot and were unable to detect BC1MS2A RNA expression (BC1MS2A mice # 10 and 12, see 3.11).

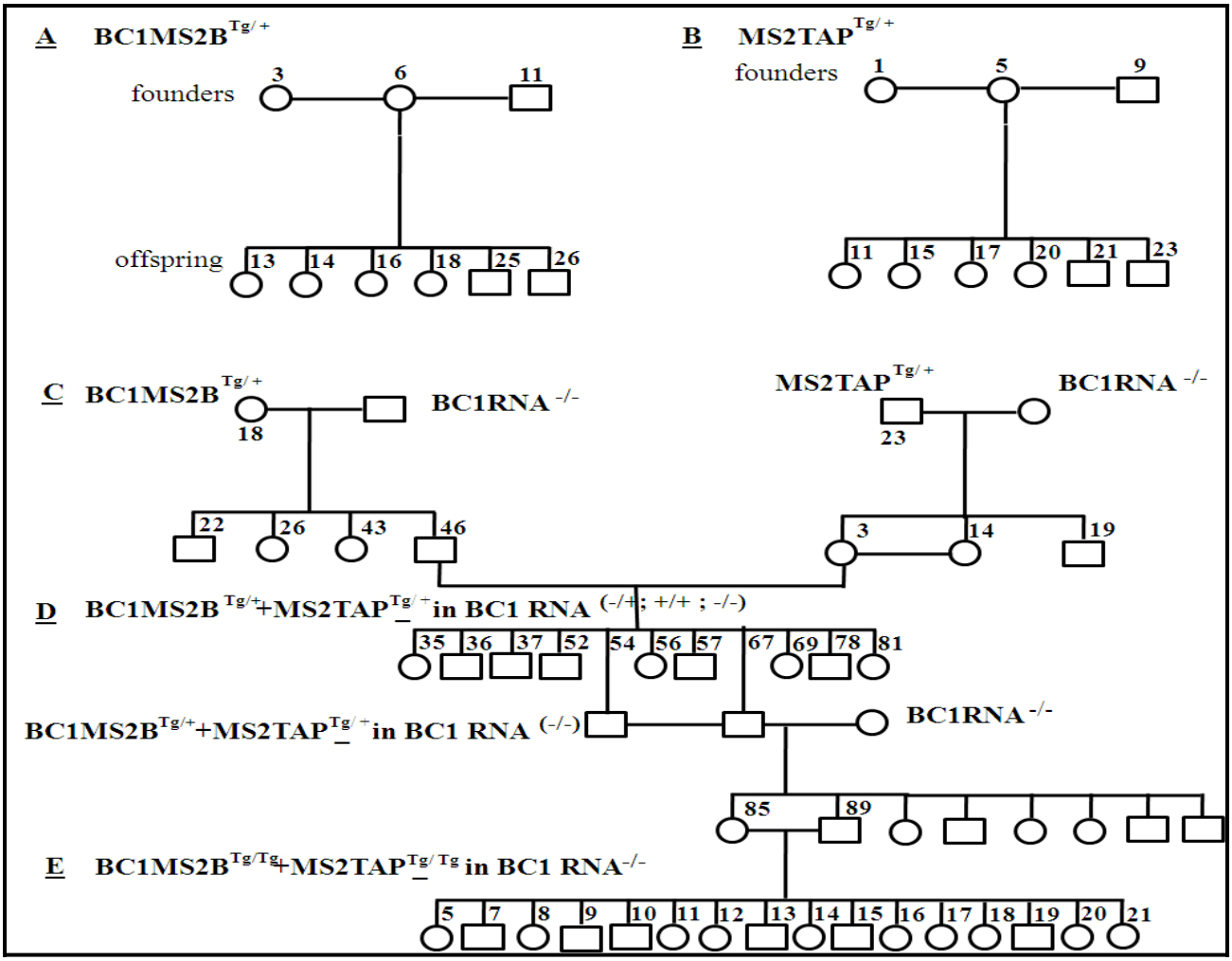


Fig 9 Pedigree of BC1MS2B, BC1MS2B_in BC1 RNA$^{-/-}$ BC1MS2B+MS2TAP_ in BC1 RNA$^{-/-}$ transgenic mice. Transgenic mice used in this study was obtained from different breedings, for analyzing Norhern blot, Western blot and Southern blot are shown in the pedigree. BC1MS2B transgenic mice was generated by injecting the DNA fragment (BC1MS2B, Appendix, **Fig 7**)**.** The resulted mouse founders (BC1MS2B mice, **Fig 8A** # 3, 6, 11) were crossed among each transgenic mouse. The resulted offspring was used for Northern blot analysis (BC1MS2B mice # 14, 16, 25, see **Fig 3.11**). Among the offspring of three founder mice, one mouse **(**BC1MS2B mouse #16, see **Fig 3.11**) showed relatively high levels of BC1MS2B RNA expression. Therefore expanded that transgenic mouse line. MS2TAP transgenic mice were generated by injecting DNA (Appendix, **Fig 5**).The siblings from the BC1MS2B mouse #16 (BC1MS2B mouse #18) was crossed with MS2TAP transgenic mice, observed further improvement in BC1MS2B RNA expression (BC1MS2B + MS2TAP, mouse # 18, see **Fig 3.11**). The resulted transgenic mouse (BC1MS2B + MS2TAP) was bred with BC1 RNA$^{-/-}$ and observed three types of BC1MS2B transgne integration into the mouse genome (see **Fig 3.12**). Among the three types of integration low copy number transgenic mouse (BC1MS2B+MS2TAP mice, **Fig 8C** # 54 and 67, see **Fig 3.12**), BC1MS2B RNA expression was observed. Therefore expanded and further crossed with BC1 RNA$^{-/-}$ transgenic mice to attain homozygosity (BC1MS2B+MS2TAP_in.BC1 RNA$^{-/-}$, mice # 5 and 10, see **Fig 3.13)**. The resulted transgenic mice was bred for 3-4 generations to obtain homozygous for both the transgenes [BC1MS2B$^{Tg/Tg}$ and MS2TAP$^{Tg/Tg}$ genes], (BC1MS2B$^{Tg/Tg}$ and MS2TAP$^{Tg/Tg}$ _in.BC1 RNA$^{-/-}$, mice # 5, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21).