Stony Brook University

OFFICIAL COPY

The official electronic file of this thesis or dissertation is maintained by the University Libraries on behalf of The Graduate School at Stony Brook University.

© All Rights Reserved by Author.

Characterization of MEN ε/β long non-coding RNAs

as essential components of nuclear paraspeckles

A Dissertation Presented

by

Hongjae Sunwoo

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Molecular and Cellular Biology

(Cellular and Developmental Biology)

Stony Brook University

May 2009

Stony Brook University
The Graduate School

Hongjae Sunwoo

We, the dissertation committee for the above candidate for the
Doctor of Philosophy degree, hereby recommend
acceptance of this dissertation

Dr. David L. Spector – Dissertation Advisor
Professor and Director of Research
Cold Spring Harbor Laboratory

Dr. Gary Zieve – Chairperson of Defense
Associate Professor, Department of Pathology
Stony Brook University

Dr. Adrian R. Krainer
Professor
Cold Spring Harbor Laboratory

Dr. Paul A. Fisher
Professor, Department of Pharmacological Sciences
Stony Brook University

Dr. Michael Q. Zhang
Professor
Cold Spring Harbor Laboratory

This dissertation is accepted by the Graduate School

Lawrence Martin
Dean of the Graduate School

Abstract of the Dissertation

Characterization of MEN ε/β long non-coding RNAs

as essential components of nuclear paraspeckles

by

Hongjae Sunwoo

Doctor of Philosophy

in

Molecular and Cellular Biology

(Cellular and Developmental Biology)

Stony Brook University

2009

Studies of the transcriptional output of the human and mouse genomes have revealed that there are many more transcripts produced than can be accounted for by predicted protein-coding genes. Using a custom microarray, we have identified 184 non-coding RNAs that exhibit more than 2-fold up- or down-regulation upon differentiation of C2C12 myoblasts into myotubes. Here, I focus on the *Men ε/β* locus, which is up-regulated 3.3-fold during differentiation. Two non-coding RNA isoforms are produced from a single RNA polymerase II

promoter, differing in the location of their 3' ends. *Men* ε is a 3.2-kb polyadenylated RNA, whereas *Men* β is a ~20-kb transcript containing a genomically encoded poly(A)-rich tract at its 3' end. The 3' end of *Men* β is generated by RNase P cleavage. The *Men* ε/β transcripts are localized to nuclear paraspeckles. During this study, I developed and characterized mouse monoclonal antibodies against PSPC1, a paraspeckle marker protein, that can be utilized for immunoblotting, immunofluorescence microscopy, and immunoprecipitation. Using one of hybridomas, I demonstrate that the *Men* ε/β transcripts directly interact with Nono. Knock-down of *MEN* ε/β expression results in the disruption of nuclear paraspeckles. Furthermore, the formation of paraspeckles, after release from transcriptional inhibition by DRB treatment, was suppressed in *MEN* ε/β depleted cells. These findings indicate that the *MEN* ε/β non-coding RNAs are essential structural/organizational components of paraspeckles.

This work is dedicated to my wife, Joonhee Kim

Table of Contents

List of Figures and Tables

List of Abbreviations

3’ UTR	3’ Untranslated Region
abd-A	abdominal-A
Abd-B	Abdominal-B
Airn	*Antisense Igf2r RNA*
ASO	Antisense Oligonucleotide
Co-IP	Co-Immunoprecipitation
CTN-RNA	*Cat2 Transcribe Nuclear RNA*
DAPI	4',6-diamidino-2-phenylindole
DHFR	Dihydrofolate Reductase
DRB	5, 6-dichlorobenzimidazole 1-β-D-ribofuranoside
ELISA	Enzyme-Linked ImmunoSorbent Assay
EMT	Epithelial-mesenchymal Transition
EST	Expressed Sequence Tag
Ezh2	Enhancer of zeste homolog 2
GST	Glutathione S-Transferase
HOTAIR	*HOX Antisense Intergenic RNA*
iab	infraabdominal
IB	Immunoblotting
ICR	Imprinting Control Region
IF	Immunofluorescence microscopy

IPTG	Isopropyl β-D-1-thiogalactopyranoside
IRES	Internal Ribosome Entry Site
Kcnq1ot1	*Kcnq1 overlapping transcript1*
MALAT1	*Metastasis Associated in Lung Adenocarcinoma Transcript1*
mascRNA	*MALAT1-associated small cytoplasmic RNA*
MEN ε/β	*Multiple Endocrine Neoplasia epsilon/beta*
ncRNA	non-coding RNA
NFAT	Nuclear Factor of Activate T cells
NRON	*Non-coding Repressor of NFAT*
ORF	Open Reading Frame
PCA3	*Prostate Cancer Antigen 3*
PcG	Polycomb group protein complex
PML body	Promyelocytic Leukemia body
PRC2	Polycomb-Repressive Complex 2
PRE	Polycomb Response Element
PSPC1	Paraspeckle Protein Component 1
Q-PCR	Realtime Quantitative-Polymerase Chain Reaction
RACE	Rapid Amplification of cDNA End
RepA	*Repeat A*
RISC	RNA-induced Silencing Complex
RNA FISH	RNA Fluorescence *in situ* Hybridization

RNAi	RNA interference
rncs-1	*RNA non-coding, starvation up-regulated*
RNP	Ribonucleoprotein Particle
RPA	RNA Protection Assay
RRM	RNA Recognition Motif
RT-PCR	Reverse Transcribed-Polymerase Chain Reaction
SINE	Short Interspersed Element
TLS	Translocated in Liposarcoma
TncRNA	*Trophoblast ncRNA*
TrxG	Trithorax group protein complex
Tsix	*X-inactive specific transcript, antisense*
TSU	Trophoblast STAT utron
UAS	Upstream Activation Sequences
Ubx	Ultrabiothorax
Xist	*X-inactive specific transcrip*t
Xlsirts	*Xenopus laevis short interspersed repeat transcripts*

Acknowledgements

This dissertation could not be done without the help of many people. First of all, I am very grateful to my advisor, David L. Spector, for his consistent encouragement and invaluable advice throughout the course of my dissertation study. He is an outstanding mentor and our discussions have fundamentally nourished my growth as a cell biologist. I would like to acknowledge my collaborators John Mattick and Marcel Dinger for providing me with a handful of promising candidates to work on. Without them, this study would not have been initiated. I would also like to express my appreciation to all current and former members of the Spector laboratory for many insightful discussions and technical assistance. In particular, Kannanganattu Prasanth, Jeremy Wilusz, and Megan Bodnar deserve special thanks. When I joined the Spector lab, I learned many techniques used in these studies from Prasanth. Jeremy is a great benchmate and collaborator. We sit next to each other in the lab and study two non-coding RNAs that sit next to each other in the genome. Megan put a lot of her time and efforts into reading and correcting my dissertation.

I would like to thank all of my committee members for their advice and support throughout my research. I appreciate that Gary Zieve kindly took a responsibility as a chairperson of committee. For every committee meeting, he and Paul Fisher gladly drove to CSHL.

I thank Angus Lamond, Yasuyuki Kurihara, Johannes Zuber, Sidney Altman, and Masayuki Nashimoto for PSPC1 antibody, Sfpq antibody, retroviral vectors, human RNase P, and human RNase Z, respectively. I also appreciate Zhenyu Xuan for helpful discussions of bioinformatic analysis and Carmelita Bautista for her help in generating monoclonal antibodies. I thank C. Frank Bennett, Sue Freier, and Chris Black at ISIS Pharmaceuticals for ASOs, one of the most valuable reagents used during this study, and for helpful discussions.

Most of all, I thank my wife for her continuous love and encouragement. Soon after our wedding, we moved to the U.S. Now I am defending my dissertation and we have two lovely babies, Sophia and Stanley. My parent, brothers, parents-in-law, and sister-in-law support me in many ways. I would not have finished my study without them.

Publications

Wilusz, J.E., **Sunwoo, H.**, and Spector, D.L. Long non-coding RNAs: Functional surprises from the RNA World. *Genes & Development*. *In press*.

Sunwoo, H., Dinger, M.E., Wilusz, J.E., Amaral, P.P., Mattick, J.S., and Spector, D.L. (2009). *MEN ε/β* nuclear retained non-coding RNAs are up-regulated upon muscle differentiation and are essential components of paraspeckles. *Genome Research* **19**: 347-359.

I. Introduction

The number of genes in the human genome was once estimated over 120,000 based on the analysis of the EST (Expressed Sequence Tag) database (Liang et al. 2000). Therefore, the result of the human genome project provided quite a surprise to many when it was determined that there are only ~20,000 protein-coding genes, representing less than 2% of the total genomic sequence (International Human Genome Sequencing Consortium 2004). While the single cell eukaryote *S. cerevisiae* has quite fewer protein-coding genes (6,532), all multicellular organisms harbor a very similar number of protein-coding genes regardless of their developmental or physiological complexities; fruitfly (14,141), *C. elegans* (20,176), Zebrafish (17,330), mouse (23,019) and human (21,343) (http://www.ensembl.org). Thus, the biological complexity of organisms can not be explained simply by the number of protein-coding genes. In contrast to this relatively stable size of protein coding genes, the proportion of non-coding DNA increases dramatically from less than 20% in prokaryotes, 30 – 45% in simple eukaryotes, to 98% in humans. Based on such analyses, A hypothesis was proposed that there must be another layer of mechanism, never been appreciated, but yet provides the biological complexity observed in higher organisms (for review, see Mattick 2004).

This has prompted an explosion of research addressing possible functional roles of genomic sequences that do not encode proteins. Recent reports indicate that transcription is not limited to genic regions but is instead pervasive throughout the mammalian genomes, as demonstrated by large-scale cDNA cloning projects (Carninci et al. 2005) and genomic tiling arrays (Bertone et al. 2004; Birney et al. 2007; Cheng et al. 2005; Kapranov et al. 2007a). In fact, the human transcriptome covers over 90% of the

human genome (Birney et al. 2007), and is composed of a complex network of transcripts that includes tens of thousands of long RNAs with little or no protein-coding capacity (for review, see Kapranov et al. 2007b).

There is still some debate as to whether this pervasive transcription represents largely useless transcription by misfiring of RNA polymerase II (transcriptional noise) (Ebisuya et al. 2008; Struhl 2007; Wang et al. 2004) or if these non-coding RNAs (ncRNAs) have functions that simply have not yet been assigned (for review, see Mattick 2004). However, considering that it has long been known that the most abundant classical non-coding transcripts, such as ribosomal RNAs, transfer RNAs, and spliceosomal RNAs, are critical components of many cellular machines, it seems highly likely that additional ncRNAs play key regulatory and functional roles. Supporting a hypothesis of functional ncRNAs, multiple studies have shown that significant numbers of long ncRNAs are regulated during development (Blackshaw et al. 2004; Dinger et al. 2008; Rinn et al. 2007), exhibit cell type-specific expression (Mercer et al. 2008; Ravasi et al. 2006), localize to specific subcellular compartments (Clemson et al. 2009; Hutchinson et al. 2007; Sasaki et al. 2009; Sone et al. 2007; Sunwoo et al. 2009), and are associated with diseases (for review, see Costa 2005; Prasanth and Spector 2007). In addition, some long ncRNAs have been found to be under evolutionary selection (Guttman et al. 2009; Pheasant and Mattick 2007; Pollard et al. 2006; Ponjavic et al. 2007).

Here, I summarize recent studies that have identified numerous paradigms for how long non-coding RNAs function on the molecular level (Figure 1.1).

Non-coding RNAs balance gene expression between male and female

Dosage compensation in female mammals involves an inactivation of one X chromosome to balance the level of gene expression from the X chromosome between female and male (for review, see Erwin and Lee 2008; Heard and Disteche 2006). Two well known ncRNAs *Xist* (*X-inactive specific transcript*) and *Tsix* (*X-inactive specific transcript, antisense*) are essential regulators of X inactivation. At the stage of X inactivation in mammalian development, the expression of *Tsix* is repressed on the future inactive X chromosome, thereby allowing up-regulation of *Xist* (Luikenhuis et al. 2001; Stavropoulos et al. 2001). The *Xist* transcripts subsequently coat the chromosome *in cis* (Clemson et al. 1996) and induce perinucleolar localization of the X chromosome during S phase (Zhang et al. 2007), representing part of the silencing mechanism. In parallel on the future active X chromosome, *Tsix* is persistently expressed, which prevents up-regulation of its antisense transcript *Xist* (Luikenhuis et al. 2001; Stavropoulos et al. 2001). In turn, *Tsix* is regulated by another ncRNA *Xite*, located ~10-kb upstream of *Tsix* (Stavropoulos et al. 2005).

Although the importance of *Xist* (Marahrens et al. 1997; Penny et al. 1996) and *Tsix* (Lee et al. 1999a) during X inactivation was shown in genetic studies, their interacting proteins have not been identified. Recently, the *Xist* transcripts and the endogenous *RepA* (*Repeat A*) ncRNA originating within the *Xist* locus were shown to directly bind to Ezh2 (Enhancer of zeste homolog 2), a histone methyltransferase and member of Polycomb-repressive complex 2 (PRC2), suggesting that at least in this case ncRNAs may have taken over the function of the PRE (Polycomb response element) in mammals (Zhao et al. 2008).

There is also emerging evidence that *Xist* and *Tsix* may also be processed to yield small RNAs (Ogawa et al. 2008). Developmentally regulated small RNAs (25 – 42-nt) map to the *Xist* and *Tsix* loci. Because *Xist* and *Tsix* form a double-stranded RNA (dsRNA) duplex *in vivo* and the level of the small RNAs is diminished when Dicer is deleted, it was suggested that processing of dsRNA by Dicer generates these small RNAs (Ogawa et al. 2008). However, Dicer is not currently known to cleave RNAs to small transcripts in this size range, suggesting that Dicer may play an indirect role in the biogenesis of these small RNAs and that *Xist* and *Tsix* are processed via other mechanisms. Further discovery of ncRNA binding proteins will elucidate how *Xist* and *Tsix* control X inactivation in mammals at the molecular level.

Regulation of Hox genes by non-coding RNAs

Hox genes are critical for correct body patterning of the axes during development. The ~330-kb *BX-C* locus in *Drosophila* encodes three Hox genes; Abdominal-B (Abd-B), abdominal-A (abd-A) and Ultrabithorax (Ubx). The spatial and temporal expression pattern of BX-C determines thoracic and abdominal segmentations (for review, see Maeda and Karch 2006). While the Polycomb group protein complex (PcG) maintain the default repressed state of Hox genes, the Trithorax group protein complex (TrxG) counteracts PcG-mediated silencing, resulting in activation of Hox genes (Klymenko and Muller 2004). It was reported more than two decades ago that transiently expressed ncRNAs during early embryogenesis, originating from the *bithoraxoid (bxd)* locus, can regulate the expression of Ubx (Hogness et al. 1985; Lipshitz et al. 1987). Since then, the *BX-C* locus has been intensively studied and several models have been suggested to

explain how ncRNAs can regulate the expression of Hox genes (for review, see Lempradl and Ringrose 2008).

Correct expression of the Abd-B and abd-A requires the 100-kb *infraabdominal* (*iab*) region that contains an array of seven domains, *iab2 – 8*. Each *iab* domain is responsible for development of a corresponding abdominal segment (Karch et al. 1985). Non-coding transcripts in the *iab* domains are expressed only in corresponding abdominal regions slightly prior to segmentation, suggesting that the *iab* transcripts may regulate segmentation (Bae et al. 2002). Two miRNAs, iab-4-3p and iab-4-5p, are generated from the *iab-4* antisense transcript (Aravin et al. 2003). Later, the 3' UTR of *Ubx* mRNA was identified as a direct target of iab-4-5p (Ronshaugen et al. 2005). The loss of *iab-7* transcript was implicated with repression of Abd-B during embryogenesis (Hogga and Karch 2002). It was also argued that transcription of ncRNAs through Polycomb response element (PRE) *in cis* regulatory domains can activate Hox genes via chromatin remodeling (Rank et al. 2002). Moreover, maintenance of active state requires persistent transcription, regardless of its orientation, suggesting that transcription, rather than ncRNA product, is essential (Schmitt et al. 2005).

On the other hand, three ncRNAs, *TRE1 – 3* were transcribed from the *bxd* locus (Sanchez-Elsner et al. 2006). *TRE1 – 3* ncRNAs bind and recruit TrxG protein Ash1, a histone methyltransferase, to the *Ubx* promoter, resulting in expression of Ubx. Surprisingly, ectopic expression of *TRE1 – 3* showed the same activation of Ubx, suggesting s trans mechanism. However, there is a discrepancy about the role of the *bxd* ncRNAs (for review, see Lempradl and Ringrose 2008). The expression of Ubx was repressed in cells where when non-coding transcripts, different from *TRE1 – 3*, from the

bxd locus were detected (Petruk et al. 2006). In addition, knock-down did not have any phenotype, suggesting that transcription may be more important than RNA product *per se* (Petruk et al. 2006). Although multiple ncRNAs may function differently, further studies are required to clarify how ncRNAs regulate the expression of Ubx.

In the human genome, there are 4 clusters of 39 HOX genes, called *HOXA – D*, on 4 different chromosomes. A bioinformatic study identified several antisense non-coding transcripts from human and mouse HOX clusters (Sessa et al. 2007). The expression of several HOXA transcripts were correlated with nearby HOXA genes in tissues as well as in an *in vitro* differentiation system (Sessa et al. 2007). A recent more in depth study of transcripts from the Human *HOX* loci led to identification of *HOTAIR* (*HOX Antisense Intergenic RNA*) in addition to several hundred more ncRNAs (Rinn et al. 2007). *HOTAIR* is a 2.2-kb ncRNA originating from the *HOXC* locus. Although *HOTAIR* expression correlates with HOXC genes, knock-down of *HOTAIR* results in derepression of the *HOXD* locus, accompanied by the clearance of PcG proteins and the loss of H3K27 trimethylation (Rinn et al. 2007). By showing that *HOTAIR* interacts with the PcG protein SUZ12, it was suggested that *HOTAIR* can regulate HOX genes even on a different chromosome *in trans* (Rinn et al. 2007). Although many HOX ncRNAs were identified, only a few of them have been shown to regulate the expression of HOX genes via various mechanisms including transcription interference, chromatin remodeling, producing miRNAs, and so on. Future studies will address the importance of other ncRNAs and bring us a more complete picture of how HOX genes are regulated, especially by ncRNAs.

Non-coding RNAs are key players in imprinting

Genomic imprinting denotes that the expression of genes is determined by the parent-of-origin (for review, see Sha 2008). An imprinting domain often includes a set of several protein-coding genes under the same regulation. Interestingly, ncRNAs have been suggested to play a critical role in establishing monoallelic expression of surrounding protein-coding genes (for review, see O'Neill 2005; Peters and Robson 2008). Among them, only a few including *Kcnq1 overlapping transcript1* (*Kcnq1ot1*) and *Airn* (*Antisense Igf2r RNA*, also known as *Air*) ncRNAs have proven functions in animal models.

The 1-Mb *Kcnq1* domain maps to the distal end of the mouse chromosome 7, which harbors eight maternally expressed protein-coding genes and one paternally expressed ncRNA, *Kcnq1ot1*. The imprinting control region (ICR) within the 10^{th} intron of the *Kcnq1* gene is methylated exclusively in the maternal allele, resulting in expression of *Kcnq1ot* from the paternal allele (Lee et al. 1999b; Smilinich et al. 1999). Deletion of this ICR caused loss of imprinting (Fitzpatrick et al. 2002). *Kcnq1ot1* is ~90-kb ncRNA transcribed from the ICR and partially overlaps with the *Kcnq1* gene in an antisense manner. Interestingly, the length of *Kcnq1ot1* is critical for silencing of nearby genes (Kanduri et al. 2006; Mancini-Dinardo et al. 2006). Recently a 890-nt fragment close to the 5' end of the *Kcnq1ot1* transcript has been assigned as a silencing domain (SD) that can silence flanking genes in an orientation-dependent and position-independent manner (Mohammad et al. 2008). This SD, more specifically the 12-nt A2 motif, was the only functional entity identified within the 5' 50-kb of *Kcnq1ot1* (Mohammad et al. 2008). Surprisingly, two different silencing mechanism were utilized by *Kcnq1ot1*: a

downstream locus was silenced by transcription interference and an upstream locus was silenced via chromatin remodeling (Mohammad et al. 2008). The *Kcnq1ot1* transcript was recently shown to interact with PRC2 complex members like Ezh2, Suz12, and G9a histone methyltransferase (Pandey et al. 2008). As observed in the case of *Xist* ncRNA (Zhang et al. 2007), this SD can induce perinucleolar localization of a reporter episome during mid S phase, suggesting a mechanism by which imprinted status is preserved during DNA replication (Mohammad et al. 2008). The human *KCNQ1OT1/LIT1* transcript has also been shown to coat the neighboring region of chromatin (Murakami et al. 2007).

The imprinted *Igf2r/Airn* locus includes three maternally expressing protein-coding genes, *Igf2r*, *Slc22a2*, and *Slc22a3* (Zwart et al. 2001). *Airn* ncRNA is an extremely long RNA pol II transcript, ~110-kb, and is expressed only from the unmethylated paternal allele (Lyle et al. 2000). The promoter for *Airn* exists within the second intron of *Igf2r* gene. Truncation of *Airn* ncRNA by premature polyadenylation results in the loss of imprinting, suggesting that transcription of downstream sequences is essential (Sleutels et al. 2002). *Airn* is retained in nucleus (Seidl et al. 2006) and shown to preferentially accumulate at the *Slc22a3* promoter (Nagano et al. 2008). *Airn* interacts with G9a histone methyltransferase, leading to methylation and silencing of the paternal *Slc22a3* promoter, but not on *Slc22a* or *Igf2r* promoters, suggesting more than one silencing mechanisms may be emplyed at this locus (Nagano et al. 2008). It is not yet known how transcription of *Airn* can silence other two protein-coding genes, Slc22a and Igf2r. Moreover, imprinted expression of the *Igf2r/Airn* locus is not observed in mouse brain; Igf2r is expressed from both alleles in mouse brain where *Airn* also is expressed

(Hu et al. 1999). We are just beginning to understand how this epigenetic phenomenon is regulated by ncRNAs through studies of *Airn* and *Kcnq1ot1*. Interestingly, their mode of action seems very similar; differentially methylated ICR, antisense ncRNAs, chromatin remodeling of upstream genes, and probably transcription interference of downstream genes. It is tempting to speculate as to whether the cases of *Airn* and *Kcnq1ot1* provide us with a general mechanism of imprinting.

When the act of non-coding RNA transcription alone may be enough

The transcription of ncRNAs can have profound consequences on the ability of nearby genes to be expressed (Katayama et al. 2005). For example, transcription of a ncRNA across the promoter region of a downstream protein-coding gene can directly interfere with transcription factor binding, and thus prevent the protein-coding gene from being expressed (Martens et al. 2004). In fact, transcriptional interference mechanisms have been shown to regulate key developmental decisions, such as where Hox genes are expressed (see above) and whether *S. cerevisiae* enters into meiosis (Hongay et al. 2006). Even if not directly interfering with a nearby promoter, transcription of non-coding RNAs can induce histone modifications that repress transcription initiation of overlapping protein-coding genes, as demonstrated at the yeast *GAL1-10* gene cluster (Houseley et al. 2008). This is because transcriptional elongation causes histone marks to be added that prevent spurious transcription initiation from sites within the body of the transcript (Carrozza et al. 2005). Non-coding transcription has even been shown to induce the formation of heterochromatin at the *p15* tumor suppressor gene locus that persisted after non-coding transcription was turned off, suggesting that the transient expression of

ncRNAs can have long-lasting heritable effects on gene expression (Yu et al. 2008).

Transcription of long ncRNAs from two upstream activation sequences (UAS) of the *S. pombe fbp1*$^{+}$ locus induces chromatin remodeling that is critical for transcriptional activation of the downstream protein-coding gene (Hirota et al. 2008). Interestingly, ncRNA transcription was found to initiate in a stepwise manner from 5' to 3' along the *fbp1*$^{+}$ UASs towards the promoter, causing the chromatin to progressively convert from a closed to open configuration. The insertion of a transcriptional terminator within these ncRNAs prevents downstream chromatin remodeling, resulting in reduced transcription factor binding at the *fbp1*$^{+}$ promoter, thereby minimal induction of the mRNA. The same group showed a similar stepwise remodeling of chromatin by ncRNAs at the *S. pombe ade6-M26* locus (Hirota and Ohta 2009).

The question now is whether chromatin remodeling occurs due to the act of upstream transcription (implying that the ncRNAs are simply non-functional by-products) or whether the ncRNAs themselves actively play a role, e.g. by recruiting chromatin remodeling or histone modifying enzymes. At least at the yeast *PHO5* locus, it appears to be the act of non-coding transcription rather than the ncRNA itself that contributes to rapid kinetics of chromatin remodeling (Uhler et al. 2007). The 2.4-kb antisense ncRNA originating from the 3' end of the *PHO5* locus is unstable and expressed only under phosphate rich condition (Uhler et al. 2007). Although the expression of this ncRNA *in trans* had no effect, a premature termination of this antisense transcript impedes histone eviction at the *PHO5* locus (Uhler et al. 2007). Therefore, it was suggested that the act of transcription affects the local rate of nucleosome exchange and/or turnover, allowing nucleosome eviction (and thus *PHO5* transcription) to occur

much more rapidly in response to phosphate starvation (Uhler et al. 2007). It should be noted that just because a transcript is rapidly degraded does not mean that it is non-functional, as unstable transcripts have, for example, been shown to repress transcription of the yeast Ty1 retrotransposon *in trans* (Berretta et al. 2008).

Therefore, depending on the gene locus, non-coding transcription can have profound effects, both negatively and positively, on the ability of neighboring protein-coding genes to be expressed. In some cases, the act of transcription is sufficient to have functional consequences, but it is likely that many of the ncRNAs produced may play yet-to-be-identified active regulatory roles.

Long non-coding RNAs modulate the activity of protein partners

Long ncRNAs can serve as key coactivators of proteins involved in transcriptional regulation. The *Evf-2* ncRNA, which is transcribed from an ultraconserved region, forms a complex with the homedomain-containing protein Dlx2 (Feng et al. 2006). Using reporter-based assays, it was shown that Dlx2 acts as a transcriptional enhancer only when the *Evf-2* ncRNA is also present. Similarly, the ncRNA *HSR1* (*heat shock RNA-1*) forms a complex with HSF1 (heat-shock transcription factor 1) enabling the transcription factor to induce expression of heat-shock proteins during the cellular heat-shock response (Shamovsky et al. 2006), and the ncRNA *SRA* (*steroid receptor RNA activator*) functions as a transcriptional coactivator of steroid receptors (Lanz et al. 1999). Conversely, non-coding transcripts derived from SINEs (short interspersed elements) bind to RNA polymerase II during heat shock to inhibit transcription of other mRNAs (Allen et al. 2004; Espinoza et al. 2004; Mariner et al. 2008).

Non-coding RNAs produced from the *cyclin D1* (*CCND1*) promoter region have recently been shown to function as allosteric effectors of an RNA-binding protein, TLS (Translocated in Liposarcoma) (Wang et al. 2008). TLS is a RNA binding protein that has been implicated with several cellular processes including transcriptional regulation (Uranishi et al. 2001), RNA splicing (Yang et al. 1998), and DNA repair (Kuroda et al. 2000). These ncRNAs are variable in their lengths and of low abundance (generally less than 2 copies per cell), but are induced in response to DNA damage and remain bound to the chromatin in the *CCND1* promoter region (Wang et al. 2008). Upon binding these ncRNAs, the TLS protein changes from an inactive to an active conformation, such that it binds and inhibits the enzymatic activities of histone acetyltransferases including CREB-binding protein (CBP) and p300 (Wang et al. 2008). By inhibiting active chromatin marks from being placed at the *CCND1* promoter, TLS bound to the ncRNAs inhibits transcription of *CCND1* (Wang et al. 2008).

At the human *DHFR* locus, a majority of the human dihydrofolate reductase (DHFR) mRNA originates from the downstream major promoter while the upstream promoter produces multiple minor ncRNA isoforms (Masters and Attardi 1985). Minor ncRNAs directly interact with and sequester transcription factor Sp3, thereby inhibiting of the formation of a transcription pre-initiation complex at the major promoter (Blume et al. 2003). A more recent report further confirmed an inhibitory role of the minor transcript (Martianov et al. 2007). The repression of the major promoter depends on whether the minor transcripts contain the entire major promoter sequences, but not on its length or whether it was produced *in cis* or *in trans* (Martianov et al. 2007). Interestingly, the major promoter and minor ncRNAs form a DNA-RNA triplex, suggesting that

ncRNAs can guide promoter-specific repression by dissociating the preinitiation complex from the major promoter (Martianov et al. 2007).

Long ncRNAs have also been shown to modulate the activity of proteins by regulating their subcellular localization. The transcription factor NFAT (Nuclear Factor of Activated T cells) localizes to the cytoplasm until calcium-dependent signals cause it to be imported into the nucleus, where it activates transcription of target genes (for review, see Hogan et al. 2003). One of the key regulators of NFAT trafficking happens to be a ncRNA known as *NRON* (*Non-coding Repressor of NFAT*) that is alternatively spliced (Willingham et al. 2005). By binding to members of the nucleocytoplasmic trafficking machinery, *NRON* specifically inhibits the nuclear accumulation of NFAT, but not that of other transcription factors such as p53 and NFκB (Willingham et al. 2005).

Long non-coding RNAs as precursors for small RNAs

Recent genome-wide studies suggest that the function of a significant fraction of long unannotated transcripts may be to serve as precursors for small RNAs less than 200 nucleotides in length (Fejes-Toth et al. 2009; Kapranov et al. 2007a). In addition to well characterized miRNAs (Cai et al. 2004; Lee et al. 2004) and piRNAs (for review, see Aravin et al. 2007), there are many more small RNAs whose functions and mechanisms of biogenesis are less clear. For example, small RNAs have been found to cluster near the 5' and 3' ends of genes (Core et al. 2008; Fejes-Toth et al. 2009; He et al. 2008; Kapranov et al. 2007a; Neil et al. 2009; Preker et al. 2008; Seila et al. 2008; Xu et al. 2009). Transfection of RNA mimetics to promoter-associated small RNAs (PASRs) were found to reduce expression of the corresponding mRNA promoter, indicating that these newly

identified small RNAs may impact on gene expression (Fejes-Toth et al. 2009).

It now appears that many protein-coding mRNAs and long ncRNAs can be post-transcriptionally processed to yield many small RNAs that, curiously, have a 5' cap (Fejes-Toth et al. 2009). Interestingly, significant numbers of CAGE (cap-analysis of gene expression) tags were found in exonic regions, and in some cases, even across splice junctions, suggesting they have arisen from at least partially processed mRNAs. Therefore, it has been proposed that long transcripts are processed post-transcriptionally to yield small RNAs, which are then modified by the addition of a cap structure (Fejes-Toth et al. 2009) (Figure 1.2A).

A recent work provides a clear example of how a nascent transcript can be processed to yield two ncRNAs that localize to distinct subcellular compartments (Wilusz et al. 2008). *MALAT1*(*Metastasis Associated in Lung Adenocarcinoma Transcript 1*) is a long (~7-kb) ncRNA that is misregulated in many human cancers (Ji et al. 2003; Lin et al. 2007). Searching for small RNAs mapping to the *MALAT1* locus led to the identification of a highly conserved 61-nt tRNA-like small RNA, *mascRNA* (*MALAT1-associated small cytoplasmic RNA*) (Wilusz et al. 2008). In contrast to the long nuclear retained *MALAT1* transcript, *mascRNA* is exclusively localized to the cytoplasm (Wilusz et al. 2008).

Interestingly, *mascRNA* is generated via processing of the *MALAT1* nascent transcript by members of the tRNA processing machinery (Wilusz et al. 2008) (Figure 1.2B). RNase P recognizes the tRNA-like structure in the nascent RNA polymerase II transcript and then cleaves to simultaneously generate the 3' end of the mature *MALAT1* transcript and the 5' end of *mascRNA*. Additional enzymes involved in tRNA biogenesis, RNase Z and the CCA-adding enzyme, then further process the small RNA prior to its

export to the cytoplasm. Based on the genome-wide studies (Fejes-Toth et al. 2009; Kapranov et al. 2007a), it is likely that many loci generate multiple non-coding transcripts that may not fit into well-characterized classes of ncRNAs.

The RNA interference (RNAi) machinery has well-characterized roles in the generation of miRNAs and siRNAs that regulate gene expression post-transcriptionally (for review, see Jaskiewicz and Filipowicz 2008; Lee et al. 2006). However, a recent report implicates these enzymes in the processing of long transcripts to small RNAs that likely do not function as miRNAs. In mice, a 2.4-kb unspliced, polyadenylated ncRNA known as *mrhl* is processed by Drosha to yield an 80-nt small RNA (Ganesan and Rao 2008). Unlike Drosha products, the 80-nt transcript is not further processed by Dicer *in vivo*, probably because it is retained in the nucleus.

miRNAs and piRNAs have received the most attention of late. However, it is becoming increasingly clear that long transcripts are processed to yield many other classes of small RNAs with likely very different and unique functions.

Long non-coding RNAs affect the processing of other RNAs

Long ncRNAs can be processed to yield small RNAs, but they can also affect how other transcripts are processed, for example by modulating their ability to be cut into small RNAs or changing their pre-mRNA splicing patterns. The *C. elegans rncs-1* (*RNA non-coding, starvation up-regulated*) ncRNA inhibits the production of small RNAs from other transcripts *in trans*. *rncs-1* is an 800-nt, spliced and polyadenylated transcript that is induced in response to starvation (Hellwig and Bass 2008). Despite of a double-stranded helix of ~300-bp, Dicer can not process *rncs-1* due to its secondary structure outside the

helix. Instead, *rncs-1* functions *in trans* to inhibit Dicer activity. Upon over-expressing or deleting *rncs-1*, the expression levels of certain siRNAs were found to decrease or increase, respectively, with a corresponding change in the mRNA levels of their targets (Hellwig and Bass 2008). Therefore, it has been proposed that *rncs*-1 binds to Dicer or accessory double-stranded RNA binding proteins to compete with other double-stranded RNAs involved in gene silencing.

Although there are no known examples to date, it is likely that certain long ncRNAs are able to base pair with small RNAs to modulate their activities. For example, by interacting with miRNAs, long non-coding transcripts could competitively inhibit the ability of microRNAs to interact with their mRNA targets, analogous to how artificial miRNA sponges function (Ebert et al. 2007).

Several recent reports indicate that transcripts derived from pseudogenes can surprisingly cause mRNAs from the functional protein-coding copy of the gene to be processed to small RNAs (Tam et al. 2008; Watanabe et al. 2008). This is because antisense transcripts from pseudogenes hybridize to their corresponding spliced mRNAs, resulting in the formation of double-stranded RNAs that are cleaved by Dicer to endogenous small interfering RNAs (endo-siRNAs). Endo-siRNAs may direct RISC (RNA-induced silencing complex) to cleave additional copies of the mRNA transcript. These results thus show that pseudogenes are not simply non-functional elements that eventually will be lost, but instead are key regulators of gene expression when transcribed.

Like pseudogenes, some natural antisense transcripts (NATs) are able to hybridize to overlapping genes and generate endo-siRNAs (Czech et al. 2008; Ghildiyal et al. 2008; Okamura et al. 2008; Watanabe et al. 2008). In addition, there are numerous examples of

NATs modulating the alternative splicing patterns of their overlapping genes (Krystal et al. 1990; Munroe and Lazar 1991; Yan et al. 2005), such as at the *Zeb2/Sip1* gene locus. Zeb2/Sip1 is a transcriptional repressor of E-cadherin whose expression is tightly regulated during epithelial-mesenchymal transition (EMT) (Beltran et al. 2008). An internal ribosome entry site (IRES) is utilized to translate the Zeb2/Sip1 protein; however, in epithelial cells, the IRES is spliced out of the mature mRNA. Upon EMT, a NAT complementary to the 5' splice site of this intron is produced, thus blocking the splicing of the IRES and enabling expression of the Zeb2/Sip1 protein (Beltran et al. 2008). In summary, long ncRNAs are able to base pair with other transcripts and affect RNA processing patterns.

Long non-coding RNAs serve as structural RNAs

RNA has also been found to serve a structural role in the organization and maintenance of the cellular cytoskeleton as well as the mitotic spindle. In *Xenopus* oocytes, the proper organization of the cytokeratin cytoskeleton is dependent on two RNAs, the *Xlsirts* (*Xenopus laevis short interspersed repeat transcripts*) ncRNA and the *VegT* mRNA, which are integrated within the cytoskeleton (Kloc et al. 2007; Kloc et al. 2005). Depletion of either transcript using antisense oligonucleotides disrupts the cytokeratin network, but not the actin cytoskeleton. Interestingly, although *VegT* transcript is a protein-coding mRNA, blocking its translation had no effect on the cytokeratin network (Heasman et al. 2001; Kloc et al. 2005), arguing that the RNA, itself, is functioning to maintain the cytoskeleton. Likewise, a large number of RNAs including 6S ribosomal RNAs, have been found to associate with the mitotic spindle (Blower et al.

2005). RNase A treatment disrupts spindle assembly and causes the spindle to collapse, although treatment with translation inhibitors has no effect, arguing that these RNAs play a translation-independent role in spindle assembly in M phase. However, it has not been determined which transcripts provide structural integrity to spindle yet. Considering the great variety of mRNA localization patterns that are observed during early *Drosophila* embryogenesis (Lecuyer et al. 2007), it is tempting to speculate that many more RNAs (especially ncRNAs) may have structural and organization roles in the cell.

Non-coding RNAs and diseases

Numerous long ncRNAs are misregulated in various diseases, especially cancer (for review, see Costa 2005; Prasanth and Spector 2007). For, example, *MALAT1* was originally identified in individuals exhibiting a high risk for metastasis of non-small cell lung tumor (Ji et al. 2003). More recently, *MALAT1* ncRNA was also shown to be present at higher levels in many other cancers, including uterine endometrial stromal carcinoma and hepatocellular carcinoma (Lin et al. 2007; Yamada et al. 2006). Increased expression of another ncRNA, *PCA3* (prostate cancer antigen 3, also known as *DD3*), has been observed in individuals with prostate cancer (Bussemakers et al. 1999). Later, PCA3 has been found to be a very sensitive and specific marker of prostate tumor (de Kok et al. 2002). However, the mechanisms by which these transcripts affect tumor initiation and/or progression are currently unknown. Long ncRNAs thus remain a relatively unexplored area in disease research, which may allow us to identify new therapeutic targets. Recent work on Alzheimer's disease has identified a ncRNA antisense to the *β-secretase* (*BACE1*) gene, which generates amyloid beta (Aβ), that may aid in driving the disease

(Faghihi et al. 2008). The ncRNA is induced in response to numerous cell stressors, including serum starvation and Aβ peptides, and, unfortunately, increases the stability of the *BACE1* mRNA, thus leading to even more Aβ peptides and the deleterious feed-forward cycle of disease progression. However, treatment with siRNAs against the ncRNA reduces the levels of Aβ peptides, suggesting that this non-coding transcript may serve as an attractive drug target candidate for Alzheimer's disease.

Not so fast - Short ORFs can be translated

Many RNA transcripts are classified as "non-coding" on the basis of not having open reading frames (ORFs) longer than 50 to 100 amino acids. However, the *Drosophila tarsal-less* (*tal*) gene provides a telling example that this assumption cannot be applied to all transcripts with relatively short ORFs. The *tal* mRNA is a 1.5-kb transcript that contains only ORFs of less than 50 amino acids and was, therefore, originally classified as a long non-coding RNA. However, several 33-nt ORFs within the *tal* gene are actually translated into 11 amino acid long peptides that control key tissue morphogenesis and pattern formation events during *Drosophila* development (Galindo et al. 2007; Kondo et al. 2007; Pueyo and Couso 2008). Due to practical and statistical reasons, ORFs as short as these in *tal* are generally systematically eliminated from gene annotations, but clearly need to be considered when addressing the function of an unannotated transcript. Additionally, some ncRNA genes, such as *SRA*, appear to yield multiple RNA isoforms, some of which can be translated (for review, see Leygue 2007), thus allowing a gene to have functions carried out by both RNA and protein.

The examples of ncRNAs discussed above indicate that these transcripts are not just transcriptional noise, but are able to provide an additional layer of regulation to the biological systems. Although genome-wide studies have identified a tremendous amount of ncRNAs in the mammalian genomes, we have only a handful of examples of functional ncRNAs. Characterization of additional ncRNAs will provide us with many more insights into the functions of ncRNAs.

To systemically understand the regulation of ncRNAs upon various differentiation stimuli, the Mattick group in Australia designed a custom microarray. In the case of C2C12 myoblast differentiation into myotubes, 184 ncRNAs exhibited more than 2-fold change in their cellular levels. Through collaboration with the Mattick group, I started to characterize some of these ncRNAs. After an initial localization study of 14 ncRNAs, I focused on *Men ε/β* ncRNAs that are up-regulated 3.3-fold upon myoblast differentiation into myotubes. These ncRNAs exhibit a punctate localization in the nucleus.

In following chapters, I discuss about the characterization of *MEN ε/β* ncRNAs (Chapter 2) and the development of mouse monoclonal antibodies to PSPC1, a protein associated with these ncRNAs (Chapter 3). Two ncRNA isoforms *Men ε* and *Men β* are produced from a single RNA polymerase II promoter but differ in the location and processing of their 3' ends. The 3' end of the *Men* β transcript is generated by RNase P. I demonstrate that *MEN ε/β* ncRNAs are localized to paraspeckles and that the depletion of *MEN ε/β* transcripts resulted in the disruption of nuclear paraspeckles. In chapter three, I describe how monoclonal antibodies to paraspeckle protein were developed and characterized. One of them, hybridoma 9-99, recognizes Nono in mouse cells and NONO/PSPC1 in human cells. This antibody was extensively utilized during my study.

Antibodies developed during this study can be utilized to label paraspeckles and to identify additional paraspeckle components. This study provides new insights into a role for ncRNAs as structural components of a specific nuclear domain and contributes to our current understanding of the functional and structural organization of the mammalian cell nucleus.

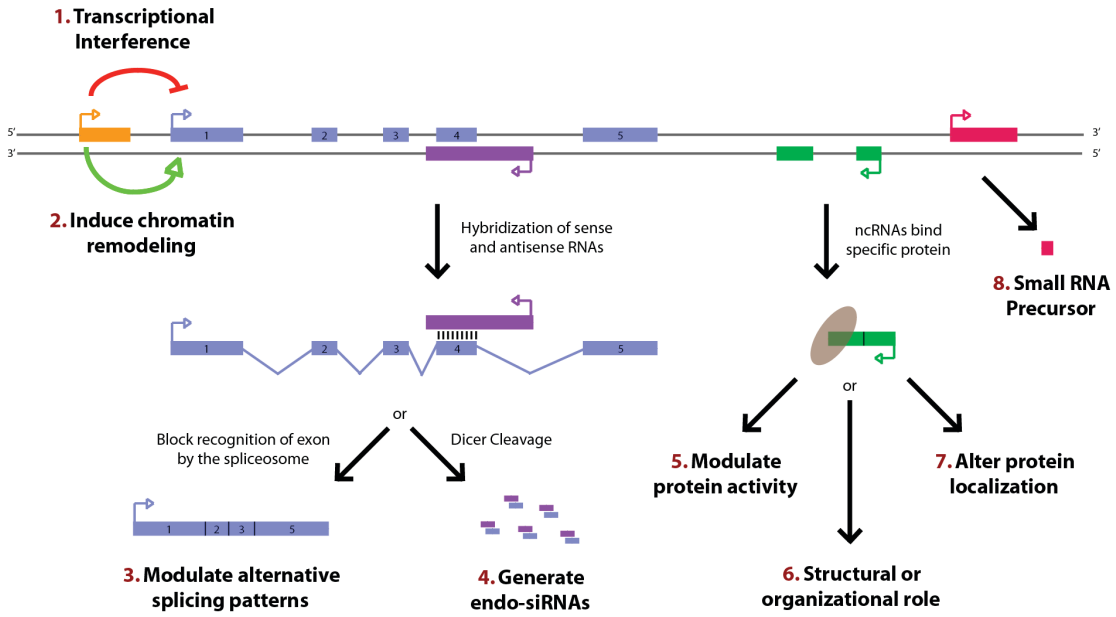


Figure 1.1. Paradigms for how long non-coding RNAs function. Recent studies have identified a variety of regulatory paradigms for how long ncRNAs function, many of which are highlighted here. Transcription from an upstream non-coding promoter (orange) can (1) negatively or (2) positively affect expression of the downstream gene (blue) by inhibiting RNA polymerase II recruitment or inducing chromatin remodeling, respectively. An antisense transcript (purple) is able to hybridize to the overlapping sense transcript (blue) and (3) block recognition of the splice sites by the spliceosome, thus resulting in an alternatively spliced transcript. Alternatively, hybridization of the sense and antisense transcripts can (4) allow Dicer to generate endogenous siRNAs. By binding to specific protein partners, a non-coding transcript (green) can (5) modulate the activity of the protein, (6) serve as a structural component that allows a larger RNA-protein complex to form, or (7) alter where the protein localizes in the cell. Long ncRNAs (pink) can (8) be processed to yield small RNAs, such as microRNAs, piRNAs, and other less well-characterized classes of small transcripts. From: Wilusz et al. 2009

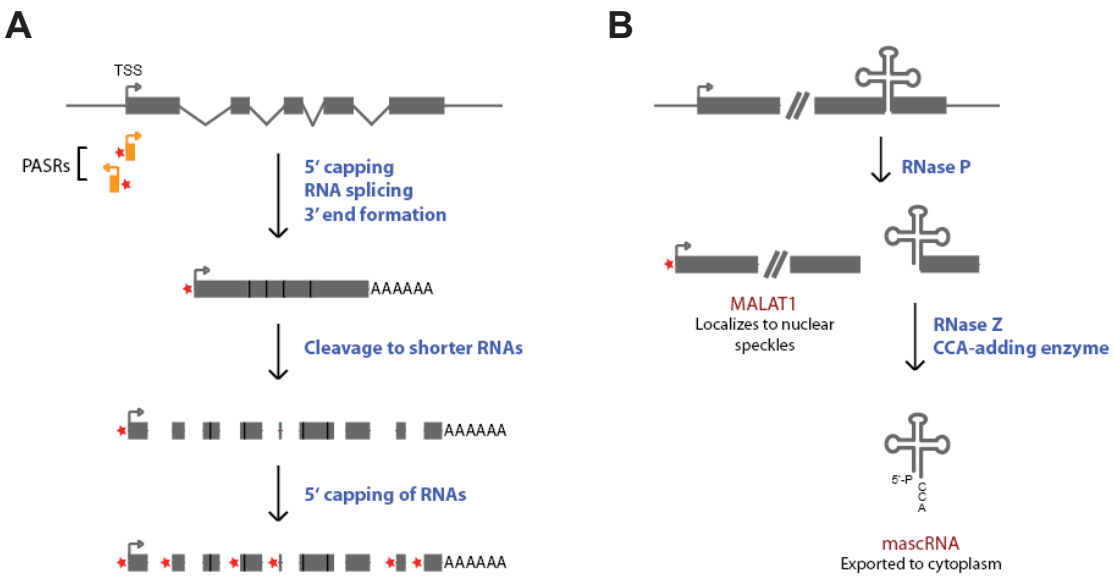


Figure 1.2. Long non-coding RNAs are processed to yield small RNAs. (A) Recent work by Fejes-Toth and colleagues (2009) suggests that many long processed transcripts can be post-transcriptionally cleaved to generate small RNAs. A cap structure (denoted by red star) is then added to the 5' ends of many of these small RNAs. In addition, capped small RNAs, known as PASRs, map near the transcription start site (TSS) of many genes. **(B)** The nascent *MALAT1* transcript is processed to yield two non-coding RNAs that localize to different subcellular compartments. Cleavage by RNase P simultaneously generates the 3' end of the mature *MALAT1* transcript and the 5' end of *mascRNA*. From: Wilusz et al. 2009

II. Characterization of MEN ε/β long non-coding RNAs

Summary

Studies of the transcriptional output of the human and mouse genomes have revealed that there are many more transcripts produced than can be accounted for by predicted protein-coding genes. Using a custom microarray, we have identified 184 non-coding RNAs that exhibit more than 2 fold up- or down-regulation upon differentiation of C2C12 myoblasts into myotubes. Here, we focus on the *Men ε/β* locus, which is up-regulated 3.3 fold during differentiation. Two non-coding RNA isoforms are produced from a single RNA polymerase II promoter, differing in the location of their 3' ends. *Men ε* is a 3.2-kb polyadenylated RNA, whereas *Men β* is a ~20-kb transcript containing a genomically encoded poly(A)-rich tract at its 3' end. The 3' end of *Men β* is generated by RNase P cleavage. The *Men ε/β* transcripts are localized to nuclear paraspeckles and directly interact with Nono. Knock-down of *MEN ε/β* expression results in the disruption of nuclear paraspeckles. Furthermore, the formation of paraspeckles, after release from transcriptional inhibition by DRB treatment, was suppressed in *MEN ε/β* depleted cells. Our findings indicate that the *MEN ε/β* non-coding RNAs are essential structural/organizational components of paraspeckles.

Introduction

Sequencing of the human and other mammalian genomes has revealed the number of protein-coding genes to be in the range of 20,000 – 25,000 (International Human Genome Sequencing Consortium 2004; Waterston et al. 2002), representing less than 2% of the total genomic sequence. However, recent studies of the mammalian transcriptome have shown that the majority of the genome is transcribed and that many transcripts lack protein-coding capacity (Birney et al. 2007; Carninci et al. 2005; Kapranov et al. 2007a). Such analyses have prompted considerable discussion as to whether these non-coding RNAs (ncRNAs) simply represent transcriptional noise or are involved in cellular functions (for review, see Mattick and Makunin 2006). Interestingly, large-scale studies of ncRNAs have shown that many are dynamically regulated during differentiation and exhibit cell- and tissue-specific expression patterns (Dinger et al. 2008; Mercer et al. 2008; Ravasi et al. 2006). These observations support the contention that ncRNAs are likely to have functional roles in the cell, some of which may serve in regulatory and/or structural paradigms (for review, see Mattick 2004).

Although the number of ncRNAs identified has increased exponentially, very few ncRNAs have thus far been assigned a cellular function (for review, see Costa 2005; Prasanth and Spector 2007). Interestingly, several ncRNAs have been shown to be involved in the regulation of the transcriptional state of a locus or at the level of a single chromosome. For example, expression of the long (>100kb) ncRNA *Airn* (Antisense *Igf2r* RNA, also known as *Air*) is associated with silencing of the *Igf2r*, *Slc22a2* and *Slc22a3* genes in mice (Sleutels et al. 2002). In another case, it was suggested that the transport of nuclear factor of activated T cells (NFAT) into the nucleus is repressed by

NRON (*non-coding repressor of NFAT*) ncRNAs, a series of transcripts ranging in size between 0.8 – 3.7-kb (Willingham et al. 2005). In fission yeast, a recent report argued that transcription of ncRNAs at the promoter region can induce chromatin remodeling at the *fbp1*$^{+}$ locus (Hirota et al. 2008). Perhaps the best studied ncRNAs are *Xist* (*X-inactive specific transcript*) and *Tsix* (*X-inactive specific transcript, antisense*), key players in dosage compensation of the mammalian X chromosome. In females, *Tsix* is an antisense transcript of *Xist* and its expression determines which X chromosome will be inactivated for dosage compensation (for review, see Erwin and Lee 2008; Heard and Disteche 2006). At the stage of X inactivation in mammalian development, *Xist/XIST* RNA, 15 – 17-kb in mouse and ~19-kb in human, respectively, is transcribed from one of the two X chromosomes. This ncRNA subsequently coats the chromosome from which it is transcribed and represents part of the mechanism by which transcriptional inactivation of the coated chromosome is achieved (for review, see Heard and Disteche 2006; Plath et al. 2002).

Determining the subcellular localization of ncRNAs is important for providing insights into their potential partners and the functional pathways with which they interact. The mammalian nucleus is a highly organized organelle with a number of membraneless sub-compartments such as Cajal bodies, nucleoli, paraspeckles, PML bodies, and speckles, to name just a few (for review, see Spector 2001; Spector 2006). Among them, paraspeckles were initially discovered via the identification and characterization of a paraspeckle-associated protein in a proteomic analysis of human nucleoli (Fox et al. 2002). Paraspeckles generally appear in clusters and are frequently localized in the nucleoplasm close to nuclear speckles where pre-mRNA splicing factors are stored,

assembled, and/or modified. Inhibition of RNA polymerase II transcription results in its protein components to relocate to the periphery of nucleoli. As paraspeckles are sensitive to RNase A treatment (Fox et al. 2005; Prasanth et al. 2005), RNA(s) may be a critical component of this nuclear structure. Interestingly, Prasanth et al. (2005) identified *Cat2* transcribed nuclear RNA (*CTN-RNA*), a ~8-kb transcript that is encoded by the *Slc7a2* (also known as *mCat2*, mouse cationic amino acid transporter 2) gene and is retained in the nucleus in paraspeckles. *CTN-RNA* is produced via utilization of an alternative promoter coupled with the utilization of the distal most polyadenylation site, resulting in an extended 3' UTR. Upon stress, *CTN-RNA* is cleaved releasing the upstream *Slc7a2* ORF such that it can transit to the cytoplasm to be translated (Prasanth et al. 2005). By cleaving the "stored" nuclear retained RNA, and bypassing the need for initiating *Slc7a2* transcription during stress, the nuclear retained form of this RNA represents a rapid response mechanism for gene expression. Although representing an RNA component of paraspeckles, knock-down of *CTN-RNA* does not result in any change in the morphology of paraspeckles, suggesting that it does not confer any structural integrity to this nuclear domain (Prasanth et al. 2005). Therefore, although paraspeckles are sensitive to RNase A treatment, thus far, a specific RNA(s) involved in its structural integrity has not been identified.

In the present study, we utilized a custom microarray to identify ncRNAs that exhibit altered levels upon C2C12 myoblast differentiation into myotubes. One of the identified loci, the *Men1* (*multiple endocrine neoplasia 1*) locus on mouse chromosome 19qA, exhibited a 3.3-fold increased level of RNA in myotubes versus myoblasts. This locus produces two ncRNA isoforms: the mouse *Men epsilon (ε)* transcript is 3.2-kb and

the *Men beta (β)* transcript is ~20-kb in length. Recently, this locus was also shown to be up-regulated during bovine muscle development (Lehnert et al. 2007). RNA fluorescence *in situ* hybridization (FISH) revealed that these transcripts exhibit a punctate pattern in the nucleus that corresponds to paraspeckles. This distribution is consistent with a previous report of these RNAs localizing adjacent to nuclear speckles (Hutchinson et al. 2007). Interestingly, the *Men ε/β* transcripts directly interact with the Nono (also known as p54/nrb) complex, a known protein component of paraspeckles. Knock-down of these RNAs induced the disruption of paraspeckles, suggesting that the *MEN ε/β* transcripts are essential for the structural/organizational integrity of this nuclear domain. In addition, inhibition of RNA polymerase II transcription resulted in a redistribution of PSPC1 (also known as PSP1α) as well as the *MEN ε/β* transcripts without altering the level of these ncRNAs or protein, indicating that active transcription of the *MEN ε/β* locus is required to establish paraspeckles. In summary, we have identified a ncRNA locus that is up-regulated upon myoblast to myotube differentiation and whose RNA products are critical structural/organizational components of paraspeckles.

Results

Identification of ncRNAs that are differentially expressed upon myoblast differentiation into myotubes

We used a custom microarray to examine the expression profiles of 4,694 non-coding and 13,432 protein-coding RNAs during myoblast differentiation. RNA was isolated from C2C12 myoblasts at three stages of differentiation: 50% confluent myoblasts, 24 hours post-induction, and 5 days post-induction where the cells appear predominantly as fully differentiated myotubes. Across the three stages of differentiation, 8,442 coding and 1,425 non-coding transcripts were expressed above background, respectively. Of these transcripts, 1,814 (21.5%) coding and 184 (13.0%) non-coding transcripts were significantly differentially expressed (fold-change > 2; B-statistic > 3; Table 2.1).

After the 184 ncRNA candidates were manually examined in terms of overall expression levels, extent of differential expression, and the possibility of overlapping with any known protein-coding genes, 14 full-length ncRNA gene loci were selected for subcellular localization studies using NIH3T3 cells. Several ncRNAs exhibited distinct localization patterns when examined by RNA FISH (Figure 2.1). AK028745 showed a distinct punctate distribution in the nuclei of NIH3T3 cells. AK030860 is exclusively nuclear retained and homogenously diffused throughout nucleoplasm except nucleoli. AK036616 showed a few cluster of fine hot foci in nuclei. AK080054 are localized to ~2 foci per nuclei in ~30% population of cells, potentially transcription sites. AK085418 is marginally nuclear retained. Interestingly, AK028745 exhibited a similar punctate localization pattern in several mouse and human cell lines (Figure 2.2). While AK028745

was exclusively retained in the nucleus in the majority of cell lines, there was additional diffusive localization in the cytoplasm only in HEK293 cells (Figure 2.2B). In addition, punctate distribution pattern was observed during mitosis although less number of foci (Figure 2.3). Therefore I focused on further characterization of the AK028745 ncRNA.

Upon differentiation of myoblasts to myotubes, AK028745 was present at a 3.3-fold higher level based on our microarray analysis (Table 2.1). The punctate foci were enlarged and more abundant in myotube nuclei compared to myoblast nuclei (Figures 2.4A and B), consistent with the increased levels observed in our microarray results. To further confirm our microarray results, Q-PCR was performed (Figure 2.4C) using several primer sets around the mapped position of EST clone AK028745 (*Men ε/β* locus; see Figure 2.6A for primer positions). While primers targeted upstream of the AK028745 promoter failed to amplify a product as expected, primers to the non-coding transcripts showed that the expression of this locus increased 3.2 – 4.9 fold upon differentiation of myoblasts to myotubes (Figure 2.4C). As described below, two ncRNAs (*Men ε* and *Men β*) are generated from this locus and the expression of both transcripts increases upon differentiation.

The localization of *Men ε/β* ncRNA was also examined in mouse tissues including ovary and kidney. In contrast to tissue culture cells where the *Men ε/β* ncRNAs are expressed in every cell, the expression of *Men ε/β* ncRNAs varies by cell type in mouse ovary (Figure 2.5A). *Men ε/β* ncRNAs are not expressed in follicles where oogenesis occurs. However, cells in the corpus luteum, which are derived from follicles after ovulation, contain intense foci of *Men ε/β* ncRNA. In kidney, the expression of the *Men ε/β* was not observed in all cells (Figure 2.5B). These data suggest that in addition to their

potential role in muscle differentiation, the up-regulation of the *Men ε/β* ncRNAs may also be implicated in other developmental processes. However, future studies will be required to elucidate how the *Men ε/β* ncRNAs are regulated in tissues.

Two long non-coding RNAs are generated from the *Men ε/β* locus

EST clone AK028745 corresponds to a 3.2-kb ncRNA known as *Men ε* (also known as *Neat1*) (Hutchinson et al. 2007) that maps to an intergenic region on mouse chromosome 19qA (Figure 2.6A). No known genes exist within 23-kb downstream of *Men β,* a longer isoform of *Men ε* (see below). However, the protein coding gene *Frmd8* is located less than 6-kb upstream of the 5' end of *Men ε*. RNA FISH analysis using a probe to the coding region of *Frmd8* did not reveal a nuclear punctate pattern, implying that *Men ε* is not an unannotated exon of *Frmd8* (data not shown). Instead, *Men β* is transcribed from a separate transcriptional unit that has a CpG island in its promoter region (Figure 2.6A). The *Men ε* transcript is unspliced, polyadenylated, lacks repetitive elements, and is highly conserved among mouse, rat, and human (Figure 2.6A). Northern blot analysis as well as 5' and 3' RACE confirmed the previously annotated ends of the 3.2-kb *Men ε* transcript (Figure 2.7; data not shown). Using Northern blot analysis, we also determined that *Men ε* is broadly expressed in many mouse tissues (Figure 2.6B). The highest *Men ε* transcript levels were observed in kidney and lung (Figure 2.6B).

The human ortholog *MEN ε* was previously identified as one of several transcripts produced from the *MEN1* locus on human chromosome 11q13.1 (Guru et al. 1997). In addition to the 3.7-kb polyadenylated human *MEN ε*, a longer transcript *MEN β* was also previously identified (Guru et al. 1997). *MEN ε* and *MEN β* share the same RNA

polymerase II promoter, but differ in the location of their 3' ends. Upon analyzing the many human EST clones mapping to this genomic region, we determined that human *MEN β* is a ~23-kb unspliced transcript that contains many repetitive elements (Figure 2.8A). In contrast, there are only a few EST clones that map in a discontinuous manner to the mouse *Men β* genomic region (Figure 2.6A), and it has been previously argued that the mouse *Men β* isoform is not transcribed (Hutchinson et al. 2007). However, it was recently shown that transcriptional elongation marks (H3K36me3) extend through the *Men β* locus in mouse embryonic stem cells (Mikkelsen et al. 2007) where the *Men ε/β* ncRNAs are expressed (Dinger et al. 2008; data not shown). Indeed, we were able to detect expression of the ~20-kb mouse *Men β* transcript in C2C12 myoblast cells by RNA FISH (Figure 2.8) and RT-PCR (data not shown). The tissue expression profile of the mouse *Men β* transcript was examined by an RNase protection assay (Figure 2.6C). Like *Men ε*, *Men β* was broadly detected in mouse tissues, with the highest *Men β* transcript levels in colon and ovary (Figure 2.6C).

An annotated human microRNA, miR-612, maps near the 3' end of the *MEN β* transcript, suggesting that *MEN β* may function as a microRNA precursor. However, miR-612 is poorly conserved and has not previously been detected in mouse cells. In addition, we failed to detect human miR-612 expression by Northern blot analysis (data not shown), suggesting that miR-612 may be misannotated and that *MEN β* is not likely a microRNA precursor.

MEN ε/β are nuclear retained non-coding RNAs

RNA FISH data suggested that the *MEN ε/β* transcripts are localized to several

foci in cell nuclei (Figures 2.1 and 2.2). To confirm that *MEN ε/β* transcripts are not exported to the cytoplasm, Northern blot analysis was performed using RNA extracted from nuclear and cytoplasmic fractions in HeLa cells and HEK293 cells. As expected, the *MEN ε* and *MEN β* transcripts are completely retained in HeLa nuclei (Figure 2.9A). Although the *MEN β* transcript is not subject to a canonical polyadenylation (see below), *MEN β* was also detected in poly(A)$^+$ fraction due to a short genomically encoded poly(A) tract (Figure 2.9A). Interestingly, although the *MEN ε/β* transcripts were highly enriched in nuclear fraction of HEK293 cells, a small amount of *MEN ε* was also detected in the cytoplasmic fraction, suggesting that *MEN ε* transcripts may be regulated differently between HeLa and HEK293 cells (Figure 2.9C).

Although the *MEN ε/Men ε* transcript was presumed to be a non-coding RNA (Hutchinson et al. 2007), its coding potential has not yet been directly tested. Bioinformatic analysis of the *Men ε* sequences revealed several putative open reading frames (ORFs) that may encode peptides longer than 50 amino acids (Figure 2.10A). Since functional peptides as short as only 11 amino acids have been reported (Galindo et al. 2007), it was essential to test whether any of these putative ORFs can be translated in a cellular context. The three longest ORFs were chosen and EYFP sequences were inserted at the end of ORF but before the stop codon (Figures 2.10A and B). The presence of EYFP fluorescence in cells would thus indicate that the ORF in question had been translated into a fusion protein. As expected, no fusion constructs were able to produce any detectable EYFP fluorescence signals in cells while control transcripts (EV) were efficiently exported to the cytoplasm and translated into EYFP protein (Figure 2.10C). Each fusion transcript exhibited a wide range of localization (Figure 2.10C). Two fusion

transcripts containing *Men ε* 1 – 1857-nt and *Men ε* 1 – 2626-nt are retained in the nucleus completely and mostly, respectively. These data suggests that there is a nuclear retention motif between 1070 – 1857-nt. In contrast, the shortest fusion transcript containing *Men ε* 1 – 1069-nt was efficiently exported to the cytoplasm with a few intense foci in the nucleus. However, there was still no EYFP protein produced, suggesting an additional mechanism that prevents the *Men ε* transcripts from being translated other than nuclear retention of these transcripts.

The 3' end of *Men β* is generated by RNase P

MEN β was originally suggested to be an alternative polyadenylation isoform of *MEN ε* (Guru et al. 1997). By Northern blot analysis using multiple probes, we were able to detect *MEN β* expression in HeLa cells and roughly define the location of its 3' end (Figure 2.11A). Surprisingly, there are no canonical cleavage/polyadenylation signals located in the immediate vicinity of where *MEN β* ends (Figure 2.11C and 2.12), suggesting that the 3' end of *MEN β* is generated via a different mechanism. Indeed, when we carried out RNase H digestion followed by Northern blot analysis to better define the 3' end of *MEN β*, it was determined that *MEN β* 3' end cleavage occurs at a defined nucleotide position and that there is no subsequent addition of nucleotides as occurs during classical polyadenylation (Figure 2.11B). Rather than a classical poly(A) tail, the mature *MEN β* transcript has a short poly(A)-rich tract at its 3' end, which is genomically encoded (Figure 2.11C).

Upon searching the human and mouse genomes for sequences similar to the 3' end of *MEN β*, we found that it is similar to the 3' end of *MALAT1* (Figure 2.12). We

have recently shown that the 3' end of the abundant *Malat1* transcript is not generated by a classical cleavage/polyadenylation mechanism, but is instead generated by RNase P (Wilusz et al. 2008). A tRNA-like structure present within the primary *Malat1* transcript is recognized by RNase P, allowing the enzyme to cleave and generate the 3' terminus of the mature *Malat1* transcript (Wilusz et al. 2008). A similar evolutionarily conserved tRNA-like structure is present at the 3' end of the *MEN β/Men β* locus (Figures 2.12 and 2.13). To test if RNase P generates the mature 3' end of the *Men β* transcript, a 189-nt region encompassing the 3' end of *Men β* was cloned, transcribed *in vitro*, and employed for RNase P *in vitro* cleavage assays. Partially purified human RNase P was able to cleave *Men β in vitro* at the expected 3' end of mature *Men β*, showing that the *in vitro* system accurately recapitulates *in vivo Men β* processing (Figure 2.11D). The *Malat1* tRNA-like structure is further cleaved by RNase Z at its 3' end to yield a 61-nt tRNA-like transcript named *mascRNA* (*Malat1*-associated small cytoplasmic RNA) (Wilusz et al. 2008). The *Men β* tRNA-like structure can similarly be cleaved *in vitro* by recombinant human RNase Z (Figure 2.11E), suggesting that 3' end processing of the *Men β* transcript also yields a small tRNA-like transcript. While *mascRNA* is broadly expressed in tissues and cell lines (Figure 2.11F; Wilusz et al. 2008), the *MEN β/Men β* tRNA-like small RNA generally fails to accumulate to significant steady state levels. By Northern blot analysis, we were unable to detect expression of the *MEN β/Men β* tRNA-like small RNA in C2C12, EpH4-EV, or HeLa cells (Figure 2.11F; data not shown), but were able to detect expression in mouse liver. These data suggest that the *MEN β/Men β* tRNA-like small RNA may be selectively stabilized in a cell-type specific manner.

Men ε/β ncRNAs are localized to nuclear paraspeckles

The observed punctate distribution of *Men ε/β* ncRNAs was further examined to assess its co-localization with known nuclear domains. We found that *Men ε/β* and *MEN ε/β* ncRNAs are not enriched in PML bodies, Cajal bodies, or nuclear speckles in most of the mouse or human cells examined (Figure 2.14). In HeLa cells, the *MEN β* transcripts exhibited a punctate distribution as well as a less intense speckle localization (Figure 2.14B). However, consistent with the results of a recent study (Hutchinson et al. 2007), *MEN ε/β* commonly localize in domains that are adjacent to nuclear speckles, suggesting that they may be constituents of paraspeckles. We established a C2C12 cell line stably expressing PSPC1 (also known as PSP1α), a known paraspeckle component (Fox et al. 2002), fused to EYFP. To minimize cell-to-cell variation, clonal selection was performed to isolate a single cell line, designated C2C12 EYFP-PSPC1, for use in this study. By immunoblotting, we confirmed that endogenous PSPC1 is expressed in both wt C2C12 and C2C12 EYFP-PSPC1 cells (Figure 2.15A, lanes 1 and 2), whereas the EYFP-PSPC1 fusion protein is detected only in the C2C12 EYFP-PSPC1 cell line (Figure 2.15A, lanes 2 and 6). The level of endogenous Pspc1 is reduced in C2C12 EYFP-PSPC1 cells as compared to wt C2C12 cells (Figure 2.15A, lanes 1 and 2), indicating a potential cellular response to control the total level of Pspc1. To confirm that the C2C12 EYFP-PSPC1 cells allow accurate visualization of paraspeckles, a plasmid expressing mCherry fused to Nono (also known as p54/nrb), another known paraspeckle component (Fox et al. 2005), was transfected into C2C12 EYFP-PSPC1cells. As expected, EYFP-PSPC1 and mCherry-Nono co-localize in paraspeckles (Figure 2.15B).

Using the C2C12 EYFP-PSPC1 cells, we found that RNA FISH probes to *Men*

ε/β co-localize with EYFP-PSPC1, indicating that the *Men ε/β* ncRNAs are indeed localized to nuclear paraspeckles (Figure 2.15C). Since A-to-I hyper-editing within inverted repeats of the 3' UTR of CTN-RNA was shown to act as a nuclear retention signal (Prasanth et al. 2005), we examined *Men ε* transcripts for evidence of inverted repeats and/or hyper-editing. In contrast to *CTN-RNA*, the *Men ε* transcripts lack any repeats, while the long *Men β* ncRNA contains many repetitive elements. Upon searching for evidence of A-to-I hyper-editing in a mouse EST database, a single 106-nt region near the 3' end of *Men ε* contained within EST clone AA709912 exhibited a significant level of hyper-editing (11 out of 24 A were sequenced as a G, indicative of A-to-I editing). However, subsequent cDNA cloning from C2C12 cells failed to identify any additional evidence that the *Men ε* transcripts are A-to-I hyper-edited (data not shown).

Next, we investigated whether the *Men ε/β* transcripts interact with Nono. A mouse monoclonal antibody to Nono, designated 9-99, was generated (Figure 2.16A) and used for co-immunoprecipitation (Co-IP) experiments using C2C12 cells. Two other *bona fide* components of paraspeckles, Sfpq (also known as Psf) and Pspc1, were co-immnuoprecipitated along with Nono (Figure 2.16B), as has been reported previously (Fox et al. 2005; Kuwahara et al. 2006). The Co-IP fraction was next analyzed by RT-PCR, which confirmed that *Men ε/β* efficiently co-immunoprecipitated with Nono in C2C12 cells (Figure 2.16B) and in NIH3T3 cells (data not shown). Northern analysis showed that *Men ε* transcripts in fact were pulled down with the Nono complex while unrelated *β-Actin* mRNA did so minimally (Figure 2.16C). Collectively, our results demonstrate that both *Men ε/β* transcripts localize to paraspeckles and associate with Nono.

MEN ε/β ncRNAs are essential components of nuclear paraspeckles

It has previously been shown that RNase A treatment disrupts the integrity of nuclear paraspeckles (Fox et al. 2005; Prasanth et al. 2005), suggesting that one or more RNA species might serve a role in the structure and/or organization of these nuclear bodies. Thus far, the only other known RNA present in paraspeckles is *CTN-RNA*; however, knock-down of *CTN-RNA* by antisense oligonucleotides (ASOs) did not have any effect on the integrity of paraspeckles (Prasanth et al. 2005). Unlike CTN-RNA that showed paraspeckle localization only in a subset of cells (Prasanth et al. 2005), *MEN ε/β* ncRNAs are always localized to paraspeckles, even during mitosis (Figure 2.16). To address whether the *MEN ε/β* transcripts are essential for the structural integrity of nuclear paraspeckles, we depleted the level of *MEN ε/β* using ASOs in HeLa cells stably expressing EYFP-PSPC1.

Three ASOs (ASO 1, 2, and 3) were targeted near the 3' end of the *MEN ε* transcript such that they knock-down both the *MEN ε* and *MEN β* transcripts (Figure 2.17A). Similarly, one ASO (ASO 4) was targeted near the 3' end of the *MEN β* transcript such that it knocks-down only the *MEN β* isoform (Figure 2.17A). When these ASOs were independently transfected into HeLa EYFP-PSPC1 cells, approximately 70% knock-down of *MEN ε/β* (ASO 1, 2, or 3) and 50% knockdown of *MEN β* (ASO 4) was observed by Q-PCR analysis 24 hours post-transfection (Figure 2.17B). The knock-down of *MEN ε* and/or *β* transcripts by the complementary ASOs and not by a control ASO was confirmed by RNA FISH (Figure 2.17C). Interestingly, paraspeckles were disrupted in ~80% of cells transfected with ASOs 1, 2, or 3 to the *MEN ε/β* transcripts as visualized by the loss of EYFP-PSPC1 foci, while the control ASO had no effect on the integrity of

paraspeckles (Figures 2.17C and D). Knock-down of the *MEN β* transcript specifically by ASO 4 also caused the disruption of paraspeckles, although to a lesser extent (Figures 2.17C and D). A similar result was observed when wt HeLa cells were treated independently with ASOs 1 – 4 and paraspeckles were examined by immunofluorescence using the 9-99 monoclonal antibody that reacts with both PSPC1 and NONO from HeLa lysate (Figure 2.18). A small population of cells transfected with ASO 4 exhibited paraspeckles that were co-localized only with *MEN ε* transcripts, but not with *MEN β* transcripts, demonstrating that *MEN ε* transcripts unambiguously localize to paraspeckles and are sufficient for paraspeckle integrity (Figure 2.18B). These results show that both *MEN ε/β* transcripts are essential for the maintenance of nuclear paraspeckles.

One possible explanation for the observed disruption of paraspeckles is that the loss of the *MEN ε/β* transcripts causes the degradation of EYFP-PSPC1. To clarify whether the level of EYFP-PSPC1 proteins was changed after knock-down of the *MEN ε/β* transcripts, we performed immunobloting using an anti-PSPC1 antibody. While the level of *MEN ε/β* transcripts was reduced by transfection of the ASOs, PSPC1 remained relatively stable upon ASO treatment (Figure 2.17E), providing further support that *MEN ε/β* are involved in the structural organization of nuclear paraspeckles.

MEN ε/β transcripts are required for the *de novo* formation of paraspeckles

It was previously shown that upon RNA polymerase II inhibition, paraspeckles are re-localized to the periphery of nucleoli forming nucleolar caps (Fox et al. 2002). After HeLa cells were treated with 5, 6-dichlorobenzimidazole 1-β-D-ribofuranoside (DRB) for 1 hour, EYFP-PSPC1 indeed formed nucleolar caps (Figure

2.20A and 2.21A). To be certain that DRB treatment did not affect the phosphorylation state of PSPC1, its migration rate on SDS-PAGE was examined +/- alkaline phosphatase treatment. Alkaline phosphatase treatment caused a band shift (Figure 2.21B, lanes 2 and 3), suggesting that PSPC1 is phosphorylated. However, DRB treatment did not affect the phosphorylation state of PSPC1 (Figure 2.21B). Upon DRB treatment, the *MEN ε/β* transcripts failed to remain co-localized with EYFP-PSPC1 (Figure 2.20A). *MEN ε* transcripts now exhibited a diffuse nuclear localization, while *MEN β* transcripts became concentrated only at nuclear speckles (Figure 2.20A and 2.21C). Since the half life of *MEN ε/β* is ~4 and 8 hours (data not shown), respectively, the observed change in RNA localization resulted from redistribution rather than degradation. These data suggest that active transcription is required for *MEN ε/β* localization to paraspeckles. Upon release from RNA polymerase II inhibition, paraspeckles were reformed within 2 hours and contained the *MEN ε/β* transcripts (Figure 2.20A).

To next address whether the *MEN ε/β* transcripts are essential for the initial formation of paraspeckles, we depleted *MEN ε/β* expression in HeLa cells by ASO transfection followed by DRB treatment and release (Figure 2.20B). When the level of *MEN ε* or *MEN β* transcripts was assessed by Q-PCR 6 hours after ASO transfection, 40 – 80% knock-down of *MEN ε* and/or β expression was achieved by ASOs 1 – 4 (Figure 2.19C). The depletion of both *MEN ε/β* transcripts (by ASO 1, 2, and 3) or *MEN β* transcripts alone (by ASO 4) resulted in suppression of paraspeckle reformation after release from DRB treatment (Figures 2.19B and D), suggesting that the *MEN ε/β* transcripts are required for the reformation of paraspeckles. ASO 4, which exclusively targets *MEN β* transcripts, suppressed the reformation of nuclear paraspeckles, but to a

lesser degree (Figures 2.19B and D). Our data suggest that both the *MEN ε/β* transcripts are required for the initial formation of paraspeckles.

Discussion

In the present study, we have identified 184 ncRNAs that are up- or down-regulated more than 2-fold upon C2C12 myoblast differentiation into myotubes. We have focused on the *Men ε/β* locus that yields two long ncRNAs localizing to nuclear paraspeckles. *Men ε* and *Men β* ncRNAs are transcribed from the same RNA polymerase II promoter but differ in the location of their 3' ends. While the *Men ε* transcript is polyadenylated, the 3' end of the *Men β* transcript is generated by RNase P cleavage after a short genomically encoded poly(A)-rich tract. Depletion of the *MEN ε/β* transcripts from cells resulted in the disruption of paraspeckles. Our data demonstrate that the *MEN ε/β* transcripts are essential for establishing paraspeckles *de novo* as well as for maintaining the structural integrity of paraspeckles in cell nuclei.

MEN ε/β ncRNAs as a structural platform of paraspeckles

Mammalian nuclei are highly compartmentalized organelles harboring numerous nuclear domains including, but not limited to nucleoli, speckles, PML bodies, Cajal bodies, and paraspeckles (for review, see Spector 2001; Spector 2006). The molecular components responsible for organizing and maintaining the structural integrity of nuclear domains are largely unknown except that the PML (promyelocytic leukemia) protein is required for the maintenance of PML bodies (Dyck et al. 1994). The idea that RNA may serve a structural and/or organizational role in the cell nucleus surfaced early on in biochemical and cell biological studies (Berezney and Coffey 1974; Brawerman and Diez 1975; Fey et al. 1986; Herman et al. 1976; Perry et al. 1974; Smetana et al. 1963). Interestingly, paraspeckles were previously shown to be sensitive to RNase A treatment

(Fox et al. 2005; Prasanth et al. 2005). Here we show that the *MEN ε/β* ncRNAs are essential for the structural integrity of paraspeckles in cell nuclei (Figures 2.17 – 2.19). Several paraspeckle proteins including PSPC1, SFPQ, and NONO contain two RNA recognition motifs. We demonstrate that the *Men ε/β* transcripts are in the same complex as Nono, suggesting that this RNP complex is likely involved in establishing paraspeckles. Upon transcriptional inhibition, PSPC1 and SFPQ relocalize to the periphery of nucleoli (Fox et al. 2005; Fox et al. 2002). Surprisingly, the *MEN ε/β* transcripts do not follow PSPC1 and SFPQ to the nucleolar cap, but instead localize to speckles as well as being diffusely distributed within the nucleoplasm, respectively (Figure 2.19). Further studies are required to elucidate the signals received by paraspeckles allowing them to respond to the transcriptional activity of the cell and alter their nuclear organization.

Although paraspeckles were identified several years ago, their function still remains elusive (Fox et al. 2002). Prasanth et al. (2005) showed that paraspeckles serve as a storage depot for a specific nuclear retained RNA, *CTN-RNA* (Prasanth et al. 2005). However, knock-down of *CTN-RNA* did not result in an alteration of paraspeckles (Prasanth et al. 2005). Our report demonstrates that the structural integrity of paraspeckles depends on the existence of RNA, namely the *MEN ε/β* transcripts. However, we cannot rule out the possibility that additional RNAs/RNPs may also play a role in the organization of this nuclear compartment. Further characterization of the *MEN ε/β* ncRNAs will provide additional insights into the structure and function of paraspeckles.

Dynamic regulation of ncRNAs during muscle differentiation

Studies of mammalian transcriptomes have suggested that ncRNAs constitute a significant portion of the output of their genomes (Carninci et al. 2005; Kapranov et al. 2007a). Recently, the regulation of long ncRNAs has been studied in various developmental processes including bovine muscle development (Lehnert et al. 2007) and ES cell differentiation (Dinger et al. 2008). We have found large scale changes in the regulation of ncRNAs upon C2C12 muscle cell differentiation (Table 2.1) supporting the premise that ncRNAs are likely important players, in terms of genomic output, rather than transcriptional noise or non-functional RNAs. Since we only examined a small subset of the known ncRNAs, there are presumably many more that also exhibit dynamic expression change in muscle differentiation and other developmental contexts. The biological functions of these ncRNAs in muscle differentiation and cell biology remain to be investigated.

In this study, we have focused on the *Men ε/β* locus, which yields two ncRNAs that localize to paraspeckles (Figure 2.15). *Men ε/β* transcripts are 3.3-fold up-regulated during myoblast differentiation (Figure 2.4, Table 2.1). Recently, the bovine ortholog of the *Men ε* transcript was shown to be 6.8-fold up-regulated during the late stages of muscle development (Lehnert et al. 2007). However, Lehnert et al. (2007) argued that this observed up-regulation was due to the existence of multiple nuclei per myotube causing nuclear retained RNAs to be a larger portion of the total RNA present in the myotubes rather than due to increased transcriptional output of the *Men ε/β* locus (Lehnert et al. 2007). Our results indicate that the nuclear foci of *Men ε/β* ncRNAs were enlarged in myotube nuclei compared to myoblast nuclei (Figure 2.4), suggesting that *Men ε/β*

transcripts are in fact present at a higher level in myotubes. In addition to muscle differentiation-related regulation, *Men ε* transcripts have been reported to be up-regulated in mouse brain infected with Japanese encephalitis virus or Rabies virus (Saha et al. 2006), although the biological significance of this induced expression remains to be explored.

The *MEN ε* locus has previously been reported to produce a 0.5-kb ncRNA in human trophoblast that mapped to the 3' end of the *MEN ε* transcript. This RNA was called trophoblast STAT utron (*TSU*) or trophoblast ncRNA (*TncRNA*) (Geirsson et al. 2003a; Geirsson et al. 2003b; Peyman 1999). In one case, it was suggested that *TSU* binds to STAT1 in the cytoplasm to repress nuclear import of the STAT1 protein (Peyman 1999). In another case, *TncRNA* was reported to suppress class II and III transactivator promoters resulting in the absence of major histocompatibility antigens (MHC) in trophoblast (Geirsson et al. 2003a; Geirsson et al. 2003b). However, we were unable to detect this 0.5-kb RNA using human or mouse total placenta RNA (data not shown). Furthermore, the *MEN ε* transcripts are exclusively retained in the nucleus (Hutchinson et al. 2007).

Processing of the *Men ε/β* ncRNAs

The 3' end of the *Men β* transcript is generated by RNase P cleavage, rather than by the classical cleavage/polyadenylation machinery. The 3' end of another ncRNA, *Malat1*, was recently shown to be generated by the tRNA processing machinery (Wilusz et al. 2008). Curiously, the *MALAT1/Malat1* locus is located immediately adjacent to the *MEN β/Men β* locus, 55-kb or 23-kb downstream of the 3' end of *MEN β/Men β* in the

human and mouse genomes, respectively. tRNA-like structures are located immediately downstream of a genomically encoded poly(A)-rich tract at the 3' ends of both the *Men β* and *Malat1* nascent transcripts. The tRNA-like structures are recognized by RNase P which cleaves to generate the 3' end of the mature *Men β* and *Malat1* transcripts. The *Malat1* tRNA-like structure is further processed by RNase Z and the CCA-adding enzyme, resulting in a 61-nt tRNA-like *mascRNA* which is subsequently exported to the cytoplasm (Wilusz et al. 2008). Somewhat surprisingly, a small tRNA-like transcript originating from the *MEN β/Men β* locus failed to accumulate to significant steady state level in many examined tissues/cell lines, probably because it is rapidly degraded. The *mascRNA* half-life is only ~1 hour (Wilusz et al. 2008), suggesting that the tRNA-like transcript may be inherently unstable. Consistent with this notion, recent studies have revealed a large class of RNAs that are rapidly degraded by the exosome, but can be stabilized by mutations in exosome components (Davis and Ares 2006; Wyers et al. 2005). Besides *Men β* transcripts and *Malat1*, it is not known how many other RNA polymerase II transcripts are subject to 3' end cleavage by RNase P.

In addition to being processed by a similar 3' end processing mechanism, both *Men β* and *Malat*1 are stable nuclear retained RNAs. *Men β* localizes to paraspeckles while *Malat1* localizes to speckles (Figure 2.15, Hutchinson et al. 2007). A-to-I hyper-editing has been suggested as one mechanism for RNA nuclear retention (Chen et al. 2008; Kumar and Carmichael 1997; Prasanth et al. 2005; Zhang and Carmichael 2001). Interestingly, one paraspeckle protein, NONO, was previously shown to be responsible for nuclear retention of hyper-edited RNAs (Zhang and Carmichael 2001). Of the known large nuclear retained RNAs, *CTN-RNA* was shown to bind to Nono and Pspc1 and its 3'

UTR was subject to A-to-I hyperediting (Prasanth et al. 2005). More recently, NONO was shown to be co-localized with nuclear retained RNAs harboring hyperedited inverted Alu repeats (Chen et al. 2008). Although the *Men ε/β* transcripts are a part of the Nono complex, we were unable to detect any editing of the *Men ε/β* transcripts in C2C12 cells. In addition, *Men ε* and *Malat1* do not appear to contain any inverted repeats, suggesting that there are likely additional mechanisms for nuclear retention of RNA. Further characterization of the *Men ε/β* transcripts will determine whether any sequence motifs or secondary structures serve as nuclear retention or paraspeckle localization signals.

In summary, we have identified two ncRNAs that localize to paraspeckles and are responsible for establishing and maintaining these nuclear structures. Our data provide the first demonstration that ncRNAs play a critical role in nuclear organization. Future studies will focus on additional roles of ncRNAs in nuclear structure/function.

Methods

Cell culture and drug treatments

C2C12, bEND.3, MEF, MIR-90, U2OS, and HeLa cells were grown at 37°C, 5% CO_2 in DMEM supplemented with 10% fetal bovine serum (FBS) and penicillin-streptomycin (P/S). NIH3T3 and HEK293 cells were grown at 37°C, 5% CO_2 in DMEM supplemented with 10% calf serum and P/S. Differentiation of C2C12 myoblast cells into myotubes was induced by incubation in DMEM supplemented with 10% horse serum for 5 days. RNA polymerase II was inhibited by DRB (33 μg/ml; Sigma, St. Louis, MO) for 1 hour at 37°C.

Preparation of RNA for microarray analysis

C2C12 myoblasts (American Type Culture Collection, Rockville, MD) were cultured as previously described (Shen et al. 2003). In summary, actively growing myoblasts were maintained in Dulbecco's modified Eagle's medium supplemented with 4 mM L-glutamine, 1.5 g/L sodium bicarbonate, 4.5 g/L glucose and 10% fetal bovine serum. Differentiation was induced when monolayers were 50% confluent by addition of growth medium containing 10% horse serum instead of fetal bovine serum. RNA was isolated 0 hour, 24 hours and 5 days following induction, using the RNeasy Mini Kit (Qiagen, Valencia, CA) and treated with DNase I (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions. The quality of purified total RNA samples was assessed with an RNA 6000 Nano Assay Kit using the Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA) as described by the manufacturer. RNA was amplified and labeled using the Amino Allyl Message Amp II kit (Ambion, Austin, TX) following

the instructions provided by the manufacturer. Amplified aRNA from each time point was labeled with either Cy3 or Cy5 monoreactive dyes (GE Healthcare, Piscataway, NJ) according to the MessageAmp II protocol (Ambion, Austin, TX). The quality and quantity of amplified RNA samples were assessed using the Agilent 2100 Bioanalyzer as described above.

Microarray expression analysis

The design of the microarray has been described previously (Dinger et al. 2008) and is available from the ArrayExpress Data Warehouse (EMBL-EBI; ArrayExpress Accession: A-MEXP-1070). Briefly, the custom microarrays contain 22,038 65-mer oligonucleotide probes from the Mouse OligoLibrary (Compugen, San Jose, CA, USA) and 2,118 70-mer oligonucleotide probes that were designed to target ncRNAs.

Total RNA from C2C12 cells that had been induced to form myotubes for 0 hour, 24 hours, or 5 days was isolated and treated with DNase I (Invitrogen, Carlsbad, CA). RNA was then amplified and labeled using the Amino Allyl Message Amp II kit (Ambion, Austin, TX). Cy3- or Cy5-labeled RNA was hybridized in all pair-wise combinations to individual microarrays. Three technical replicates (including dye swaps) and two biological replicates were performed for each pair-wise comparison. Blocking, hybridization and washing was performed according to the manufacturer's instructions (Full Moon BioSystems, Sunnyvale, CA, USA). Slides were scanned at 5 μm resolution using a DNA microarray scanner (Agilent Technologies, Santa Clara, CA). Feature extraction was performed using ImaGene software (BioDiscovery, El Segundo, CA, USA) with manual grid adjustment and auto-spot finding and segmentation. Data was

exported from ImaGene as text files, then uploaded and analyzed using the Linear Models for Microarray Data (LIMMA) software package via the R Project for Statistical Computing (www.r-project.org). Data was background-corrected, normalized both within and between arrays (Smyth and Speed 2003), and differential expression analysis was performed by fitting a linear model of the data to the experimental design matrix and then calculating Bayesian statistics (B statistics; posterior log odds) adjusted for multiple testing using Benjamini-Hochberg analysis (Smyth 2004). Raw and processed microarray data is available at the ArrayExpress Data Warehouse (EMBL-EBI; ArrayExpress Accession: E-TABM-575).

Microarray probe classification

Microarray probes were classified as targeting coding or non-coding transcripts as previously described (Dinger et al. 2008; Mercer et al. 2008). Briefly, the probe sequences were mapped to the February 2006 (NCBI Build 36) assembly of the mouse genome using BLAT (Kent 2002). Probes that could not be reliably mapped were excluded from the study. The probe mappings were then intersected with the genomic coordinates of all full-length cDNA transcripts as indicated by the UCSC Genome Browser (as of March 2007). Targeted transcripts were then defined as protein-coding and non-coding as described previously (Mercer et al. 2008).

RNA FISH / Microscopy

RNA FISH using nick-translated cDNA probes was performed as previously described (Prasanth et al. 2005). Cells were examined using an Axioplan 2i fluorescence

microscope (Carl Zeiss, Thornwood, NY) equipped with a 40X/1.3 N.A. and a 63X/1.4 N.A. objective lens and Chroma filters (Chroma Technology, Brattleboro, VT). OpenLab software (Improvision, Boston, MA) was used to collect digital images from an ORCA cooled charge-coupled device camera (Hamamatsu, Bridgewater, NJ). Mouse ovary section was examined using Axiobsever fluorescence miscroscope (Carl Zeiss, Thornwood, NY) equipped with a 63X/1.4 N.A. objective lens. Axiovision software (Carl Zeiss, Thornwood, NY) was utilized to collect digital images from camera and process them.

RNA isolation and Northern blotting

Trizol was used for all total RNA isolations as per the manufacturer's instructions (Invitrogen, Carlsbad, CA). For isolation of poly(A)$^+$ RNA, the Oligotex mRNA Mini Kit was used (Qiagen, Valencia, CA). Nuclear and cytoplasmic fractionation and RNA isolation were performed as described previously (Topisirovic et al. 2003). 10 μg of total RNA from mouse tissues or tissue culture cells was separated by electrophoresis through a 1% denaturing agarose gel and transferred to Hybond-N membrane (GE Healthcare, Piscataway, NJ) by capillary transfer. Prehybridization and hybridization was carried out in NorthernMax Prehyb/Hyb Buffer (Ambion, Austin, TX) at 42°C. Labeling of random-labeled probes was performed using the Prime-It RmT Random Primer Labeling Kit (Stratagene, La Jolla, CA). Oligo probes were labeled with [γ-^{32}P] ATP using T4 polynucleotide kinase (New England Biolabs, Ipswich, MA) and used with ULTRAhyb-Oligo Hybridization Buffer as per the manufacturer's instructions (Ambion, Austin, TX). Blots were visualized and quantified using the Fujifilm Life Science FLA-5100 imaging

system.

To map the 3' end of *MEN β*, 15 μg of HeLa total RNA was first mixed with 20 pmol of antisense oligo and heated to 65°C for 10 min. After allowing the antisense oligo to anneal by slow cooling, the RNA was treated with RNase H (New England Biolabs, Ipswich, MA) at 37°C for 30 min and then subject to small RNA Northern blot analysis, as previously described (Wilusz et al. 2008).

RNase Protection Assay (RPA)

Mouse *β-Actin* and *Men β* probes were internally labeled using [α-^{32}P] UTP and gel purified. RPA III kit (Ambion, Austin, TX) was used as per manufacture's instructions. *β-Actin* and *Men β* probes at 2,000 cpm and 16,000 cpm, respectively, were allowed to hybridize to 10 μg of total RNA from mouse tissues and C2C12 cells. After unprotected nucleotides were digested by RNase T1 at 37°C for 30 min, samples were electrophoresed on 4% polyacrylamide/8 M urea gels.

***In vitro* cleavage assays**

Mouse *Men β* RNA substrates were internally labeled using [α-^{32}P] UTP, gel purified, and used at 10,000 cpm per cleavage reaction. RNase P reactions were incubated at 37°C for 1 hour in 50 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, and 100 mM NH_4Cl. RNase Z reactions using Recombinanat His-tagged tRNase ZL (Δ30) were incubated at 37°C for 30 min in 10 mM Tris-HCl (pH 7.5), 1.5 mM DTT, and 10 mM $MgCl_2$. All reactions were stopped by adding gel loading dye and samples were electrophoresed on 8% polyacrylamide/8 M urea gels. The *in vitro* RNase P cleavage site

was cloned using the GeneRacer kit (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions, except that CIP and TAP treatments were omitted.

Establishment of EYFP-PSPC1 stable cell lines

EYFP-PSPC1 (Fox et al. 2002) was cloned into the pMSCV-Puro vector (Clontech, Mountain View, CA), which was then transfected into Phoenix packaging cells. C2C12 or HeLa cells were infected by virus for 24 hours and were selected under 1.3 μg/mL or 0.35 μg/mL puromycin, respectively, for 1 week. FACS analysis was performed to isolate EYFP positive cells, followed by clonal selection.

Generation of 9-99 monoclonal antibody

GST-PSPC1 was expressed in BL21(DE3) cells using the pGEX-6P system (GE Healthcare, Piscataway, NJ). PreScission Protease (GE Healthcare, Piscataway, NJ) was used to cleave the GST tag from PSPC1. PSPC1 protein was then injected into mice, hybridomas were established, and supernatants from 96 hybridomas were screened for specificity of antigen recognition by immnuofluorescence microscopy and immunoblotting. One of the supernatants, 9-99 specifically recognized mouse Nono instead of Pspc1 due to amino acid homology between the two proteins, while it recognized both human PSPC1 and NONO.

Q-PCR

For real-time quantitative PCR (Q-PCR), 1 μg of total RNA was treated with DNase I (Invitrogen, Carlsbad, CA) and reverse transcribed to cDNA using TaqMan

Reverse Transcription Reagents (Applied Biosystems, Foster City, CA). Gene-specific primer sets were designed using Primer 3 software. Q-PCR was carried out in triplicate using SYBR Green PCR Master Mix (Applied Biosystems, Foster City, CA) and *β-Actin* served as an endogenous normalization control.

Knock-down of *MEN ε/β* transcripts

Second generation 2'-O-methoxyethyl oligonucleotides were synthesized at ISIS Pharmaceuticals (Carlsbad, CA) to the target *MEN ε/β* transcripts and tested for their efficacy. Oligonucleotides were administered at a final concentration of 200 nM to HeLa cells using Lipofectamine RNAiMAX (Invitrogen, Carlsbad, CA). Cells were incubated with a mixture of Lipofectamine RNAiMAX and oligonucleotide in Opti-MEM I medium (Invitrogen, Carlsbad, CA) at 37°C, 5% CO_2. After 6 hours, the transfection mixture was aspirated from the cells and replaced with fresh DMEM supplemented with 10% FBS and incubated at 37°C, 5% CO_2 for additional 16 – 18 hours prior to assays.

Co-immunoprecipitation

C2C12 cell lysate was prepared in RIPA buffer containing Anti-RNase (Ambion, Austin, TX) and Complete mini protease inhibitor (Roche Diagnostics, Indianapolis, IN). Following lysate centrifugation, the supernatant was incubated for 1 hour at 4°C with antibody conjugated Dynabeads Protein A (Invitrogen, Carlsbad, CA). Beads were then washed four times in RIPA buffer. 80% of the IP material was used for extraction of RNA using Trizol (Invitrogen, Carlsbad, CA) and subsequent RT-PCR.

Antibodies

For immunofluorescence, antibodies were used at the following concentrations; SC35 at 1:50, Coilin at 1:100, PML at 1:15, SF2/ASF at 1:5, ANA-N (Sigma, St. Louis, MO) at 1:10, and purified 9-99 ascites at 1:20, respectively. For immunoblotting, antibodies were used at the following concentrations; FC23 (anti-Sfpq) at 1:50, anti-PSPC1 at 1:2,000 – 4,000, anti-Lamin B1 at 1:2,000, anti-GFP at 1:1,000, 9-99 supernatant at 1:1, and purified 9-99 ascites at 1:100, respectively.

siRNA transfection

siRNAs to PSPC1 or NONO were synthesized at Ambion (Austin, TX) and Sigma (St. Louis, MO), respectively. siRNAs were administered at a final concentration of 10 nM to HeLa cells using Lipofectamine RNAiMAX (Invitrogen, Carlsbad, CA). Cells were incubated with a mixture of Lipofectamine RNAiMAX and siRNA in Opti-MEM I medium (Invitrogen, Carlsbad, CA) at 37°C, 5% CO_2. After 24 hours, the transfection mixture was replaced with fresh DMEM supplemented with 10% FBS and incubated at 37°C, 5% CO_2 for additional 24 hours prior to assays.

Alkaline phosphatase treatment

HeLa cell lysate was prepared in RIPA buffer containing Complete mini protease inhibitor (Roche Diagnostics, Indianapolis, IN). Following lysate centrifugation, 10 μg of protein was incubated with 10 units of calf intestinal alkaline phosphatase (New England Biolab, Ipswich, MA) for 1 hour at 37°C in 50 mM Tris-HCl (pH 7.9), 100 mM NaCl, 10 mM $MgCl_2$, and 1mM Dithiothreitol. 3.5 μg of protein was used of immunoblotting.

Oligonucleotide sequences

AK030860 primers to generate RNA FISH probe (Figure 2.1)

Forward primer: CCTTCCTTGCACACACACAC

Reverse primer: CTTTCCTGAAAGTCACAATCCTG

Mouse *Men ε/β* primers to generate RNA FISH and Northern probe (Figures 2.2A, 2.3, 2.4, 2.5, 2.6B, 2.8B, 2.14A, and 2.15C)

Forward primer: CCGAGGAGTTAGTGACAAGG

Reverse primer: CTGTAAAGGGGAGGAAAATGGT

Human *MEN ε/β* primers to generate RNA FISH and Northern probe (Figures 2.2B, 2.8A, 2.9, 2.14B, 2.17, 2.18, 2.19, 2.20, and 2.21C)

Forward primer: TAGTTGTGGGGGAGGAAGTG

Reverse primer: TGTGCTGTAAAGGGGAAGAAA

Q-PCR primers (Figures 2.4C and 2.6A)

Upstream forward primer: TTCATTCTGCGTGCTTGAAC

Upstream reverse primer: CGCCATGAAGCATCTACTCA

Men ε/β forward primer: GGGAAGGGTGACATTGAAAA

Men ε/β reverse primer: CTCCCCAGCTTCACTTCTTG

Men β-1 forward primer: CTGGTTTATCCCAGCGTCAT

Men β-1 reverse primer: CTTACCAGACCGCTGACACA

Men β-2 forward primer: TGCTTTTGTTGCCTGAACTG

Men β-2 reverse primer: AGCACTTGCCGAGTGTTTTT

Men β-3 forward primer: TCCTGCCAGTGATGAAGATG

Men β-3 reverse primer: AACTGCTACCCACCGAGATG

Men β-4 forward primer: TGGGGATTATTGAGCTGAGG

Men β-4 reverse primer: CTCCTGAGTACCGGGATGAA

Men β-5 forward primer: TGCTTACACGGCTTGTTCAG

Men β-5 reverse primer: AACTCCAAGGTCCCTGTCCT

Gapdh forward primer: AACTTTGGCATTGTGGAAGG

Gapdh reverse primer: GGATGCAGGGATGATGTTCT

PCR primers to generate mouse *Men β* probe used in RPA (Figure 2.6C)

T7-*Men β* Forward primer (T7 promoter sequences in italics)

TAATACGACTCACTATAGG TGGAATACAACCTGGGAATGTTC

Men β Reverse primer (Extra nucleotide sequences in italics)

ATGAGTGAGCTAACTCACATTAATTGCGTT ATAGTTAGGCCTAGATGCACTG

Oligonucleotides probes (Figure 2.7)

mChr19 5845590-639

AGCCACAGAGCCAAAGGAGCCCACGAGGGGCGGGTCATGGCCTCCGTCAA

mChr19 5845540-589

AACGGAACGATTCCTCCACGGGCACTTAAAAAAAAGTTCCTCCCCGACCT

mChr19 5845490-539

TTACAACTTTTGCTTTTATACTCTTGTGTCGCGTCACCCAAACACTGCTA

mChr19 5845440-489

ACTAGCCCCAAGAGCGAGCCCTCCTTGTCACTAACTCCTACAGGTGGAGG

mChr19 5842304-353

AGCTTCAATCTCAAACCTTTATTTTGCTGTAAAGGGGAGGAAAATGGTTA

mChr19 5842252-301

TGAAGGCAAAGTGACAGAGGTCGAGAATGTAACAGCCATGACAATTCTAA

mChr19 5842202-251

AGACCTGCTCTAACTGCTTCCATGGTGCAGTGCATCTAGGCCTAACTATA

mChr19 5842102-151

GCAGTGCCATTATCCCATGACTCAGTGTCATGTCTGCAGCCATTGTCAGC

Human *MEN β* primers for RNA FISH probe 1 and Northern probe (Figures 2.8A, 2.9A, 2.18, 2.19, and 2.20)

Forward primer: CCACTCAAGCCAATGAAGGT

Reverse primer: AGAGAGGAAAGGCAGTGCAG

Human *MEN β* primers for RNA FISH probe 2 (Figure 2.8A)

Forward primer: CAGGGCTTCTGAGCTCCTTA

Reverse primer: CCACATGTCTCCTAGCATGG

Mouse *Men β* primers for RNA FISH probe 1 (Figure 2.8B)

Forward primer: GGTCCTGCAAACACTGCTCT

Reverse primer: ACACACAGGGAAGCTGGACT

Mouse *Men β* primers for RNA FISH probe 2 (Figure 2.8B)

Forward primer: TGCTTGGAGCTGGTTCTTTT

Reverse primer: AGGGACATGAAGTGCAGACA

Mouse *Men β* primers to generate RNA FISH probe 3 (Figures 2.8B and 2.15C)

Forward primer: CATGTCAGCCTTCTGCTTCA

Reverse primer: CCAACCACAGAACTGCTACCT

Mouse *Men β* primers for RNA FISH probe 4 (Figure 2.8B)

Forward primer: CTCTCCGCTACTGCGAGAAT

Reverse primer: GGCCCTCTAAGGAAACATCC

Mouse *Men β* primers for RNA FISH probe 5 (Figure 2.8B)

Forward primer: GCCCACACCTTTGTGACACT

Reverse primer: GGGGGAGGAGAAAGAACAAC

EYFP primers to generate RNA FISH probe (Figure 2.10C)

Forward primer: ATGATGCCATGGTGAGCAAGGGCGAGGAG

Reverse primer: ATGATGTTAATTAACTTGTACAGCTCGTCCATGCCG

MEN β primers to generate hChr11 64967173-8193 probe (Figure 2.11A)

Forward Primer: ATCCGATCTGCCATATCCTG

Reverse Primer: ATCATCCCTCCTCACACGTC

Oligonucleotide probes (Figure 2.11A)

hChr11 64969534-583 Antisense

GCTTTTTCCTGCTCACTCTTTCACAGATGGGAAGGTGAAAAGCAAAACCT

hChr11 64969584-633 Antisense

ACTCGAACCCCGGCCCAGCCGTGCTGGACGTGCCACCACCAGCGCCTTTT

hChr11 64969634-683 Antisense

GAGGAGAGGCCTGGAGGAGGGAGCTGGAAGGAAGCAGCAACACTGCGGGG

RNase H/Northern antisense oligonucleotides (Figure 2.11B)

Oligo 1 (hChr11 64969484-533 Antisense)

GAGTGCGGCCATGGGCTGCACTCAGTAAAAGAAACACCTGCGGCGGCTCC

Oligo2 (hChr11 64969492-541 Antisense)

CAAAACCTGAGTGCGGCCATGGGCTGCACTCAGTAAAAGAAACACCTGCG

Northern Probe

GCTTTTTCCTGCTCACTCTTTCACAGATGGGAAGGTGAAAAGCAAAACCT

PCR primers to generate *Men β* RNase P Substrate (Figure 2.11D)

T7-*Men β* RNase P Forward primer (T7 promoter sequences in italics)

TAATACGACTCACTATA GGGAGGGGCACGGAGCCGCC

Men β RNase P Reverse primer

AGCAGGAAGGAAGCACGGTAC

PCR primers to generate *Men β* RNase Z substrate (Figure 2.11E)

T7-*Men β* RNase Z Forward (T7 promoter sequences in italics)

TAATACGACTCACTATA GGCACTGGTGGCGGCACGCCC

Men β RNase Z Reverse

CAAACTTTCTCCTAGGAACTGCTAAG

Oligonucleotide probes (Figure 2.11F)

Men β tRNA-like Small RNA: CGTGCCGCCACCAGTGCC

mascRNA: AGTGCCAGCCACCAGCGTCT

U6: GCTAATCTTCTCTGTATCGTTCCAATTTTAGTATATGTGCTGCCG

RT-PCR and Q-PCR primers to confirm that the mouse *Men ε/β* transcripts interact with the Nono complex (Figure 2.16)

Men ε/β Forward primer: GGGGCCACATTAATCACAAC

Men ε/β Reverse primer: TCAGAGTGAGGGGCAAGAGT

Men β forward primer: GAGGGCCTGTGAAAGCATTA

Men β reverse primer: CATCCAGATTTTGGGAAGGA

Gapdh forward primer: AACTTTGGCATTGTGGAAGG

Gapdh reverse primer: GGATGCAGGGATGATGTTCT

ASOs to knock-down *MEN ε/β* transcripts (Figures 2.18 – 2.20)

Control ASO: CCTTCCCTGAAGGTTCCTCC

ASO 1: TCGCTCATGATTTTCAATCA

ASO 2: ATCACACATGTAGTAAAGGC

ASO 3: ATCATCCCCAAGTCATTGGT

ASO 4: AGAAACACCTGCGGCGGCTC

Human *MEN ε/β* Q-PCR primers (Figures 2.18 – 2.20)

Forward primer: GGGCCATCAGCTTTGAATAA

Reverse primer: CTTGAAGCAAGGTTCCAAGC

Human *MEN β* Q-PCR primers (Figures 2.18 – 2.20)

Forward primer: GCTGAGAAGGAAGGTGCTTG

Reverse primer: CTGGCTAGTCCCAGTTCAGC

Human *β-Actin* Q-PCR primers (Figures 2.18 – 2.20)

Forward primer: AGAAAATCTGGCACCACACC

Reverse primer: GGGGTGTTGAAGGTCTCAAA

siRNAs to knock-down PSPC1 and NONO (Figure 2.19)

PSPC1 sense: GCAGGUUGAUAGAAACAUC[dT][dT]

PSPC1 antisense: GAUGUUUCUAUCAACCUGC[dT][dC]

NONO sense: GACUAUUGACCUGAAGAAU[dT][dT]

NONO antisense : AUUCUUCAGGUCAAUAGUC[dT][dT]

Figure 2.1. Localization of ncRNA candidates by RNA FISH in NIH3T3 cells. RNA FISH was performed to study the localization of several candidate ncRNAs that showed more than 2-fold change during myoblast differentiation into myotubes in microarray experiments. Scale bar, 10 μm.

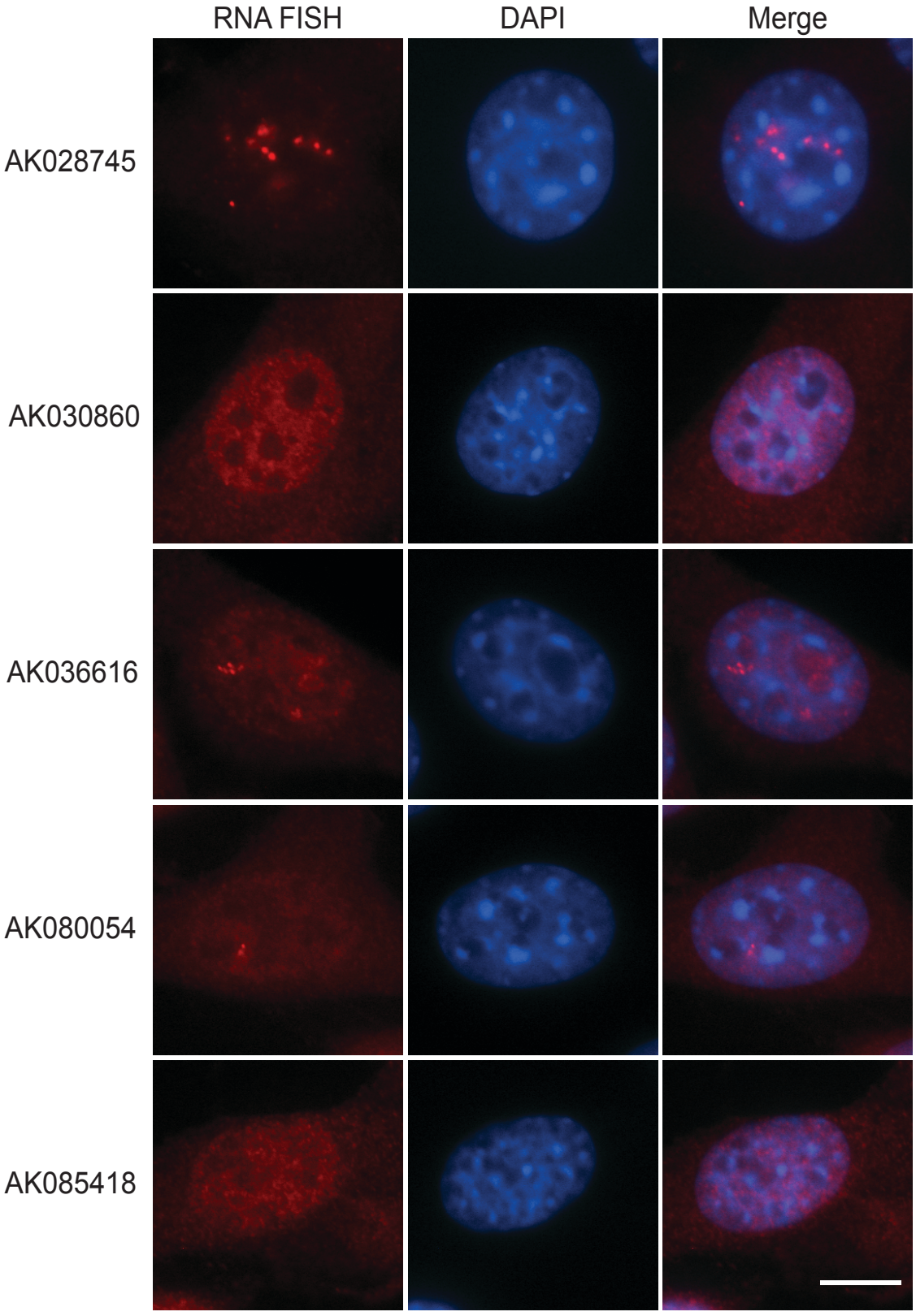
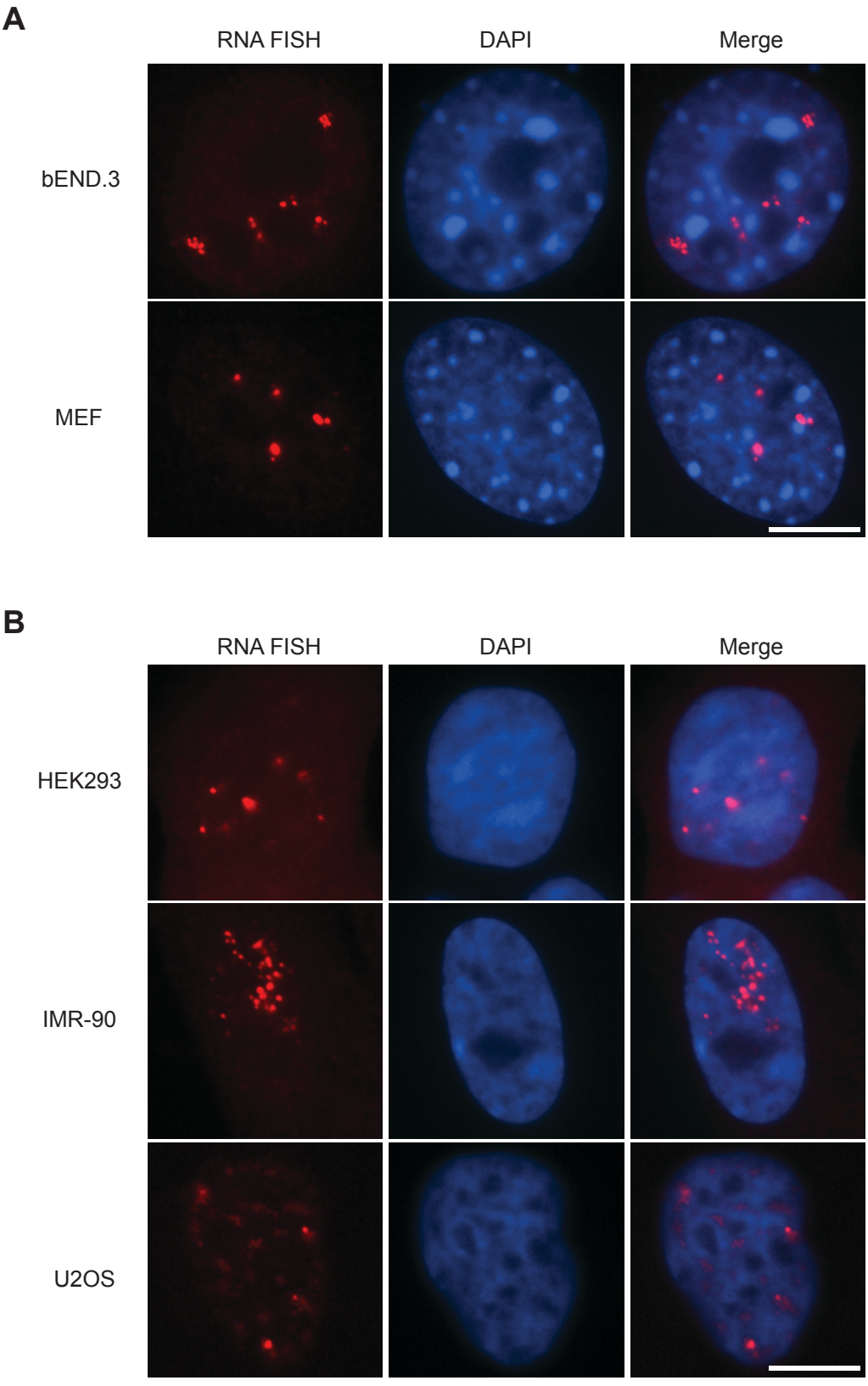

RNA FISH
DAPI
Merge
AK028745
AK030860
AK036616
AK080054
AK085418

Figure 2.2. *Men ε/β* and *MEN ε/β* localize in a punctate nuclear pattern in numerous mouse and human cell lines. (A) RNA FISH was performed on two additional mouse cell lines, bEND.3 and MEF, using a probe to the *Men ε/β* transcripts. **(B)** RNA FISH was performed on three human cell lines, HEK293, IMR-90, and U2OS using a probe to the *MEN ε/β* transcripts. Scale bars, 10 μm.


A
RNA FISH
DAPI
Merge
bEND.3
MEF
B
RNA FISH
DAPI
Merge
HEK293
IMR-90
U2OS

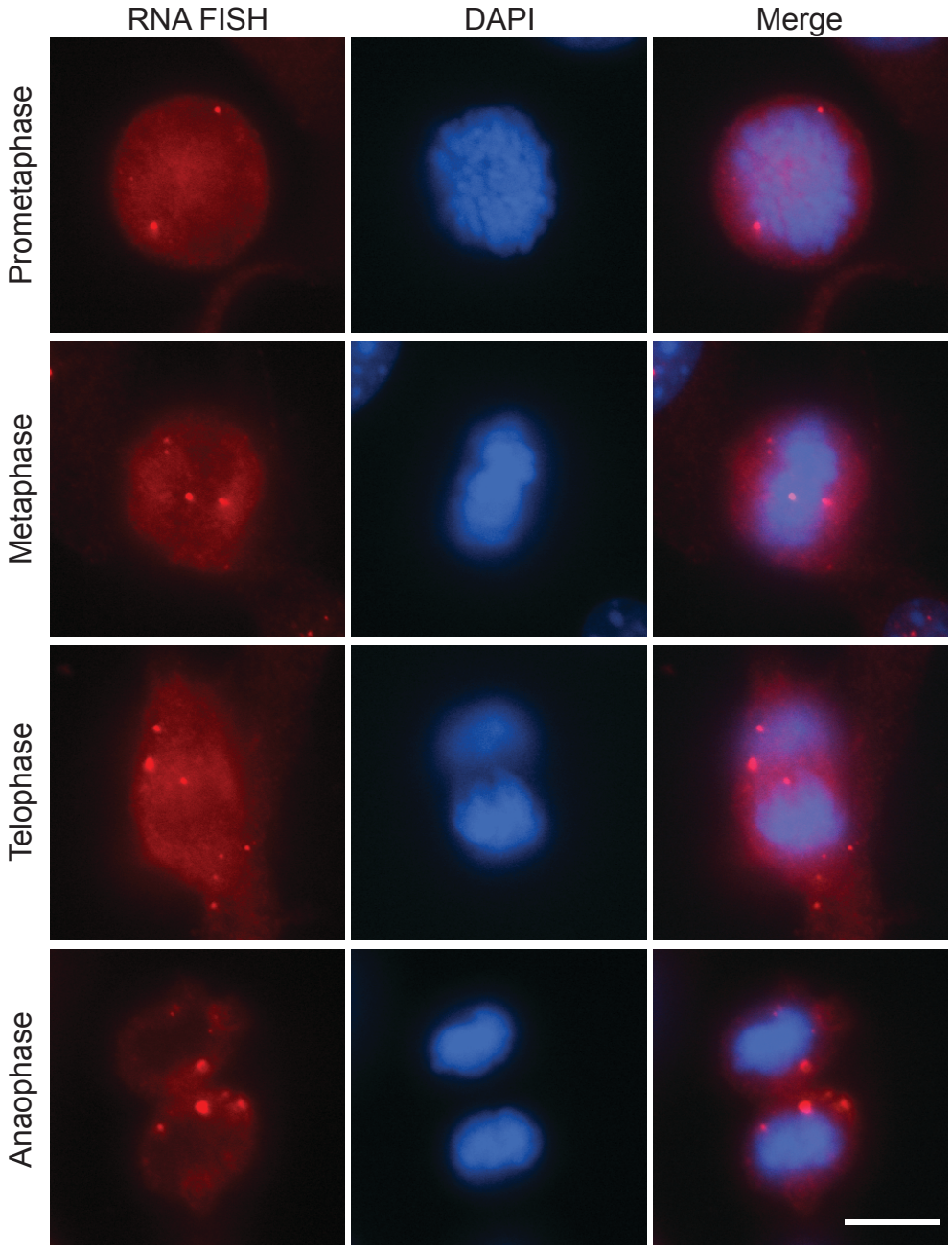
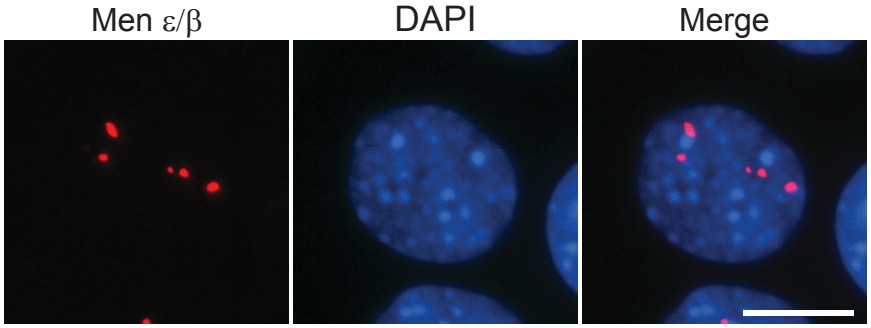


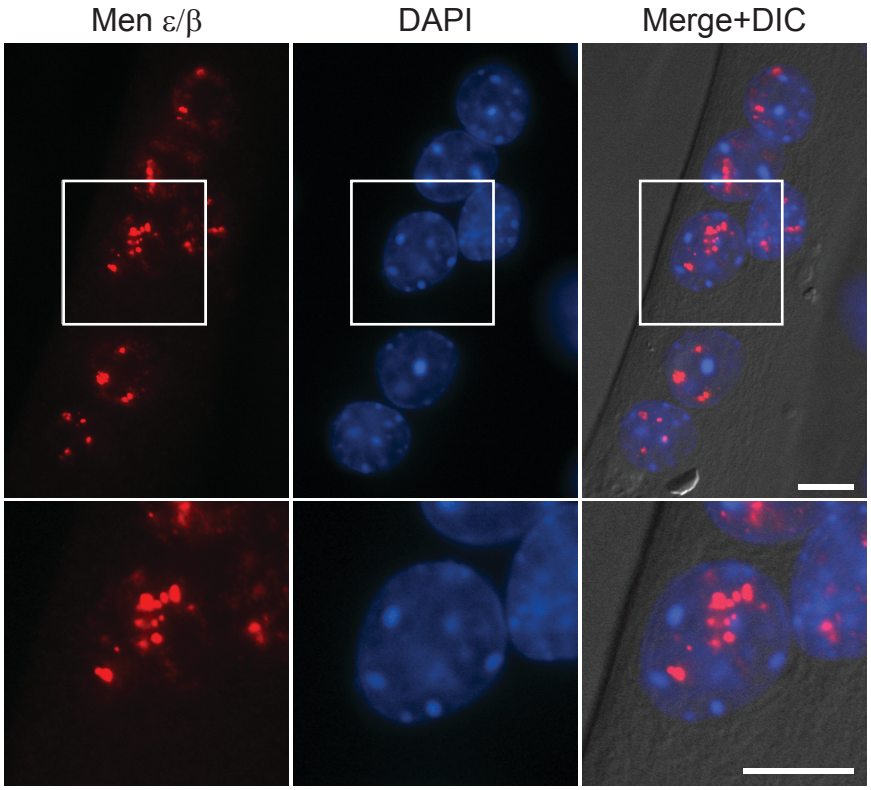
Figure 2.3 *Men* ε/β transcripts are localized to a few foci throughout mitosis in NIH3T3 cells. Although the number of *Men* ε/β foci decreases in mitotic cells compared to interphase nuclei, the foci of *Men* ε/β transcripts persist during mitosis. Scale bar, 10 μm.

Figure 2.4. *Men ε/β* ncRNA localization in C2C12 nuclei. (A) RNA FISH revealed that the *Men ε/β* transcripts are localized in discrete foci in C2C12 myoblast nuclei. **(B)** In C2C12 myotubes, the foci of *Men ε/β* ncRNAs are enlarged and present in greater numbers than in C2C12 myoblasts. Scale bars, 10 μm. **(C)** Both *Men ε* and *Men β* isoforms are up-regulated upon C2C12 myoblast differentiation into myotubes as assessed by Q-PCR. *Gapdh* was used as a normalization control. The data in the histogram are shown as mean and standard deviation values of three independent experiments.

A


Men ε/β
DAPI
Merge

B


Men ε/β
DAPI
Merge+DIC

C

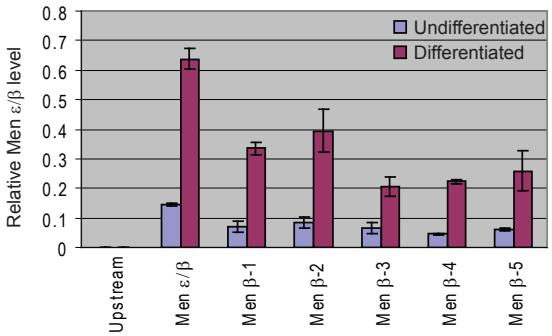

Relative Men ε/β level
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Undifferentiated
Differentiated
Upstream
Men ε/β
Men β-1
Men β-2
Men β-3
Men β-4
Men β-5

Figure 2.5. *Men ε/β* localization in mouse tissues. (A) RNA FISH was performed on a mouse ovary section that was sliced to 15 μm thickness by cryosection. *Men ε/β* ncRNAs are expressed only in Corpus luteum (CL), but not in follicle (Fc) (top panel). Lower panel represents high magnification pictures of boxed areas in top panel. Scale bars, 100 μm (top panel) and 20 μm (lower panel), respectively. **(B)** RNA FISH was performed on a mouse kidney section that was sliced to 15 μm thickness by cryosection. The expression of *Men ε/β* ncRNAs is detected only in subset of cells. Scale bar, 20 μm.

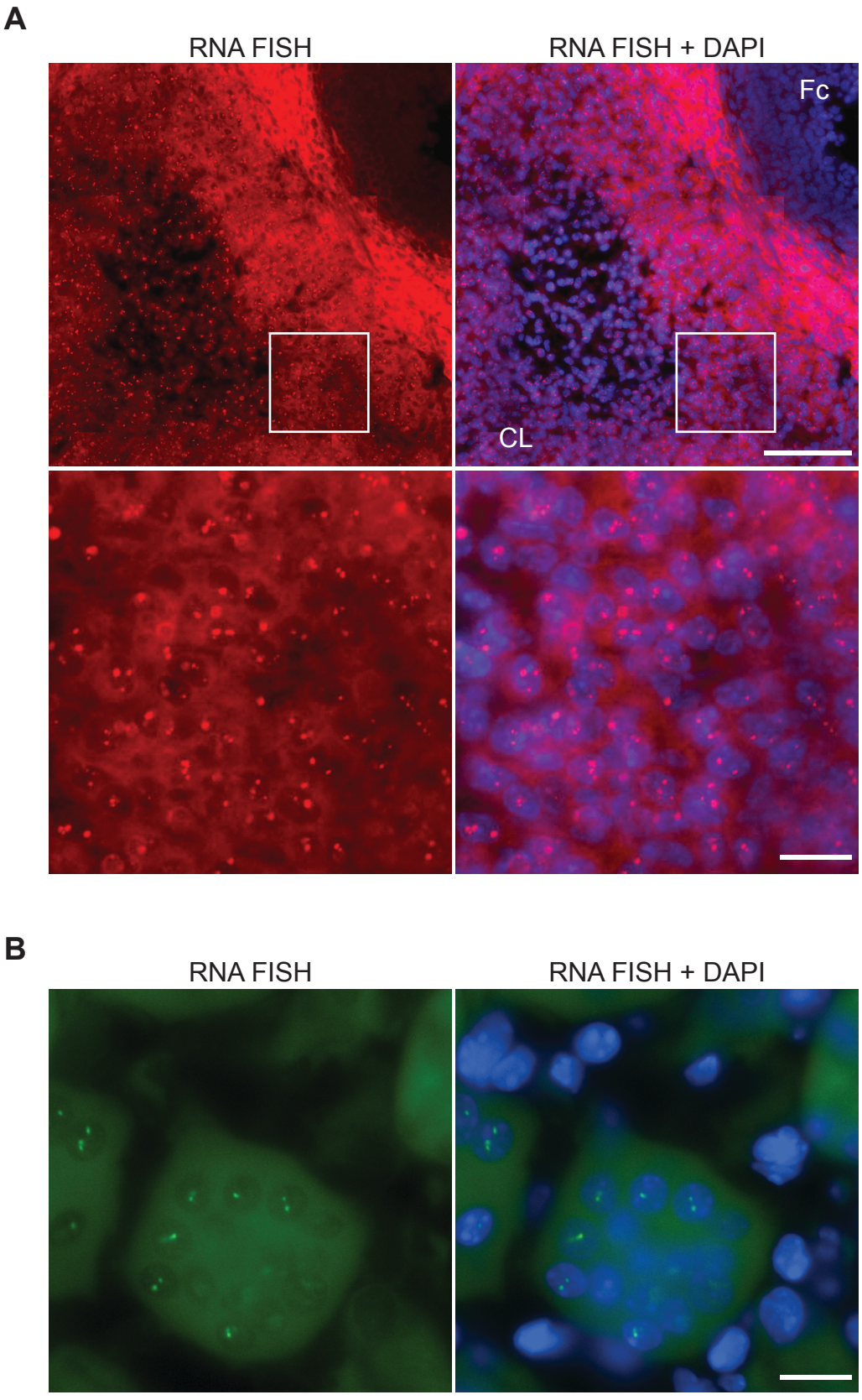

A
RNA FISH
RNA FISH + DAPI
Fc
CL
B
RNA FISH
RNA FISH + DAPI

Figure 2.6. The *Men ε/β* locus produces two non-coding RNAs. (A) The *Men ε/β* ncRNAs are transcribed from a single promoter located on mouse chromosome 19qA. *Men ε* is highly conserved among mammals and does not contain any repetitive elements. Transcription start site and polyadenylation site are denoted by arrows. Northern and RNase protection assay (RPA) probe positions are depicted by black box and ▲, respectively. The position of the primer pairs used in the Q-PCR (Figure 2.4C) is indicated. The 3' UTR of a nearby protein coding gene is located ~6-kb upstream of the transcription start site. **(B)** Northern blot analysis using 10 μg of total RNA from 8 week old C57Bl6 mice. A single band of approximately 3.2-kb (arrow) was detected using a probe complementary to the *Men ε* transcript. Relative levels of *Men ε* in various tissues are depicted as a histogram. *Beta-Actin* (also known as *Actb*) was used as loading control. **(C)** RNase protection assay was performed using 10 μg of total RNA from the same mice. Relative levels of *Men β* in various tissues are depicted as a histogram. *β-Actin* was used as loading control. Full-length *β-Actin* probe and *Men β* probe are 304-nt and 221-nt, respectively. Protected *MEN β* fragment (arrow) is 201-nt. Arrowheads depict protected *β-Actin* fragments, 245-nt major fragment and two smaller muscle specific fragments.

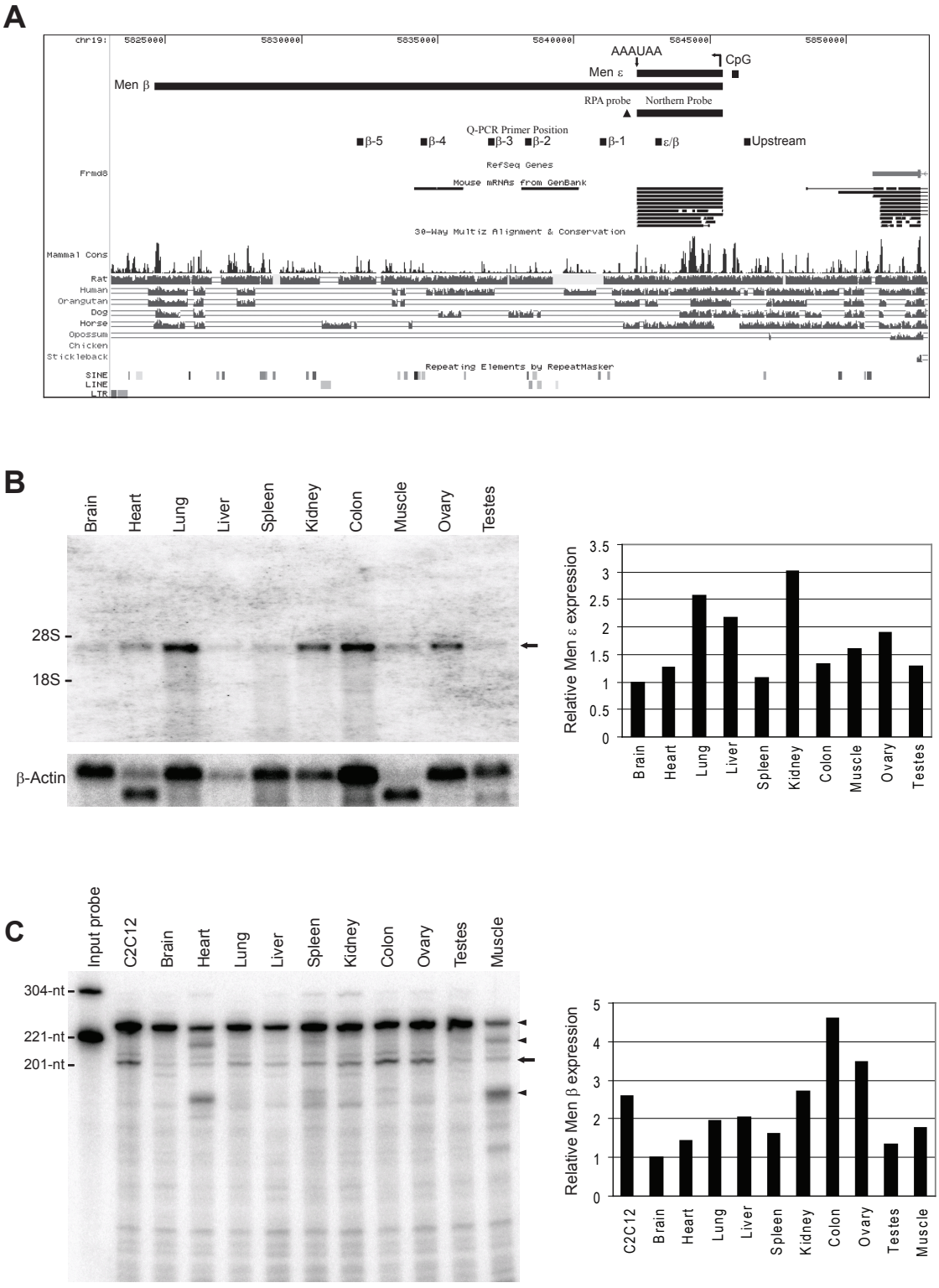

A
chr19: 5825000 5830000 5835000 5840000 5845000 5850000
AAAUAA
CpG
Men ε
Men β
RPA probe
Northern Probe
Q-PCR Primer Position
β-5 β-4 β-3 β-2 β-1 ε/β Upstream
RefSeq Genes
Frmd8
Mouse mRNAs from GenBank
30-Way Multiz Alignment & Conservation
Mammal Cons
Rat
Human
Orangutan
Dog
Horse
Opossum
Chicken
Stickleback
Repeating Elements by RepeatMasker
SINE
LINE
LTR
B
Brain Heart Lung Liver Spleen Kidney Colon Muscle Ovary Testes
28S
18S
β-Actin
Relative Men ε expression
3.5 3 2.5 2 1.5 1 0.5 0
Brain Heart Lung Liver Spleen Kidney Colon Muscle Ovary Testes
C
Input probe C2C12 Brain Heart Lung Liver Spleen Kidney Colon Ovary Testes Muscle
304-nt
221-nt
201-nt
Relative Men β expression
5 4 3 2 1 0
C2C12 Brain Heart Lung Liver Spleen Kidney Colon Ovary Testes Muscle

Figure 2.7. The 5' and 3' end of the *Men ε* transcripts were mapped by Northern blot analysis using oligonucleotides probes. (A) The *Men ε* transcript was detected by oligonucleotide probe mapped around the annotated transcription start site (lane 4, arrow), but not by further upstream probes (lane 1 – 3). Ten μg of C2C12 total RNA was loaded each lane. A nonspecific band is denoted by *. **(B)** Membranes were stripped and reprobed to map the 3' end of the *Men ε* transcripts. The *Men ε* transcript was detected by oligonucleotide probe mapped around the annotated polyadenylation site (lane 1, arrow), but not by further downstream probes (lane 2 – 4). **(C)** Ethidium bromide staining picture of the gel is shown to confirm equal loading and the quality of RNA used.

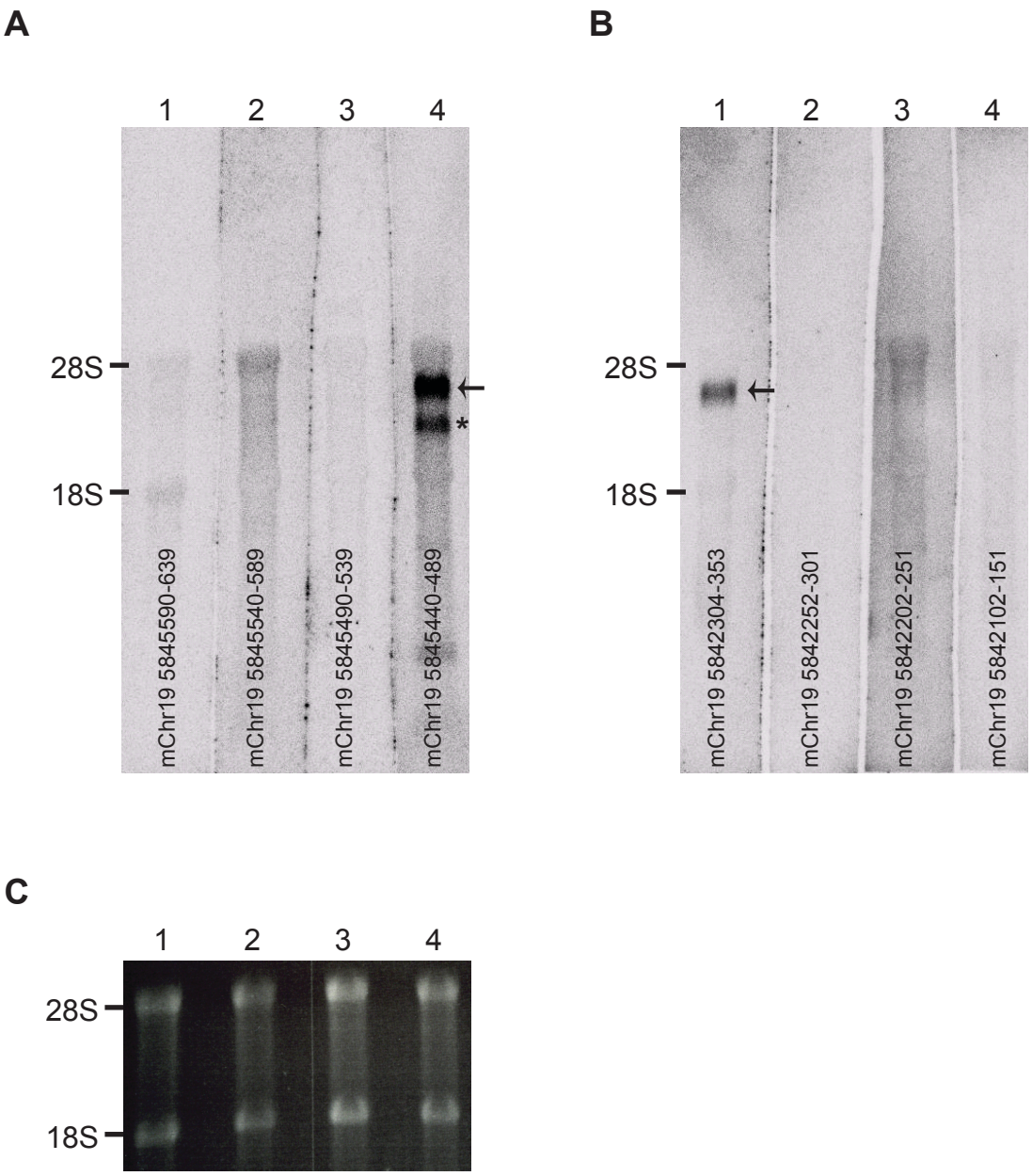

A
1
2
3
4
28S
18S
mChr19 5845590-639
mChr19 5845540-589
mChr19 5845490-539
mChr19 5845440-489
*
B
1
2
3
4
28S
18S
mChr19 5842304-353
mChr19 5842252-301
mChr19 5842202-251
mChr19 5842102-151
C
1
2
3
4
28S
18S

Figure 2.8. *MEN ε* and *Men ε* probes exhibit the same localization pattern as *MEN β* and *Men β* probes. (A) One *MEN ε/β* and two *MEN β*-specific probes were generated along the human *MEN ε/β* locus. *MEN β* transcripts are localized to discrete foci, while *MEN ε* probes exhibit additional speckle localization. **(B)** One *Men ε* probe and five *Men β*-specific probes were generated at various positions along the mouse *Men ε/β* locus where there are no repetitive elements for at least 1-kb. All *Men ε* probes co-localized with the *Men β* probe in C2C12 myoblast nuclei. Scale bars, 10 μm.

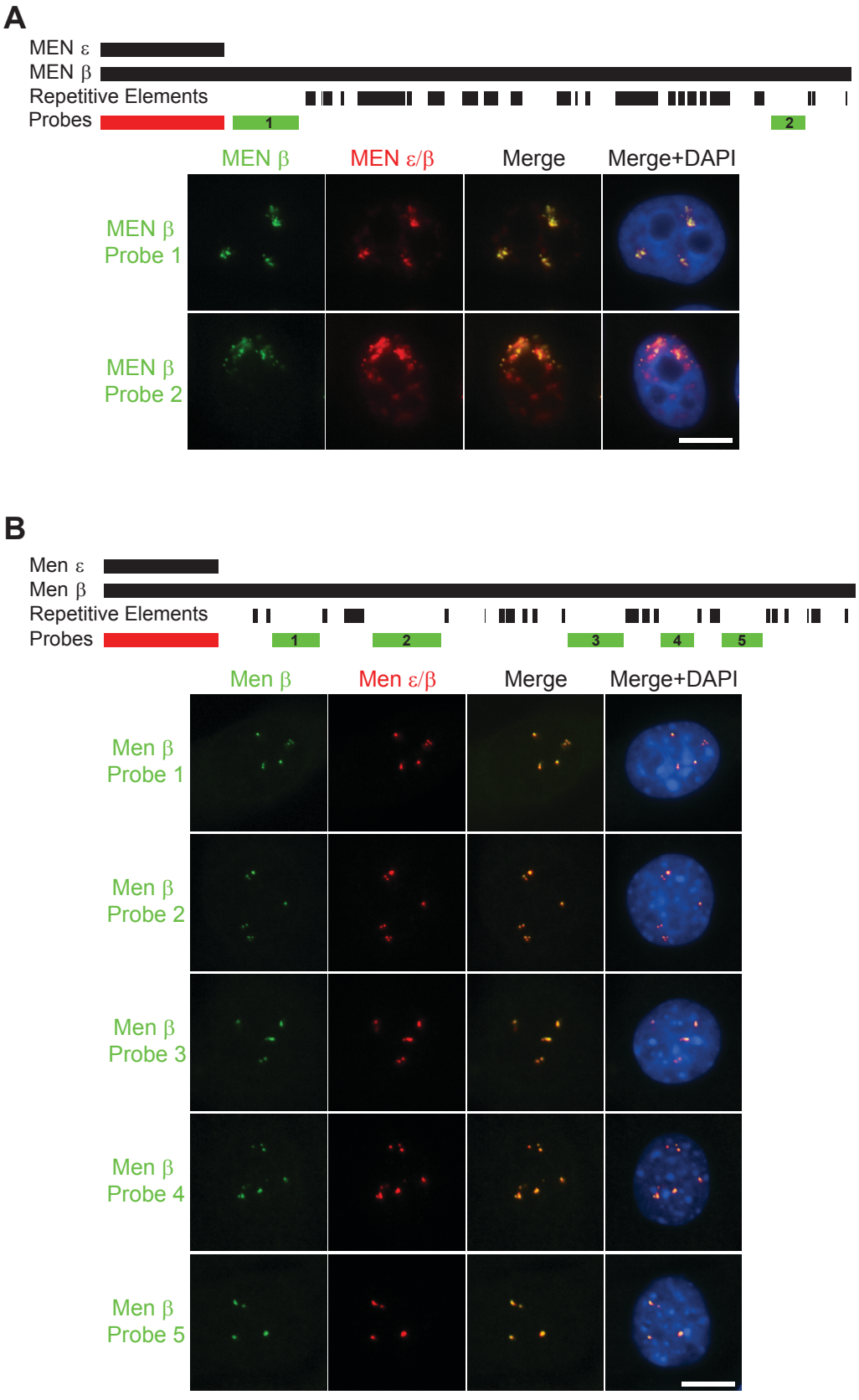

A
MEN ε
MEN β
Repetitive Elements
Probes
1
2
MEN β
MEN ε/β
Merge
Merge+DAPI
MEN β Probe 1
MEN β Probe 2
B
Men ε
Men β
Repetitive Elements
Probes
1
2
3
4
5
Men β
Men ε/β
Merge
Merge+DAPI
Men β Probe 1
Men β Probe 2
Men β Probe 3
Men β Probe 4
Men β Probe 5

Figure 2.9. *MEN ε/β* are nuclear retained RNAs. (A) Northern blot analysis confirmed that *MEN ε/β* are nuclear retained RNAs. The *MEN ε* transcripts are detected in the poly(A)$^+$ RNA fraction and nuclear fraction, but not in the cytoplasmic fraction. The membrane was stripped and reprobed to detect the *MEN β* transcripts. *MEN β* was also detected in nuclear fraction, but not in cytoplasmic fraction. Ten μg of total RNA from HeLa, 1 μg of poly A (+) RNA, 10 μg of RNA from nuclear fraction, and 10 μg of RNA from cytoplasm was used. **(B)** Ethidium bromide staining picture of the gel is shown to confirm equal loading and the quality of RNA used. **(C)** Northern blot analysis was performed using RNA extracted from HEK293 cells. The *MEN ε* transcripts are detected in both nuclear and cytoplasmic fraction. MEN b is still nuclear retained. Ten μg of total RNA from HEK293, 10 μg of RNA from nuclear fraction, and 10 μg of RNA from cytoplasm was used. **(D)** Ethidium bromide staining picture of the gel is shown to confirm equal loading and the quality of RNA used.

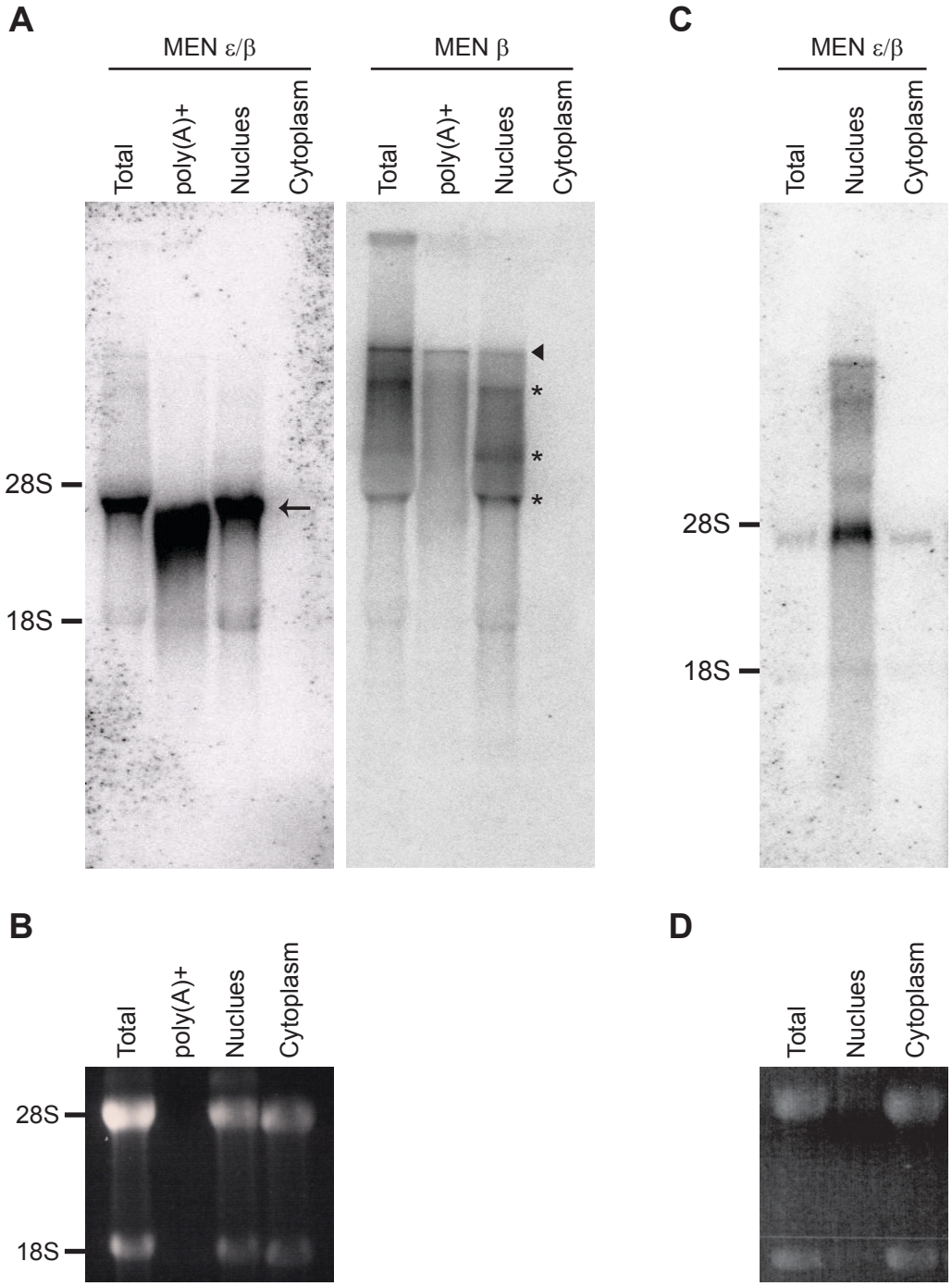
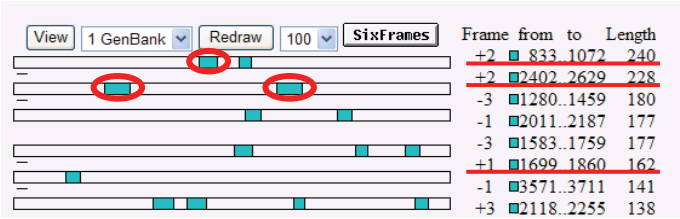
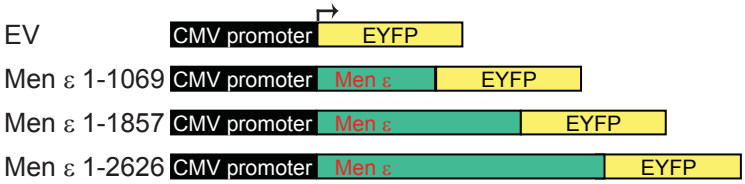

A
MEN ε/β
MEN β
Total
poly(A)+
Nuclues
Cytoplasm
28S
18S
C
B
D

Figure 2.10. Putative open reading frames (ORFs) in *Men ε/β* transcripts are not functional to encode peptides. (A) Putative ORFs in *Men ε* transcripts were bioinformatically looked for by using ORF finder. Three longest putative ORF were selected for test of their coding capability (circled and underlined). **(B)** PCR amplified partial fragment of the *Men ε* ncRNA from 5' end to stop codon of each ORF was cloned into pEYFP N vectors so that any translation will proceed into EYFP sequences. The transcription start site is depicted by an arrow **(C)** HEK293 cells were examined by RNA FISH using a probe to EYFP sequences 24 hours post-transfection. Cells transfected with control pEYFP empty vector (EV) showed EYFP signal and transcripts were detected in cytoplasm. Cells transfected with plasmids containing partial *Men ε* sequences exhibited a strong expression of fusion transcripts but there was no EYFP signal in cells. Scale bar, 10 μm.

A


View
1 GenBank
Redraw
100
SixFrames
Frame from to Length
+2 833..1072 240
+2 2402..2629 228
-3 1280..1459 180
-1 2011..2187 177
-3 1583..1759 177
+1 1699..1860 162
-1 3571..3711 141
+3 2118..2255 138

B


EV
CMV promoter
EYFP
Men ε 1-1069
CMV promoter
Men ε
EYFP
Men ε 1-1857
CMV promoter
Men ε
EYFP
Men ε 1-2626
CMV promoter
Men ε
EYFP

C

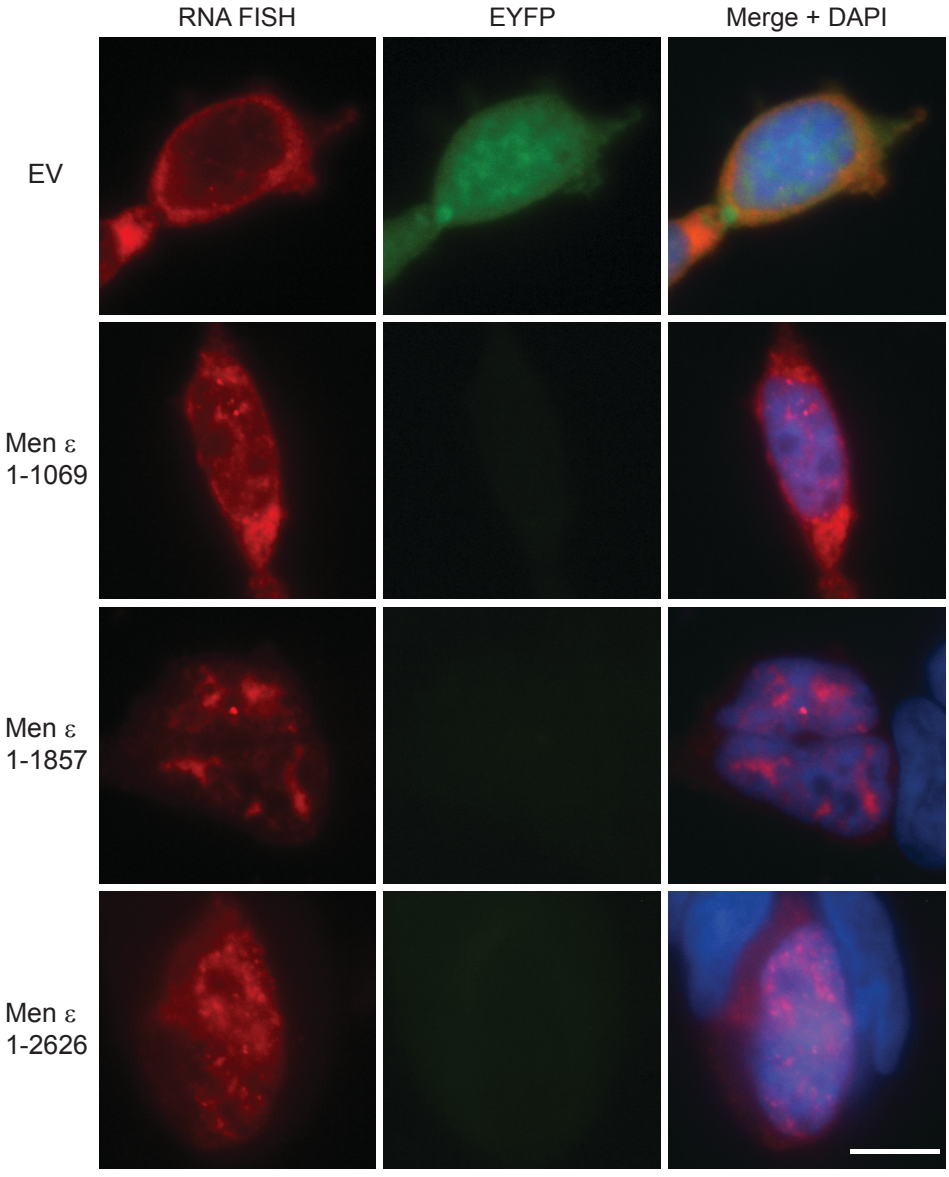

RNA FISH
EYFP
Merge + DAPI
EV
Men ε 1-1069
Men ε 1-1857
Men ε 1-2626

Figure 2.11. RNase P cleavage generates the 3' end of *Men β*. (A) Northern blot analysis using 20 μg of total RNA showed that the human *MEN β* ortholog is expressed in HeLa cells (left). The designated oligonucleotide probes were then used to roughly map the 3' end of *MEN β* (right). *Beta-Actin* was used as a loading control. **(B)** RNase H digestion followed by Northern blot analysis was used to more finely map the 3' end of *MEN β*. Oligo 1 is complementary to nt 64969484-533 of human chromosome 11. Oligo 2 is complementary to nt 64969492-541 of human chromosome 11. **(C)** The in vitro RNase P cleavage site, which corresponds to the 3' end of the mature *MEN β* transcript, was mapped by ligation-based RNA cloning procedures. Numbers at the top indicate the position on human chromosome 11. **(D)** *Men β* is a substrate for human RNase P. **(E)** Recombinant His-tagged human RNase Z cleaves *Men β in vitro*. (F) Northern blot analysis using 25 μg of total RNA from EpH4-EV, C2C12, or mouse liver showed that the *Men β* tRNA-like small RNA is selectively stabilized in liver. *U6* was used as a loading control.

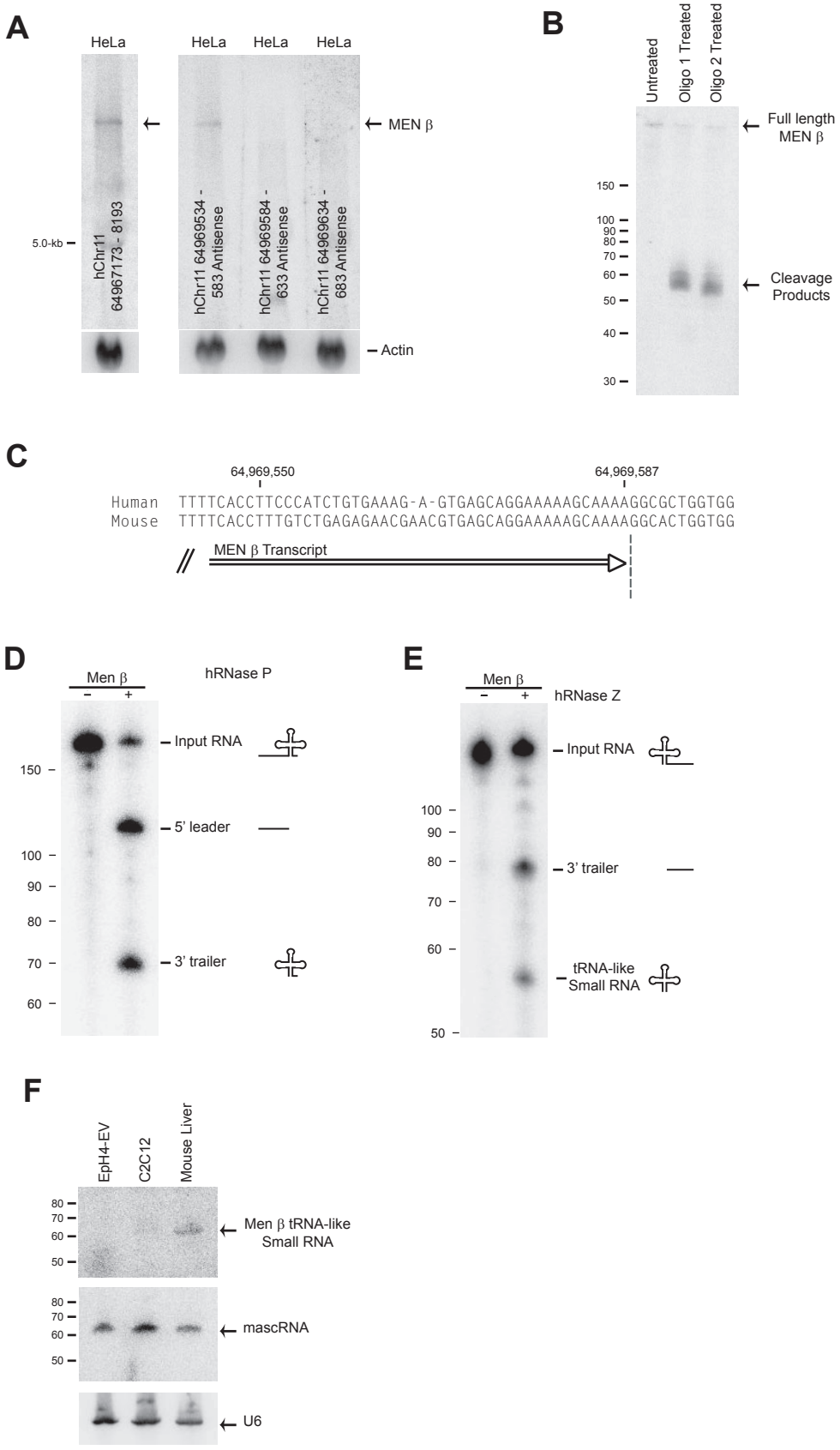
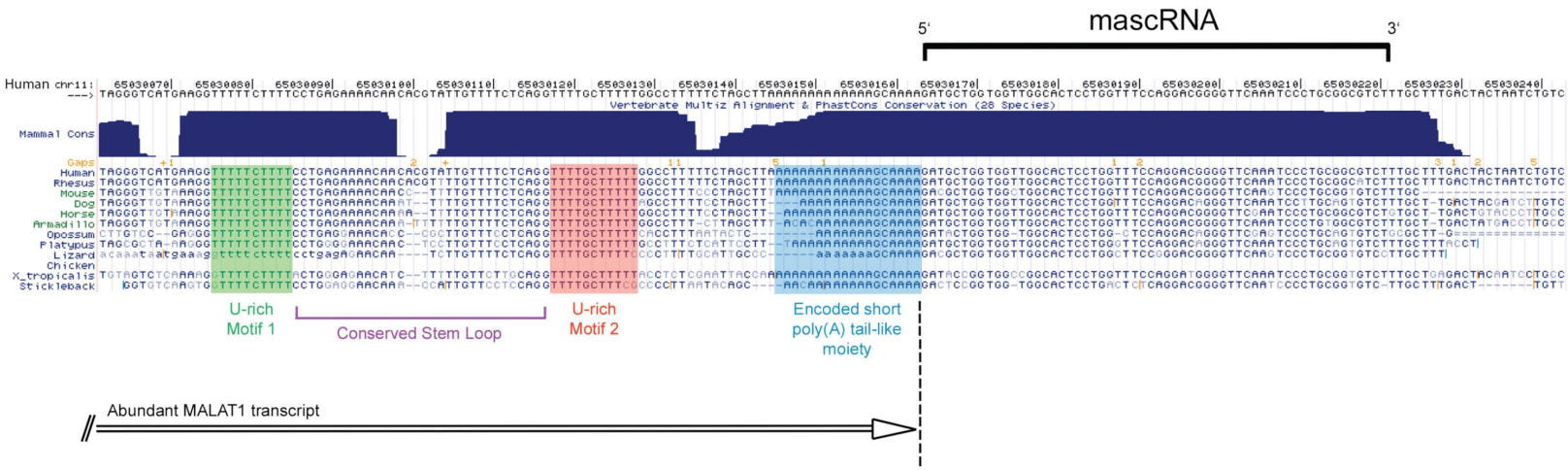
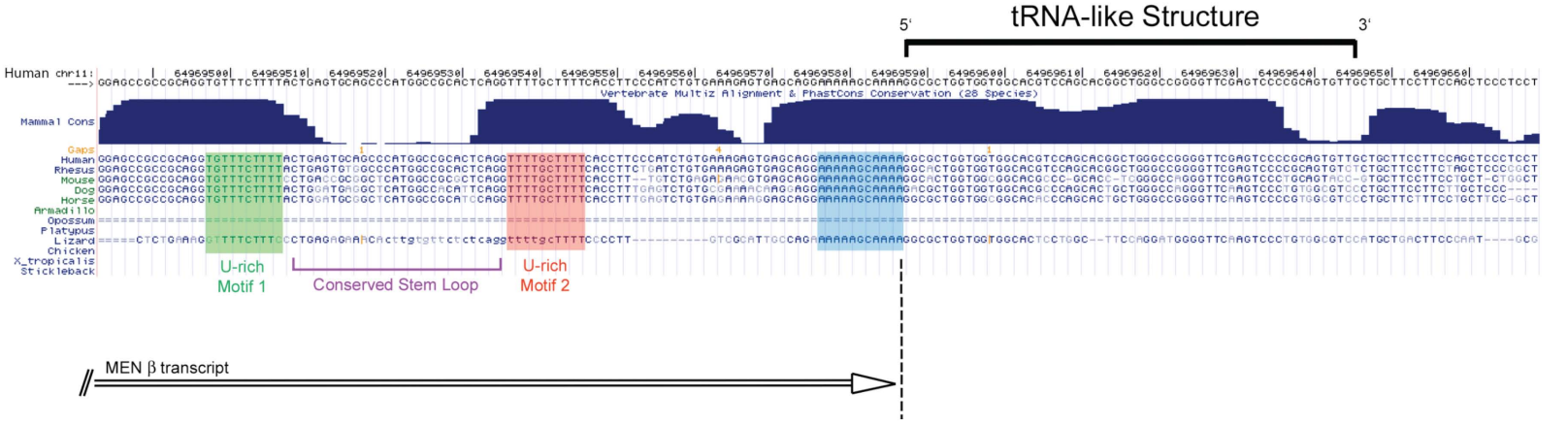

A
HeLa
HeLa
HeLa
HeLa
5.0-kb
hChr11 64967173 - 8193
hChr11 64969534 - 583 Antisense
hChr11 64969584 - 633 Antisense
hChr11 64969634 - 683 Antisense
MEN β
Actin
B
Untreated
Oligo 1 Treated
Oligo 2 Treated
Full length MEN β
Cleavage Products
150
100
90
80
70
60
50
40
30
C
64,969,550
64,969,587
Human TTTTCACCTTCCCATCTGTGAAAG-A-GTGAGCAGGAAAAAGCAAAAGGCGCTGGTGG
Mouse TTTTCACCTTTGTCTGAGAGAACGAACGTGAGCAGGAAAAAGCAAAAGGCACTGGTGG
MEN β Transcript
D
Men β
hRNase P
Input RNA
5' leader
3' trailer
150
100
90
80
70
60
E
Men β
hRNase Z
Input RNA
3' trailer
tRNA-like Small RNA
100
90
80
70
60
50
F
EpH4-EV
C2C12
Mouse Liver
Men β tRNA-like Small RNA
mascRNA
U6
80
70
60
50

Figure 2.12. The 3' end of *MEN β/Men β* is highly similar to the 3' end of *MALAT1/Malat1*. At the 3' ends of the *MALAT1* (top) and *MEN β* (bottom) loci, there are highly conserved tRNA-like structures. Immediately upstream of the tRNA-like structures are conserved poly(A)-rich tracts (designated by the blue boxes). Further upstream are two nearly perfectly conserved U-rich motifs separated by a conserved predicted stem loop. There are no AAUAAA or other canonical cleavage/polyadenylation signals present within these regions. *mascRNA* refers to *Malat1*-associated small cytoplasmic RNA (Wilusz et al. 2008).


5'
mascRNA
3'
Vertebrate Multiz Alignment & PhastCons Conservation (28 Species)
Mammal Cons
Gaps
Human
Rhesus
Mouse
Dog
Horse
Armadillo
Opossum
Platypus
Lizard
Chicken
X_tropicalis
Stickleback
U-rich Motif 1
Conserved Stem Loop
U-rich Motif 2
Encoded short poly(A) tail-like moiety
Abundant MALAT1 transcript


5'
tRNA-like Structure
3'
Vertebrate Multiz Alignment & PhastCons Conservation (28 Species)
Mammal Cons
Gaps
Human
Rhesus
Mouse
Dog
Horse
Armadillo
Opossum
Platypus
Lizard
Chicken
X_tropicalis
Stickleback
U-rich Motif 1
Conserved Stem Loop
U-rich Motif 2
MEN β transcript

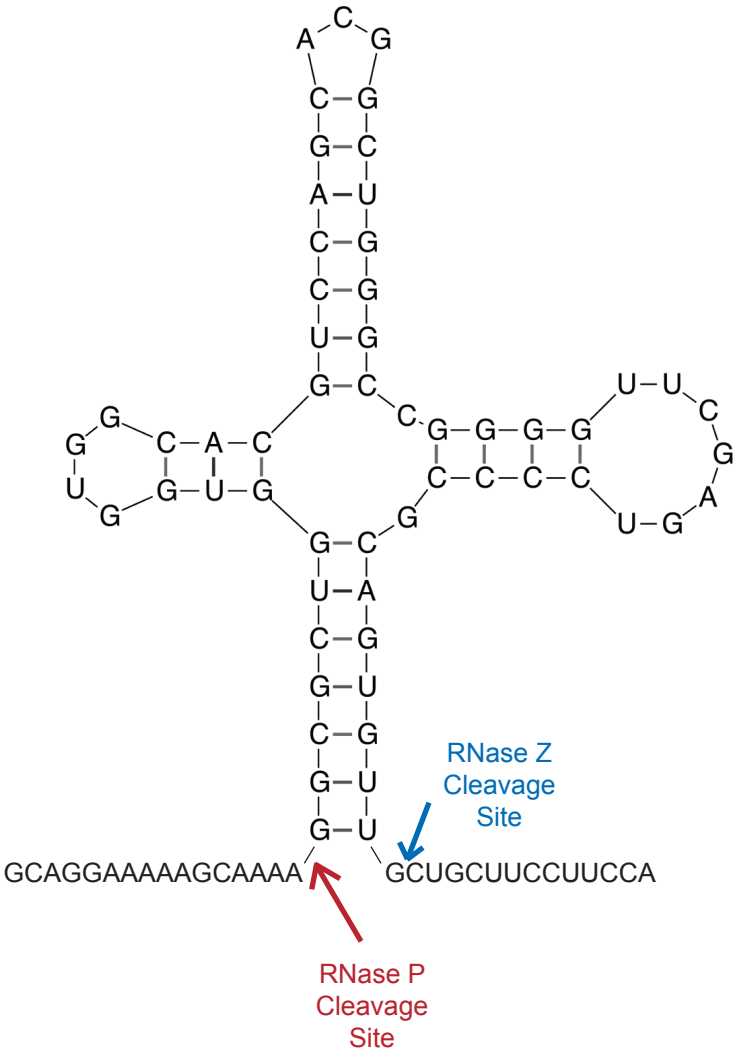


Figure 2.13. A tRNA-like structure is predicted at the 3' end of the MEN β transcript. A ~60-nt tRNA-like structure is predicted by Mfold to be present at the 3' end of the human MEN β locus. As designated, RNase P cleaves at the 5' end of the tRNA-like structure to generate the 3' end of the mature MEN β transcript. The mapped RNase Z *in vitro* cleavage site is also denoted.

Figure 2.14. *Men ε/β* are exclusively localized to paraspeckles in NIH3T3 cells while *MEN ε* localized to speckles as well as to paraspeckles in HeLa cells. (A) RNA FISH using a probe to the *Men ε/β* transcripts exhibited an exclusive co-localization with paraspeckles (EYFP-Nono, also known as p54/nrb), but not with speckles (SF2/ASF) or PML bodies in NIH3T3 cells. **(B)** In HeLa cells, the *MEN ε* transcripts are localized to speckles (SC35) in addition to paraspeckles, but not to Cajal bodies. Arrows depicts speckle localization of *MEN ε* transcripts. Scale bars, 10 μm.

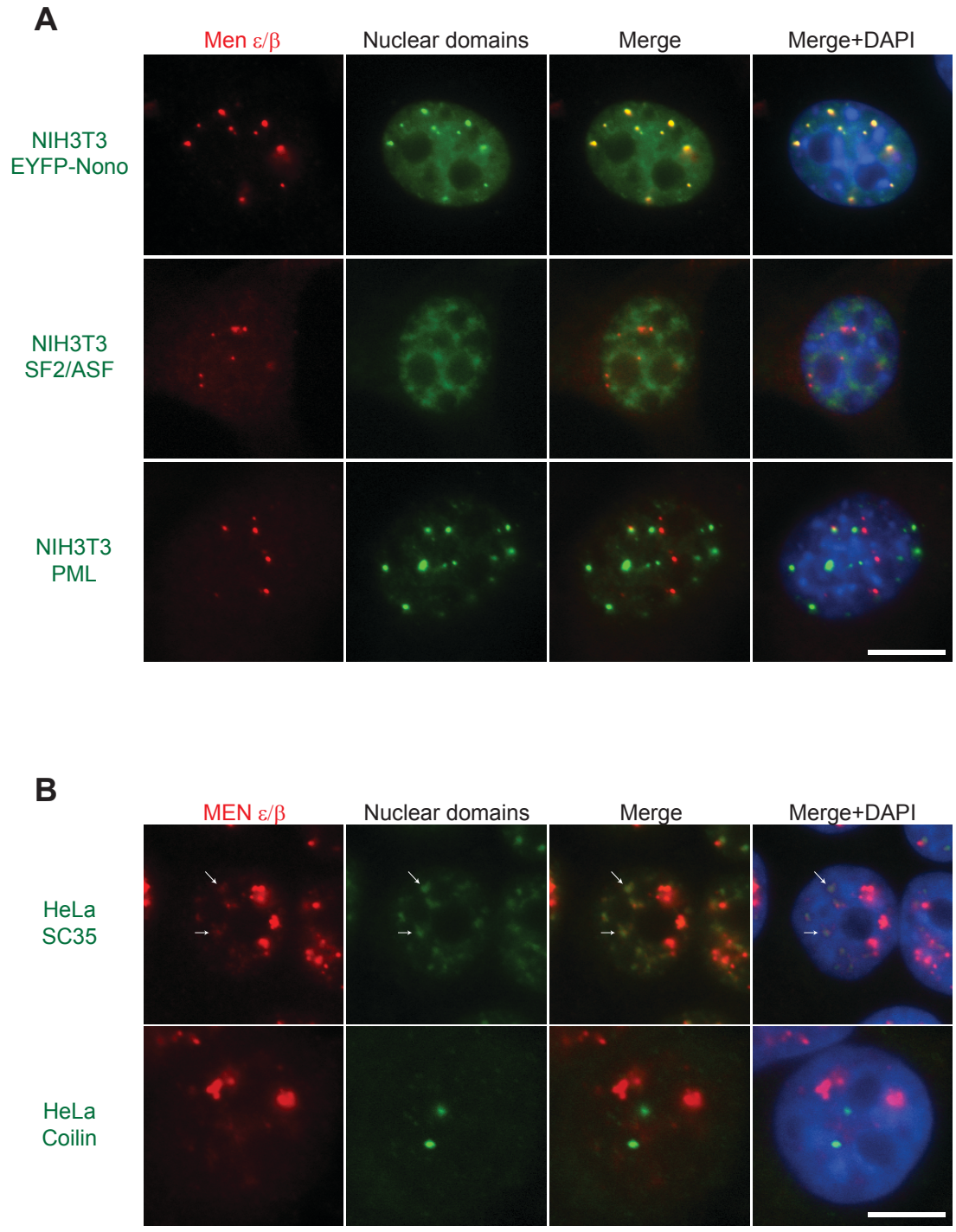
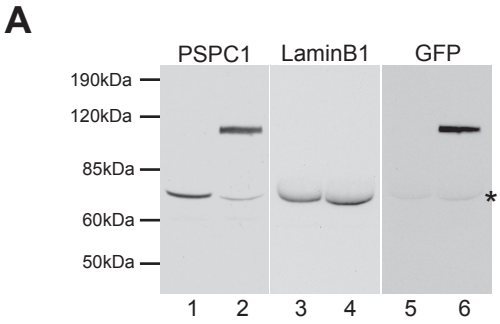
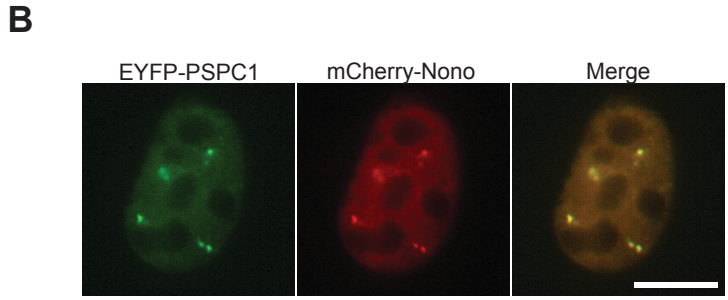

A
Men ε/β
Nuclear domains
Merge
Merge+DAPI
NIH3T3
EYFP-Nono
NIH3T3
SF2/ASF
NIH3T3
PML
B
MEN ε/β
Nuclear domains
Merge
Merge+DAPI
HeLa
SC35
HeLa
Coilin

Figure 2.15. *Men ε/β* transcripts are localized to nuclear paraspeckles. (A) A C2C12 cell line stably expressing EYFP fused to PSPC1(also known as PSP1α) was established. Immunoblotting showed that the endogenous PSPC1 level was reduced in C2C12 EYFP-PSPC1 stable cells (Lane 2, lower band) compared to in wt C2C12 cells (Lane 1). Lamin B1 served as a loading control in a duplicate blot (Lanes 3 and 4). After stripping the anti-Lamin B1 antibody, immunoblotting using an anti-GFP antibody confirmed that the band at approximately 100 kDa corresponds to EYFP-PSPC1 (Lanes 5 and 6). Residual Lamin B1 signal is denoted by *. **(B)** mCherry fused to Nono (also known as p54/nrb) was transiently expressed in C2C12 EYFP-PSPC1 stable cells. The foci of mCherry-Nono are co-localized with EYFP-PSPC1. Scale bar, 10 μm. **(C)** RNA FISH analysis showed that the *Men ε/β* transcripts are localized to paraspeckles. A probe that detected both the *Men ε* and *Men β* transcripts, as well as a probe that only detects *Men β* exhibit the same localization patterns. Scale bar, 10 μm.


A
PSPC1
LaminB1
GFP
190kDa
120kDa
85kDa
60kDa
50kDa
*
1
2
3
4
5
6


B
EYFP-PSPC1
mCherry-Nono
Merge

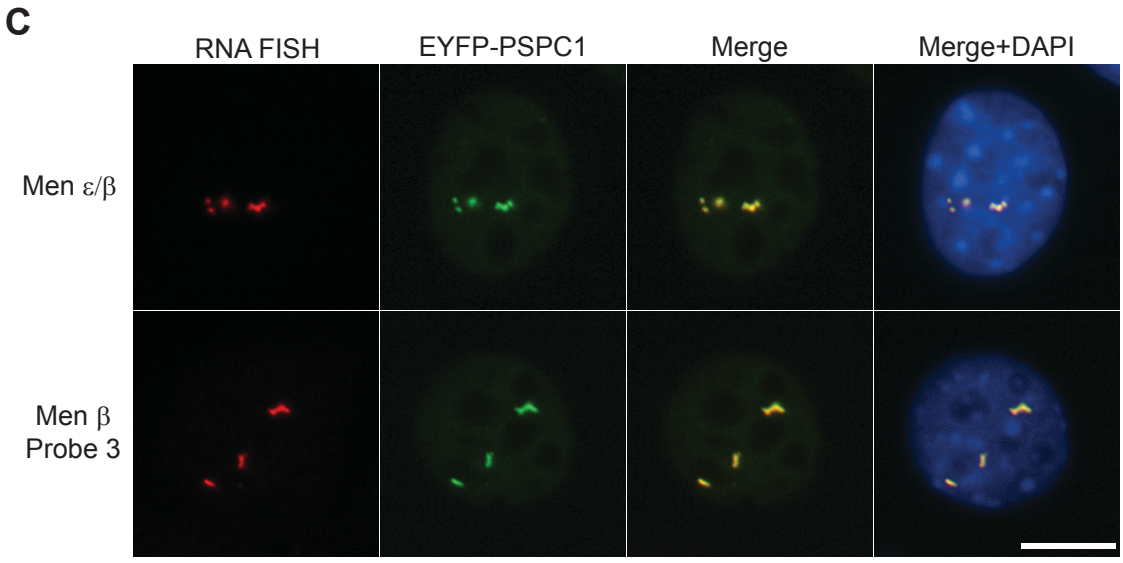
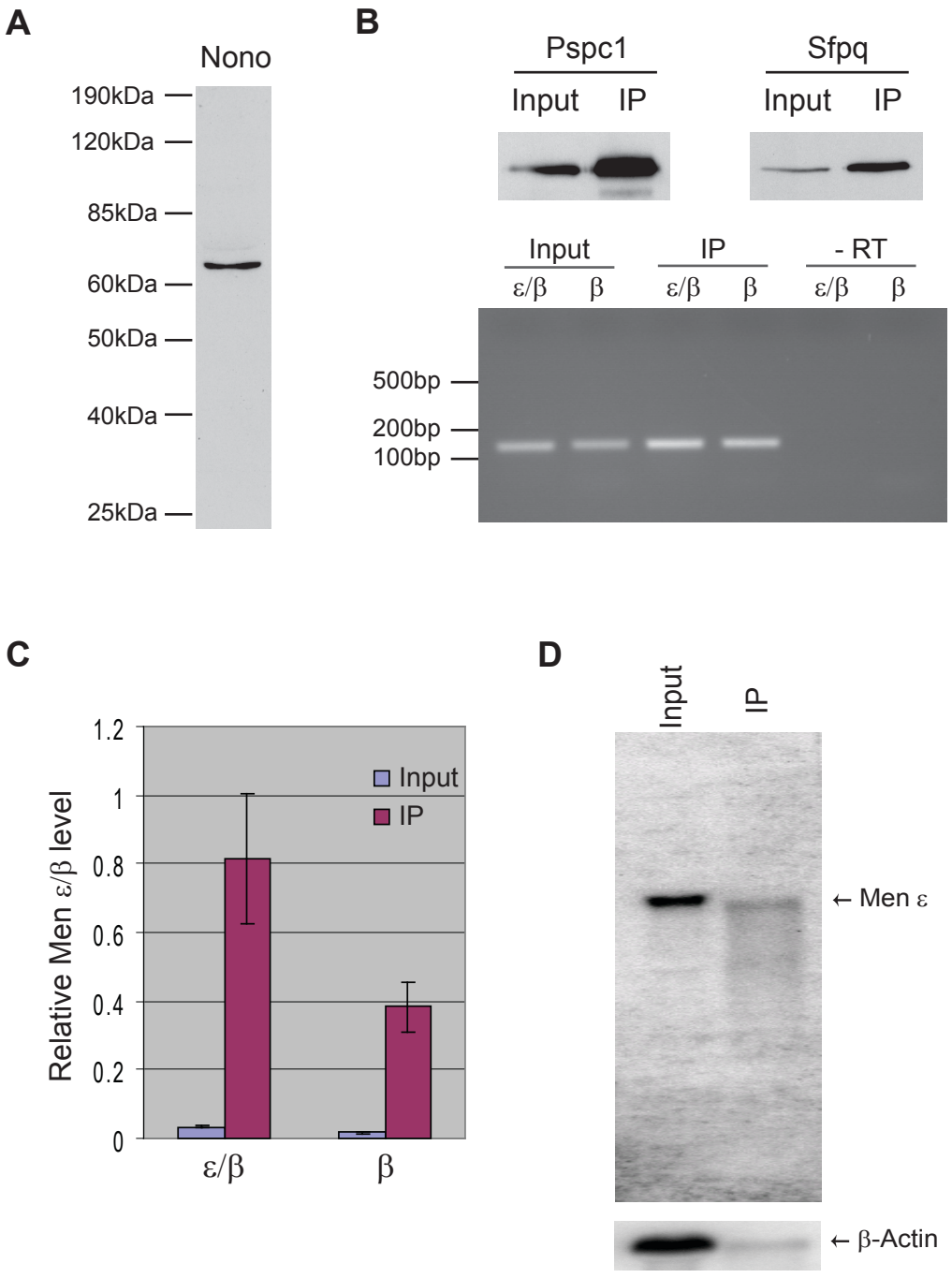

C
RNA FISH
EYFP-PSPC1
Merge
Merge+DAPI
Men ε/β
Men β
Probe 3

Figure 2.16. The *Men ε/β* transcripts directly interact with the Nono complex in C2C12 cells. (A) A mouse monoclonal antibody to Nono, designated 9-99, was generated. Immunoblotting analysis using C2C12 whole cell lysate detected a single band confirming the specificity of the 9-99 monoclonal antibody. **(B)** A co-immunoprecipitation assay revealed that the *Men ε/β* transcripts directly interact with Nono. Immunoblotting using anti-Pspc1 or anti-Sfpq antibodies after co-IP showed that Pspc1 and Sfpq (also known as Psf) directly interact with Nono. 10% input was used for immunoblotting. **(C)** cDNA was generated using random hexamers from the IP fraction. RT-PCR revealed the existence of the *Men ε/β* transcripts in the Nono protein complex. Both *Men ε/β* transcripts are ~20 fold enriched in the same IP fraction, assessed by Q-PCR. *Gapdh* was used as a normalization control in Q-PCR. The data in the histogram are shown as mean and standard deviation values of three technical replicates. **(D)** Northern analysis was performed to confirm that Men e transcripts exist in IP fraction. 10 μg of input and 2 μg of RNA from IP fraction was used. *Beta-Actin* exists at the minimal level in IP fraction.


A
Nono
190kDa
120kDa
85kDa
60kDa
50kDa
40kDa
25kDa
B
Pspc1
Input
IP
Sfpq
Input
IP
Input
IP
- RT
ε/β
β
ε/β
β
ε/β
β
500bp
200bp
100bp
C
Relative Men ε/β level
1.2
1
0.8
0.6
0.4
0.2
0
Input
IP
ε/β
β
D
Input
IP
← Men ε
← β-Actin

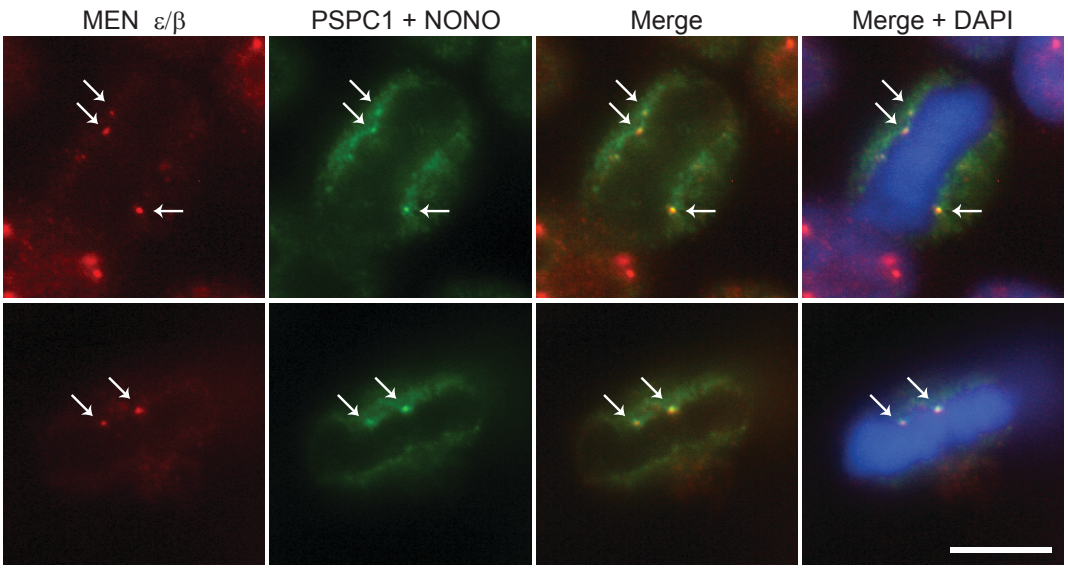


Figure 2.17 *MEN* ε/β transcripts are localized to mitotic paraspeckles. After IF using 9-99 antibody was performed on HeLa cells, *MEN* ε/β transcripts are visualized by RNA FISH. *MEN* ε/β transcripts are colocalized with PSPC1 and NONO, indicated by arrows. Two representative HeLa cells are shown here. Scale bar, 10 μm.

Figure 2.18. The *MEN ε/β* transcripts are essential for the integrity of nuclear paraspeckles. (A) Antisense oligonucleotides (ASO) were designed to knock-down *MEN ε/β* or *MEN β* expression in HeLa cells. Three ASOs (ASO 1, 2, and 3) target both *MEN ε* and *β* isoforms, while ASO 4 targets only the *MEN β* transcript. Arrows depict the positions where each ASO targets the *MEN ε/β* transcripts. **(B)** 24 hours after transfection of ASOs into HeLa cells stably expressing EYFP-PSPC1 (also known as PSP1α), a ~70% knock-down of *MEN ε/β* (ASO 1, 2, or 3) or 50% knock-down of *MEN β* (ASO 4) was achieved, as assessed by Q-PCR. *Beta-Actin* was used as a normalization control. The data in the histogram are shown as mean and standard deviation values of three independent experiments. **(C)** RNA FISH was performed 24 hours after HeLa EYFP-PSPC1 cells were transfected with a control ASO or ASOs targeting the *MEN ε/β* or β transcript, to identify cells in which the RNAs were knocked-down. In cells transfected with the control ASO, *MEN ε/β* transcripts are localized to paraspeckles. Cells transfected with ASO 3 or 4 did not show paraspeckles. Arrows indicate residual paraspeckles in a cell where knock-down of the *MEN β* transcript was not complete. Scale bar, 10 μm. **(D)** The portion of paraspeckle positive cells was reduced to ~20% by ASO 1, 2, or 3, while control ASO did not appear to influence the integrity of paraspeckles. Treatment with ASO 4 resulted in a loss of paraspeckles, although to a lesser extent. The data in the histogram are shown as mean and standard deviation values of three independent experiments. Approximately 100 cells were counted per experiment. **(E)** Knock-down of the *MEN ε/β* transcripts did not result in degradation of the PSPC1 protein. Lamin B1 serves as a loading control.

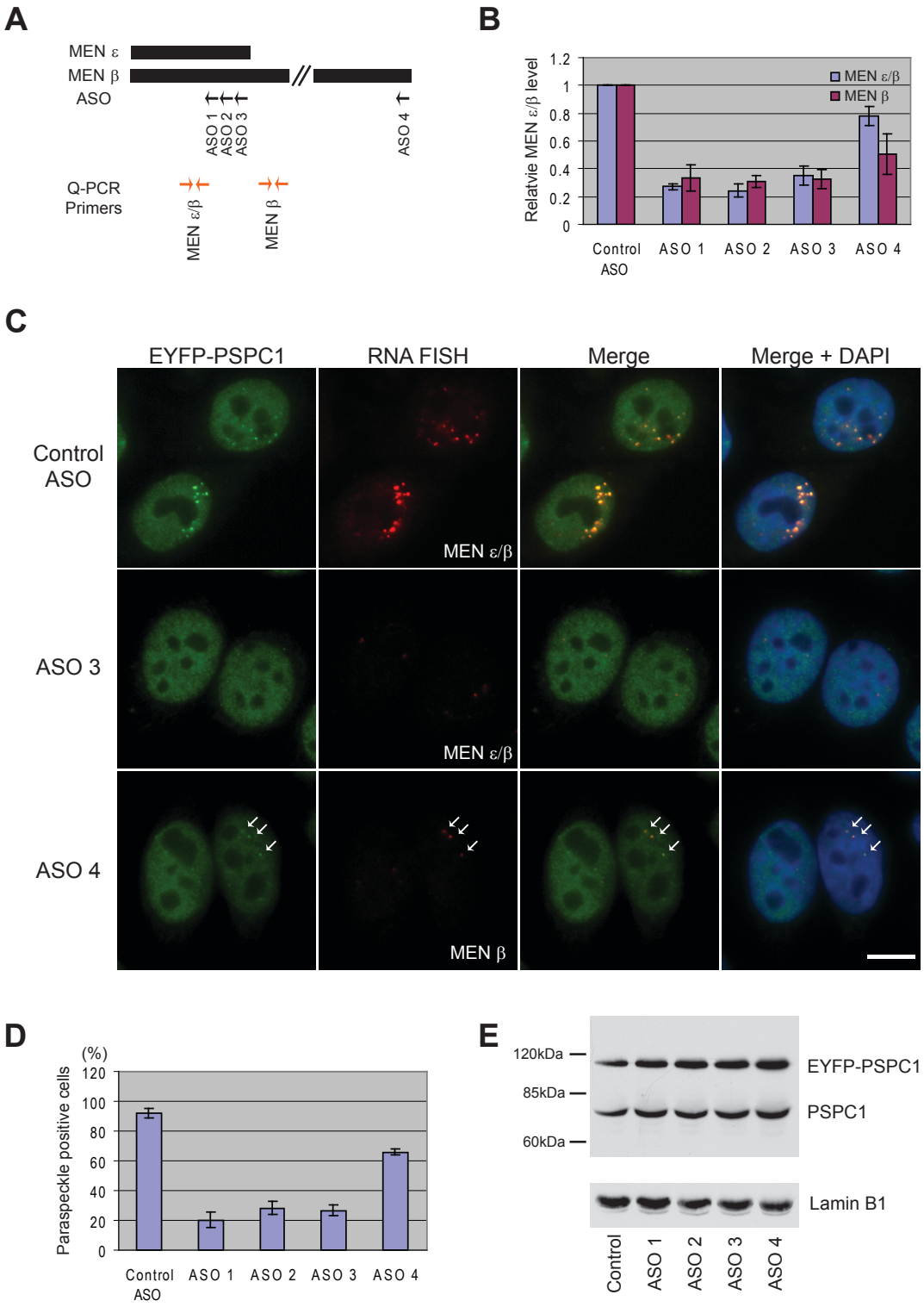
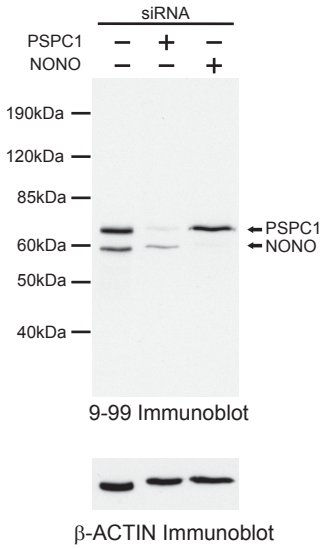

A
MEN ε
MEN β
ASO
ASO 1
ASO 2
ASO 3
ASO 4
Q-PCR Primers
MEN ε/β
MEN β
B
Relatvie MEN ε/β level
MEN ε/β
MEN β
Control ASO
ASO 1
ASO 2
ASO 3
ASO 4
C
EYFP-PSPC1
RNA FISH
Merge
Merge + DAPI
Control ASO
ASO 3
ASO 4
MEN ε/β
MEN ε/β
MEN β
D
Paraspeckle positive cells (%)
Control ASO
ASO 1
ASO 2
ASO 3
ASO 4
E
120kDa
85kDa
60kDa
EYFP-PSPC1
PSPC1
Lamin B1
Control
ASO 1
ASO 2
ASO 3
ASO 4

Figure 2.19. The *MEN ε/β* transcripts are essential for the integrity of nuclear paraspeckles. (A) A mouse monoclonal antibody, 9-99, detected both PSPC1 (also known as PSP1α, upper band) and NONO (also known as p54/nrb, lower band) in HeLa cells. PSPC1 or NONO was diminished when HeLa cells were transfected independently with siRNA to PSPC1 or NONO, respectively. *Beta-Ac*tin serves as a loading control. **(B)** Immunofluorescence using primary antibody 9-99 and Cy5 conjugated secondary antibody was performed 18 hours after HeLa cells were transfected independently with a control ASO or ASOs targeting the *MEN ε/β* (ASO1) or *β* (ASO 4) transcripts. RNA FISH using probes to *MEN ε/β* (Red) and *MEN β* (Green) was performed to identify cells in which the RNAs were knocked-down. In cells transfected with the control ASO, *MEN ε/β* transcripts are localized to paraspeckles. Cells transfected with ASO 1 or 4 did not show paraspeckles. Arrows indicate residual paraspeckles in a cell where knock-down of the *MEN ε/β* transcripts was not complete (Middle). A small population of cells transfected with ASO 4 exhibited paraspeckles that were co-localized only with *MEN ε* transcripts, but not with *MEN β* transcripts (arrows, bottom). Pseudo-colored images by OpenLab software (Improvision, Boston, MA) are shown. Scale bar, 10 μm.

A


siRNA
PSPC1 – + –
NONO – – +
190kDa
120kDa
85kDa
60kDa
50kDa
40kDa
PSPC1
NONO
9-99 Immunoblot
β-ACTIN Immunoblot

B

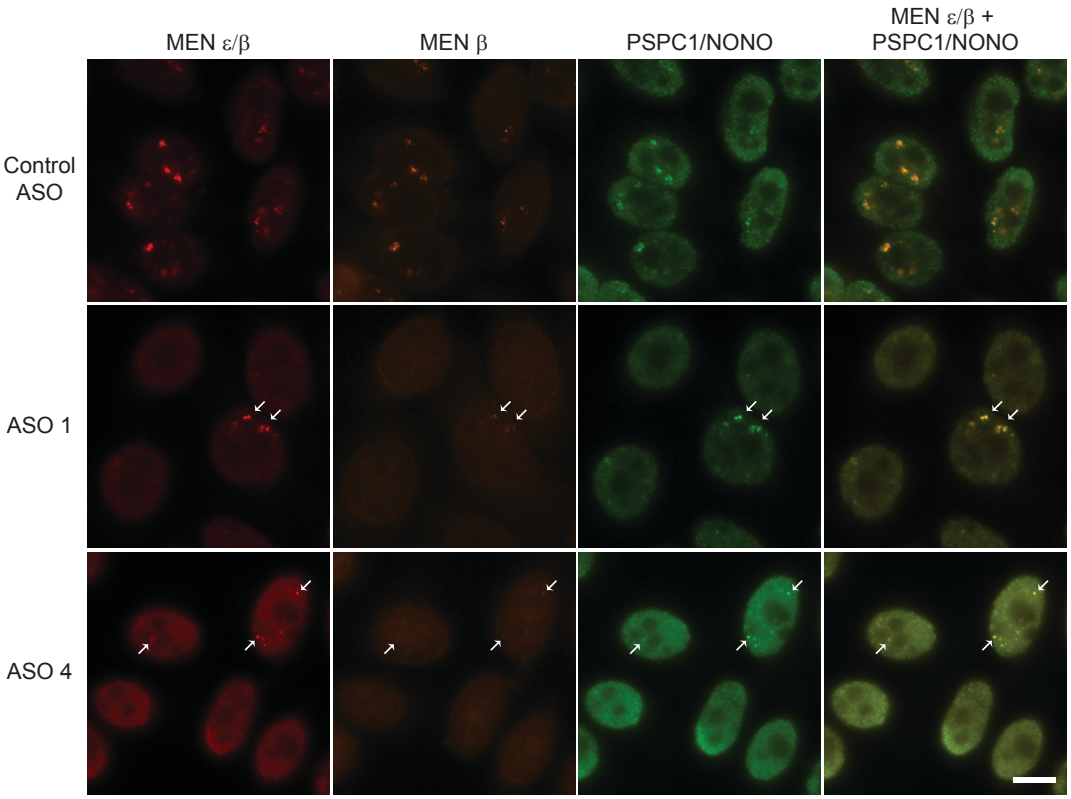

MEN ε/β
MEN β
PSPC1/NONO
MEN ε/β + PSPC1/NONO
Control ASO
ASO 1
ASO 4

Figure 2.20. The reformation of paraspeckles after release from transcriptional inhibition is suppressed in *MEN ε/β* depleted cells. (A) Upon DRB treatment for 1 hour, EYFP-PSPC1 (also known as PSP1α) relocalized to the periphery of nucleoli. The *MEN β* transcript lost paraspeckle localization, while the *MEN ε* transcript relocalized to speckles. Upon removal of DRB and recovery for 2 hours, paraspeckles reformed and co-localized with the *MEN ε/β* transcripts. **(B)** Cells were treated with ASOs to knock-down *MEN ε/β* expression prior to DRB treatment and recovery. In cells treated with a control ASO, paraspeckles reformed within 2 hours of recovery. In contrast, paraspeckles did not reform when *MEN ε/β* (ASO 3) or *MEN β* alone (ASO 4) were depleted. Arrows indicate residual paraspeckles in a cell where knock-down of the *MEN ε/β* transcripts was not complete. Scale bar, 10 μm. **(C)** Q-PCR was used to assess the ASO knock-down efficiency after 6 hours of ASO treatment. A 50 – 70% knock-down of *MEN ε/β* (ASO 1, 2, or 3) or ~70% knockdown of *MEN β* (ASO 4) was achieved. *Beta-Actin* was used as a normalization control. The data in the histogram are shown as mean and standard deviation values of three independent experiments. **(D)** The percentage of paraspeckle positive cells was reduced to 20-30 % by ASO 1, 2, or 3, while the control ASO did not influence the integrity of paraspeckles. Treatment with ASO 4 also resulted in a loss of paraspeckles, although to a lesser extent. The data in the histogram are shown as mean and standard deviation values of three independent experiments. Approximately 100 cells were counted per experiment.

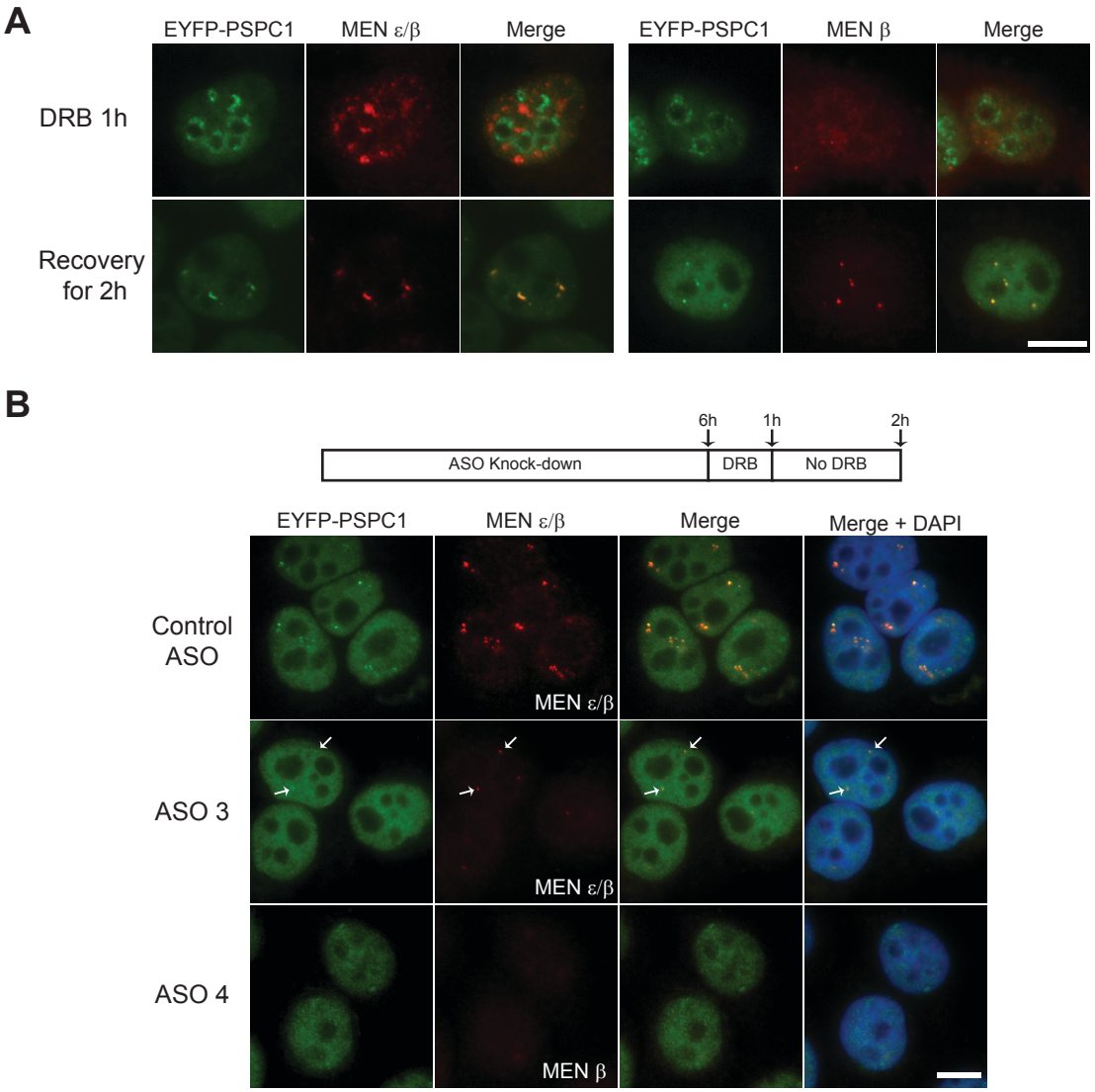

A
EYFP-PSPC1
MEN ε/β
Merge
EYFP-PSPC1
MEN β
Merge
DRB 1h
Recovery for 2h
B
6h
1h
2h
ASO Knock-down
DRB
No DRB
EYFP-PSPC1
MEN ε/β
Merge
Merge + DAPI
Control ASO
ASO 3
ASO 4
MEN ε/β
MEN ε/β
MEN β
C
Relatvie MEN ε/β level
1.2
1
0.8
0.6
0.4
0.2
0
MEN ε/β
MEN β
Control ASO
ASO1
ASO2
ASO3
ASO4
D
Parasteckle positive cells
(%)
100
80
60
40
20
0
Control ASO
ASO 1
ASO 2
ASO 3
ASO 4

Figure 2.21. The effect of DRB treatment on paraspeckle components. (A) In HeLa cells stably expressing EYFP-PSPC1 (also known as PSP1α), paraspeckles are localized to the vicinity of nucleoli. Upon DRB treatment, EYFP-PSPC1 was localized to the periphery of nucleoli, forming nucleolar caps. **(B)** When HeLa lysate was incubated with alkaline phosphatase, the migration rate of PSPC1 on SDS-PAGE changed. DRB treatment did not affect the phosphorylation state of PSPC1 (Lanes 2 and 4). Lamin B1 serves as a loading control. **(C)** *MEN ε* transcripts are localized to speckles (SC35) upon DRB treatment. Scale bars, 10 μm.

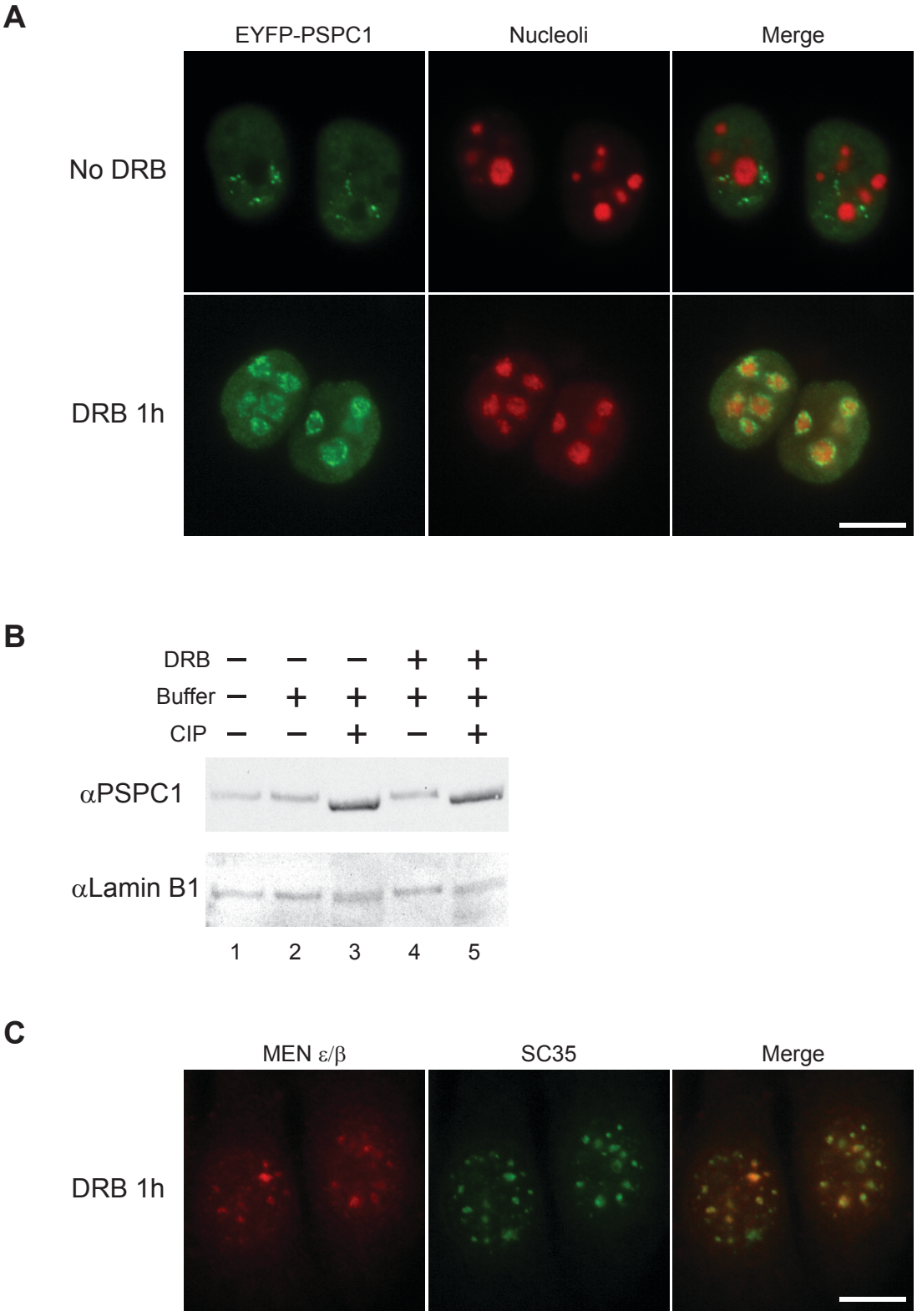

A
EYFP-PSPC1
Nucleoli
Merge
No DRB
DRB 1h
B
DRB
Buffer
CIP
αPSPC1
αLamin B1
1
2
3
4
5
C
MEN ε/β
SC35
Merge
DRB 1h

Table 2.1. Summary of microarray analysis. Non-coding transcripts showed differential expression between any two of the three time points of myoblast differentiation examined. Significantly up-regulated expression values (B-statistic > 3; M > 1) are highlighted in red and significantly down-regulated values are highlighted in green (B-statistic > 3; M < -1). The GenBank Accession IDs of the transcripts targeted by the probes are indicated. Accession IDs shown in bold blue were candidates examined by RNA *in situ* hybridization.

Target Accession ID	Target length (nt)	A-value (log2 mean expression)	24h vs Confluent		Differentiated vs 24h		Differentiated vs Confluent	
			M-value (log2 fold change)	B-statistic	M-value (log2 fold change)	B-statistic	M-value (log2 fold change)	B-statistic
AK090016	1660	13.92	-0.58	0.28	-0.42	-2.64	-1.00	7.37
AK167154	3179	13.45	1.63	21.33	0.79	8.79	2.41	28.47
AK167154	3179	13.39	1.58	20.81	0.58	4.19	2.15	26.39
AK135044	1642	13.13	3.54	30.77	1.05	9.18	4.59	35.56
AK077382	3320	12.23	1.08	12.91	0.84	8.82	1.92	23.01
AK021103	1657	12.11	1.46	22.54	0.43	2.68	1.89	27.19
AK167154	3179	12.07	1.62	22.39	1.06	14.96	2.68	31.67
AK043166	2727	12.03	-0.52	11.99	-0.81	20.10	-1.34	28.96
AK032566	5244	11.97	-0.65	4.06	-0.37	-2.10	-1.03	10.76
AK028133	795	11.90	1.88	13.67	-0.07	-7.75	1.81	12.73
AK017285	1330	11.74	1.13	26.32	-0.83	20.76	0.30	3.31
AK167154	3179	11.69	1.75	19.67	0.82	6.97	2.57	26.61
AK147816	2362	11.66	2.15	45.49	-0.29	10.23	1.86	43.63
AK004418	1181	11.59	2.46	35.14	1.73	28.70	4.19	44.92
AK167154	3179	11.44	1.53	24.10	0.54	6.23	2.06	29.57
AK076188	429	11.37	-1.23	29.16	-0.59	15.70	-1.82	36.35
AK035085	2363	11.21	2.05	21.42	-0.68	3.29	1.37	13.86
AK077756	1759	11.19	1.37	33.65	0.06	-6.52	1.43	34.44
BC032970	996	11.16	-0.67	10.21	-0.54	7.16	-1.21	20.48
AK167154	3179	11.13	1.44	20.55	0.84	11.30	2.28	28.96
AK019367	977	11.13	0.90	20.55	0.79	18.38	1.70	32.17
AK167154	3179	11.05	1.46	26.26	0.42	4.97	1.88	30.85
AK012530	1368	11.01	-0.40	5.86	-1.30	26.35	-1.70	31.20
BC046478	2364	10.98	-0.52	9.13	-0.55	10.43	-1.07	21.69
AK167154	3179	10.97	1.42	28.39	0.56	11.86	1.98	34.55
AK007736	1330	10.95	0.92	11.33	0.24	-4.28	1.16	15.12
AK013386	1255	10.92	1.21	22.05	-0.30	0.23	0.91	16.54
AK003290	1795	10.92	1.22	16.17	0.80	9.09	2.02	25.17
AK012966	2073	10.86	-0.34	1.60	-0.80	15.06	-1.13	21.01
AK044353	2558	10.82	1.24	20.26	0.21	-3.19	1.46	23.58
AK028745	3179	10.82	1.25	18.72	0.49	3.71	1.74	24.65
BC002083	2166	10.82	0.90	5.53	0.53	-0.90	1.42	12.67
AK076131	1430	10.82	1.30	27.09	0.11	-5.68	1.41	28.42
AK009590	166	10.80	1.33	14.32	1.01	9.89	2.33	24.41
AK079985	1717	10.78	-0.65	19.23	-0.61	18.10	-1.26	31.31
AK009333	1287	10.67	1.10	21.18	0.42	5.18	1.53	27.02
BC051529	1677	10.60	-0.88	6.67	-0.53	0.18	-1.41	14.18
AK079336	3004	10.54	-1.17	31.73	-0.86	26.06	-2.04	41.89
AK079985	1717	10.51	-0.69	18.48	-0.70	18.77	-1.39	31.26
AK163088	3951	10.49	1.19	18.43	0.12	-6.55	1.31	19.92
AK020483	1508	10.44	1.45	22.66	1.52	23.69	2.97	35.96
AK013363	977	10.37	0.84	20.67	0.89	21.77	1.73	33.96
AK079350	1617	10.36	-0.31	10.47	-0.72	25.77	-1.03	32.22
AK010233	1695	10.33	0.49	-2.40	1.37	11.34	1.86	16.20
AK053922	2873	10.24	-0.67	12.30	-1.02	19.90	-1.69	28.97
AK009928	1007	10.19	1.36	27.08	0.68	14.44	2.04	34.50
AK030860	3784	10.19	1.64	31.52	-0.44	8.21	1.20	25.50

AK003800	970	10.11	1.42	-2.33	2.06	1.88	3.48	8.89
AK012857	977	10.09	1.06	14.90	-0.24	-3.55	0.82	11.07
BC048953	3374	10.07	0.80	11.81	0.37	0.64	1.17	18.18
AK136368	2803	10.06	1.30	9.39	0.39	-4.26	1.68	13.52
AK036616	2727	10.05	-0.46	-1.65	-0.73	4.21	-1.19	11.18
AK077371	1806	10.04	2.11	12.25	4.18	24.72	6.28	32.11
X79508	858	10.02	-0.64	11.84	-0.42	5.64	-1.06	20.72
AK004311	803	9.99	3.87	25.74	-0.14	-7.48	3.73	24.85
AK008216	1208	9.92	0.86	13.46	1.01	16.43	1.87	27.42
BC003348	1658	9.85	1.09	32.05	0.08	-5.37	1.17	33.23
AK008278	218	9.80	-1.31	27.64	-0.82	19.28	-2.13	36.69
AK020812	2093	9.78	0.78	12.33	0.55	6.78	1.33	21.61
AK005412	876	9.78	-0.82	5.51	-0.26	-5.24	-1.08	9.43
AK014707	899	9.74	2.76	23.59	0.33	-4.87	3.09	25.50
AK189470	1040	9.69	1.86	18.25	-0.40	-3.04	1.46	13.62
AK017280	3862	9.69	1.09	12.10	-0.18	-6.24	0.91	8.83
AK006587	304	9.67	0.32	-2.61	0.97	12.19	1.29	16.76
AK021069	193	9.66	-0.51	3.84	-0.94	13.76	-1.45	21.10
AK189470	1040	9.66	1.85	20.04	-0.42	-1.73	1.42	15.05
AK020443	594	9.59	0.57	3.61	0.64	5.43	1.20	15.40
AK047833	1408	9.56	1.10	13.84	-0.25	-4.28	0.85	9.10
AK004187	1111	9.55	2.01	23.86	0.31	-3.34	2.32	26.30
AK131685	788	9.49	-0.75	4.23	-0.39	-2.85	-1.14	10.35
AK020839	969	9.47	-0.31	-2.90	-0.70	7.20	-1.01	12.64
AK084607	2943	9.46	-0.61	11.47	-0.47	7.33	-1.08	21.35
AK168411	2362	9.44	0.81	18.45	0.56	12.10	1.37	27.98
AK142897	4865	9.39	1.15	15.06	-0.63	5.44	0.52	2.22
AK080054	1792	9.38	-1.11	28.68	-0.70	20.23	-1.81	37.71
AK010197	2501	9.36	-0.85	16.54	-0.86	16.80	-1.71	29.21
AK049160	3358	9.33	0.70	-3.06	1.25	3.49	1.95	9.61
AK018410	2075	9.32	0.78	23.68	0.54	16.98	1.32	33.32
AK044589	2095	9.30	1.43	8.99	4.10	27.83	5.53	33.29
AK145182	1560	9.30	-0.89	9.50	-0.41	-0.70	-1.30	15.76
AK020181	2055	9.30	0.99	24.96	0.21	0.35	1.20	28.41
AK048970	1629	9.26	-0.90	18.76	-0.31	1.94	-1.21	23.99
AK050869	1943	9.21	-0.52	13.85	-0.51	13.47	-1.03	25.98
AK138072	4466	9.21	-1.09	16.23	-0.58	5.97	-1.67	23.81
AK006941	165	9.20	-0.90	8.72	-0.53	1.18	-1.44	16.09
AK045400	3530	9.14	1.05	28.04	-0.72	21.25	0.33	7.07
AK013046	711	9.14	-0.96	12.72	-0.79	9.69	-1.75	23.34
AK078007	1694	9.09	-0.55	11.10	-0.46	8.53	-1.02	21.76
AK004158	995	9.09	1.04	23.90	0.21	-1.03	1.25	27.10
AK034634	3267	9.02	0.08	-7.71	-1.09	12.42	-1.01	10.57
BC051496	716	8.98	1.28	17.50	0.03	-7.77	1.31	17.68
AK035405	2905	8.93	1.18	17.00	0.44	1.76	1.62	22.57
AK021103	1657	8.87	1.36	16.97	0.51	1.70	1.86	22.52
BC055348	819	8.85	0.85	15.64	0.68	11.98	1.53	26.23
AK020890	1056	8.76	1.48	16.76	0.53	1.95	2.01	22.59
AK036616	2727	8.75	-0.65	4.81	-0.71	6.27	-1.37	16.68
AK011855	941	8.75	-0.61	7.54	-0.64	8.44	-1.26	19.61
AK014762	1943	8.72	-0.70	6.51	-0.70	6.84	-1.40	18.05
AJ006837	429	8.65	-1.55	17.53	-0.59	2.41	-2.15	23.24

AK012834	1179	8.65	-0.71	10.12	-0.60	7.60	-1.30	20.67
AK009784	695	8.62	2.46	18.99	0.55	-2.34	3.01	22.30
AK085418	887	8.62	1.11	28.00	0.47	12.65	1.59	34.53
AK013643	847	8.62	-0.84	3.95	-0.40	-3.47	-1.24	9.67
BC055934	5551	8.57	0.94	8.61	0.12	-7.21	1.06	10.19
AK043270	3269	8.56	-0.70	4.10	-0.97	8.84	-1.67	17.76
AK004093	1991	8.52	1.20	14.13	1.28	15.56	2.48	27.27
AK085418	887	8.49	1.13	28.36	0.44	11.46	1.57	34.40
AK020504	1002	8.48	-1.13	14.22	-0.49	1.78	-1.62	20.50
AK015444	1054	8.47	1.68	16.69	1.38	13.44	3.07	27.57
AK137060	1123	8.46	1.69	18.78	0.82	6.64	2.51	25.85
AK142913	3185	8.45	-0.71	8.83	-0.91	13.33	-1.62	23.26
AK134627	1263	8.43	0.72	17.14	0.66	15.91	1.38	29.06
AK036616	2727	8.42	-0.73	15.17	-0.59	11.71	-1.32	25.82
AK007103	743	8.41	0.91	22.45	0.25	1.36	1.16	26.68
AK040565	3187	8.40	1.27	24.02	-0.04	-7.56	1.23	23.14
AK019951	3269	8.39	-0.72	10.47	-0.79	12.25	-1.50	23.47
AK017419	1186	8.38	0.62	2.83	0.87	7.96	1.49	16.62
BC099512	898	8.34	0.20	-0.02	0.88	23.49	1.08	27.06
AK021143	1392	8.34	0.85	18.67	0.18	-3.17	1.02	21.87
AK012332	1261	8.30	-0.96	11.30	-0.42	-0.10	-1.38	17.43
AK008369	518	8.30	-0.36	-1.36	-0.77	9.01	-1.13	14.98
AK020275	1582	8.30	1.28	19.72	-0.43	2.35	0.84	12.08
AK018285	1455	8.28	0.73	13.22	0.48	6.66	1.21	22.12
AK084964	2803	8.27	1.29	20.92	-0.39	1.65	0.91	14.21
BC050144	1574	8.26	0.25	-4.20	1.39	19.27	1.64	21.91
AK005548	489	8.26	1.72	12.26	-0.16	-7.21	1.57	11.08
AK002420	571	8.21	1.33	24.12	0.06	-7.28	1.40	24.77
AK132542	5263	8.20	2.30	21.87	0.33	-4.36	2.64	24.15
AK020467	163	8.19	2.44	14.08	-1.21	3.53	1.23	2.75
AK155993	3740	8.18	0.74	16.06	0.27	0.73	1.01	21.42
AK010897	473	8.18	0.56	3.99	0.50	2.67	1.06	13.87
AK003267	585	8.17	2.06	16.93	0.28	-5.92	2.33	18.96
AF380423	1436	8.13	1.54	14.52	0.12	-7.28	1.66	15.61
AK139055	2317	8.08	1.03	9.13	0.55	0.46	1.58	16.13
AK053519	1452	8.06	0.43	9.51	0.61	15.43	1.04	24.89
AF380423	1436	8.06	1.47	16.05	0.01	-7.84	1.48	15.90
AK009897	1265	8.02	0.64	5.41	0.48	1.89	1.12	14.34
AK003731	1043	8.00	0.50	15.52	0.56	17.49	1.06	29.05
AK015404	1274	8.00	-0.01	-8.08	2.91	32.14	2.90	32.06
AF380423	1436	8.00	1.45	18.66	0.13	-6.77	1.58	19.97
AK015435	858	7.98	1.33	9.93	1.24	9.06	2.57	21.36
AK003388	1516	7.97	0.48	-0.68	0.54	0.99	1.02	9.44
AK083217	1009	7.93	0.88	18.44	0.27	0.45	1.15	23.15
AK008679	585	7.87	-0.99	6.02	-0.63	0.35	-1.62	13.82
AK019250	752	7.86	0.86	10.63	2.09	26.70	2.94	32.67
AK008868	458	7.86	0.22	-4.68	1.37	20.34	1.59	22.64
AK021346	761	7.85	-0.60	1.83	-0.52	0.23	-1.12	10.72
AK016788	1698	7.84	0.42	3.05	0.58	8.17	1.00	16.88
BC038250	1885	7.83	0.73	14.46	0.33	2.43	1.06	20.98
AK011408	689	7.83	1.05	21.55	-0.28	0.76	0.77	15.59
AK015886	603	7.82	-1.19	8.80	-0.66	0.67	-1.85	16.01

U48388	939	7.80	2.04	15.30	-0.37	-5.16	1.68	11.55
AK008862	1852	7.78	-0.01	-8.03	1.04	14.79	1.03	14.20
AK087031	4192	7.78	1.13	25.18	0.45	9.08	1.58	31.32
AF173359	1083	7.76	1.85	26.55	-0.48	3.61	1.37	20.79
AK006240	476	7.76	1.06	22.87	-0.40	6.13	0.66	14.08
AK012741	1711	7.75	0.75	14.20	0.27	-0.65	1.02	19.37
AK082774	1238	7.74	-0.93	8.44	-0.51	0.39	-1.44	15.58
BC065396	787	7.73	0.41	-0.07	0.76	8.52	1.18	15.36
AK006324	1130	7.72	0.80	16.22	0.30	1.40	1.11	21.83
AK003710	1620	7.72	1.48	14.05	0.25	-5.66	1.73	16.57
AK011660	3660	7.71	-0.74	9.93	-0.42	1.94	-1.16	17.50
AK017368	794	7.69	1.31	9.94	0.54	-1.65	1.85	15.55
AK019533	958	7.64	1.10	26.91	-0.25	2.50	0.85	21.77
AK149390	2662	7.62	0.56	-5.96	3.00	12.33	3.55	14.58
AK009348	574	7.59	0.77	12.19	0.72	11.25	1.49	23.86
AK018927	568	7.57	-0.94	16.17	-0.52	6.11	-1.46	23.59
AK154427	1415	7.57	1.10	21.47	-0.13	-5.50	0.97	18.89
AK021368	702	7.53	1.94	7.98	0.15	-7.46	2.09	9.39
AK050598	598	7.48	0.07	-7.91	-1.10	6.97	-1.03	6.03
AK040027	1997	7.44	0.43	3.59	1.07	18.63	1.50	24.50
AK014550	2882	7.41	1.67	11.77	0.19	-7.02	1.87	13.37
AK009418	718	7.37	0.37	-4.09	2.20	19.30	2.57	21.36
AK078486	2506	7.34	-0.06	-7.88	1.49	10.27	1.43	9.07
AK171889	1146	7.32	1.24	16.47	0.38	-0.12	1.62	21.56
AK086087	1653	7.27	0.74	7.49	0.59	5.69	1.33	17.82
AK036720	2661	7.25	0.41	-3.99	0.80	3.51	1.21	8.77
BC067027	2109	7.21	0.40	-1.56	0.66	5.65	1.06	10.87
BC030050	1744	7.20	1.27	7.54	0.34	-4.41	1.61	12.67
AK020519	1274	7.19	-0.69	13.10	-0.33	2.09	-1.02	19.64
AK012993	1114	7.19	0.64	3.86	0.43	-0.22	1.07	11.95
AK053779	1396	7.19	0.22	-5.50	1.18	13.21	1.40	14.46
AK078486	2506	7.16	-0.31	-6.59	1.15	3.55	0.84	-0.22
AK014125	1523	7.16	0.45	0.23	0.63	5.15	1.09	13.02
AK038871	1667	7.13	-0.85	0.55	-0.64	-2.86	-1.49	6.36
AK155239	1983	7.12	0.69	4.44	0.53	1.25	1.21	12.21
AK021043	1024	7.11	0.77	4.71	0.58	1.81	1.34	13.81
AK160312	1572	7.11	0.95	6.82	0.43	-2.02	1.38	13.08
AK011095	879	7.10	0.12	-6.98	1.03	13.02	1.15	14.25
AK013479	531	7.10	0.39	-4.31	0.65	1.76	1.04	6.98
AK014078	882	7.09	-0.84	2.47	-0.38	-4.69	-1.21	7.39
X53631	581	7.08	1.08	12.12	0.40	-1.08	1.48	16.47
AK043958	592	7.07	0.65	1.58	0.83	5.45	1.48	12.57
AK008456	908	7.06	0.39	-1.21	0.69	6.88	1.08	13.22
AK084432	1491	7.06	0.82	8.27	0.33	-1.99	1.15	13.18
AK009927	387	7.06	0.44	-6.48	1.52	3.58	1.96	6.25
AK005958	705	7.04	0.91	3.59	0.73	2.14	1.65	13.22

III. Development of mouse monoclonal antibodies

Introduction

Antibodies are key players in the immune system in that they can specifically recognize antigens. Antibodies have become an essential reagent for many applications in modern cell and molecular biology including immunoblotting (IB), co-immunoprecipitation (Co-IP), immunofluorescence microscopy (IF), immunogold labeling, and so on. Since the *MEN ε/β* ncRNAs are localized to paraspeckles (Figure 2.15), a good antibody to mark paraspeckles in mammalian nuclei is essential for further characterization of *MEN ε/β* ncRNAs.

PSPC1 was originally identified during a proteomic study of nucleoli (Fox et al. 2002). However, PSPC1 is not localized to nucleoli. Instead, it exhibits a clustered localization of several foci in cell nuclei, leading to the discovery of paraspeckles (Fox et al. 2002). Since then, PSPC1 has been widely utilized to label paraspeckles (Fox et al. 2005; Fox et al. 2002; Prasanth et al. 2005). Although other two paraspeckle proteins NONO (also known as p54/nrb) and SFPQ (also known as PSF) have been used to label paraspeckles, they have not been considered as marker proteins of paraspeckles. There is one PSPC1 antibody available from the Lamond group (Fox et al. 2002), which is a rabbit polyclonal antibody that showed a strong activity for immunoblotting but not for immunofluorescence microscopy. Since only limited quantities can be obtained from a rabbit as a polyclonal antibody, and in addition it was neither suitable for immunoprecipitation nor other experiments that may require a large quantity of antibody.

I decided to use PSPC1 as an antigen to develop mouse monoclonal antibodies that can be utilized especially for IF and Co-IP in addition to IB. While mouse

monoclonal antibodies may have less avidity than polyclonal antibodies, their biggest advantage is that an unlimited quantity of antibody can be obtained from an established hybridoma. After several rounds of screening hundreds of hybridomas, 15 monoclonal antibodies were successfully confirmed for their specificity and activities in several applications. One of them, 9-99 recognizes Pspc1 in C2C12 lysates and PSPC1/NONO in HeLa lysates. Since it showed a strong activity in IB, IF, and Co-IP applications, I have extensively utilized this antibody during my study (Chapter 2).

Results

Preparation of PSPC1

PSPC1 (Fox et al. 2002) was subcloned into two groups of bacterial expression vectors containing either GST (Glutathione S-tranferase) tag or His tag that can be utilized during affinity purification of fusion proteins. As a pilot experiment, the expression of the fusion protein in *E. coli* BL21 (DE3) strain was tested on a small scale. After Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to 0.1 mM to induce the expression of PSPC1 fusion protein, the expression of the fusion protein was monitored by SDS-PAGE. While the expression of His-PSPC1 was not apparent (Figure 3.1B), the GST-PSPC1 fusion protein was detected after 2 hours of induction (Figure 3.1A). The high expression of GST-PSPC1 may be due to a relatively large 26 kDa module of GST helping PSPC1 to fold more easily in bacteria. In contrast, the folding of His-tagged protein may not occur properly, resulting in quick degradation of His-PSPC1 fusion protein. The highest level of GST-PSPC1 was reached within ~4 hours post induction.

The experiment was then scaled up to 400 mL culture to purify a sufficient quantity of PSPC1 to induce an immune responses in mice. The high level expression of GST-PSPC1 was again observed within 3 hours of induction by 0.1 mM IPTG (Figure 3.2A, lane 2). Since the GST module may serve as an unwanted antigen in mice, it is beneficial to cleave the GST moiety from the fusion protein before injecting into mice. The bacterial expression vector, pGEX 4T harbors a Thrombin cleavage site within linker sequences between GST and PSPC1. After cleavage, PSPC1 was eluted along with Thrombin (Figure 3.2A, lane 8). The HiTrap Benzamidine column was then utilized to remove Thrombin from elution fractions. Unfortunately, PSPC1 has an internal Thrombin

cleavage site near the N terminus and cleaved PSPC1 was not recovered after passing through HiTrap Benzamidine column (Figure 3.2A, lane 9).

Another GST expression vector pGEX 6P that utilizes Prescission protease instead of Thrombin, was tried. Prescission protease is a fusion protein of GST and a viral protease. Since Prescission protease remains bound to GST beads during the cleavage reaction, it is not necessary to further purify PSPC1 protein from the protease. Prescission protease is also modified to have an optimal temperature of 4°C so that minimal degradation of protein occurs during cleavage. PSPC1 that was cleaved away from the GST moiety eluted freely from the beads (Figure 3.2B, lanes 7 – 9). As expected, the cleaved GST moiety remained bound to the beads (Figure 3.2B, lane 6). Immunoblotting using an anti-PSPC1 antibody confirmed the identify of purified protein as PSPC1 (Figure 3.2C). From 400 mL culture, ~1 mg of PSPC1 was obtained.

Establishment of hybridomas

Serum was obtained from the tail of each mouse that had been immunized with PSPC1. IB and IF were performed using HeLa cells to test whether antibodies to PSPC1 were raised. In IB, All sera could detect PSPC1 from HeLa lysate in addition to several other nonspecific proteins (Figure 3.3A). These antibodies also showed activity in immunofluorescence microscopy when HeLa cells were pre-extracted by 0.5% TX-100 on ice for 5 – 10 min before fixation (Figure 3.3B). However, without pre-extraction, antibodies label the entire nucleoplasm without foci of potential paraspeckles (data not shown), suggesting that there is an excess amount of diffusive pool of PSPC1 protein. Since the serum from Mouse #1 showed the strongest activity in IB and the most specific

activity in IF (Figure 3.3), I decided to establish hybridomas using this mouse.

From a single attempt to fuse splenocytes with myeloma cells, unexpectedly, more than 2,000 hybridomas were established. After screening by Enzyme-Linked ImmunoSorbent Assay (ELISA) using purified PSPC1 protein, the best 96 hybridomas were selected for further screening. When IF was performed on HeLa cells, 51 hybridomas showed a paraspeckle-like pattern in nuclei determined as clusters of fine foci (Figure 3.4). Although one hybridoma showed speckle-like pattern, it was unfortunately lost during later stages (Figure 3.4). Clonal selection of these 52 hybridomas was performed to assure their identity as monoclonal antibodies, 44 hybridoma lines were successfully established and tested again for their specificity in IB and IF (data not shown). Finally 15 hybridomas were confirmed for their specificities and activities in IF (Figure 3.5) as well as IB (Figure 3.6) using HeLa cells and C2C12 cells. For IF application, 10 hybridomas showed activity in HeLa cells. Three hybridomas showed activity in C2C12 cells. Two showed activity in both HeLa and C2C12 cells (Table 3.1). For IB application, while most monoclonal antibodies recognize only PSPC1, two antibodies detect additional proteins. 1-321 appears to recognize PSF in addition to PSPC1 in HeLa cells (Figure 3.6A, lane 12) and another unknown protein in C2C12 cells (Figure 3.6B, lane 7). Interestingly, 9-99 recognized only Nono in C2C12 lysate, while it recognizes PSPC1 as well as NONO in HeLa lysate (Figures 2.19A and 3.6B)

Characterization of 15 hybridomas

Mouse monoclonal antibodies are classified into several isotypes depending on heavy chain and light chain (Harlow and Lane 1988). There are five heavy chains

(γ, α, μ, δ, and ε) and two light chains (κ and λ). IgG is further grouped into IgG_1, IgG_{2a}, IgG_{2b}, and IgG_3. The isotype of the mouse monoclonal antibody is useful information to maximize its activity in many applications including IB, IF, and Co-IP. In the Co-IP application, two different types of beads, protein A beads and protein G beads, are broadly used to pull down the antibody-protein complex. While protein A beads are recommended for IgG_{2a}, IgG_{2b}, and IgG_3, protein G beads are used to bind IgG_1 (Harlow and Lane 1988). The heavy chain isotypes of these 15 hybridomas were determined (Table 3.1). All light chains appreared to be κ light chain (data not shown). PSPC1 harbors two RRMs (RNA recognition motif) and one NOPS motif that is found at the carboxyl terminus of NonA (Drosophila homolog of SFPQ) and PSPC1 (Figure 3.7A). Several deletion mutant of PSPC1 were made and their expression was monitored by SDS-PAGE (Figure 3.7B). By immunoblotting using these deletion mutants, epitopes were roughly determined (Figure 3.7C). All hybridomas specifically recognize either the N terminus or the C terminus of PSPC1 (Figure 3.7C), suggesting that extra proteins detected by mouse sera may have RRMs or NOPSs (Figure 3.4A).

Immunoprecipitation is a very powerful technique to purify a specific protein in a complex with other proteins, DNAs, or RNAs. Seven hybridomas were tested in IP applications using C2C12 lysate. Although most hybridomas were not able to pull down the protein complex, 4 monoclonal antibodies including 22-67, 27-107, 31-89, and 48-47, pulled down Nono along with Pspc1 (Figure 3.8A). Hybridoma 9-99 efficiently pulled down Sfpq along with Nono while 31-89 and 48-47 did so marginally. Since 9-99 showed higher specificity in IB and stronger activity in Co-IP, it has been extensively utilized during my dissertation study (Chapter 2).

Hybridoma 9-99

When injected into mice, hybridomas can grow as a form of liquid tumor, called ascites (Harlow and Lane 1988). While only a low concentration of antibody (up to 50 μg/mL) can be obtained from the supernatant of hybridoma culture, ascites can generate an extremely large quantity of antibodies (up to 9 mg/mL of specific antibody, 10 – 20 mL per mouse). Hybridomas 9-99 was utilized to generate ascites. Although the majority of antibodies in ascites consist of monoclonal antibodies generated from injected hybridomas, there are still endogenous pools of heterogeneous antibodies in addition to impurities from the blood plasma. A high concentration of a salt such as ammonium sulfate can reduce the solubility of protein in water, resulting in precipitation of proteins. Ammonium sulfate precipitation of antibodies has been widely adopted to remove these impurities that may cause degradation of antibodies. After ammonium sulfate precipitation of 5 mL of ascites, ~30 mg of antibodies were recovered. However, there were still high molecular weight contaminants (Figure 3.9). Antibodies were further purified using protein A beads. From 0.75 mL (8.5 mg of protein) of ammonium sulfate precipitated fraction, ~3.9 mg of pure antibodies were recovered (Figure 3.9).

Ascites 9-99 was quantitatively tested for its capacity to pull down the Nono complex. Under harsh conditions of Co-IP using RIPA buffer, 9-99 efficiently pulled the Nono complex. The amount of Co-IP fraction increased with the amount of antibodies used (Figure 3.8B).

Discussion

I have successfully developed mouse monoclonal antibodies using PSPC1 as an antigen. These antibodies can be used in many applications including IB, IF, and IP. While most antibodies showed specificity to PSPC1/Pspc1, a few can recognize additional proteins including NONO and SFPQ; 9-99 recognized PSPC1 and NONO in HeLa lysate but only Nono in C2C12 lysate. 1-321 recognized SFPQ in additional to PSPC1 in HeLa lysate. 9-99 monoclonal antibody was intensively utilized to label paraspeckles in HeLa cells and pull down the Nono complex from C2C12 lysate during my study (Chapter 2). However, antibodies still remain to be tested for their IP application using human cell lysate.

Antibodies developed during this study have great potential for use in other applications. Since 9-99 recognize specifically Nono, but not Pspc1 in mouse cells, this antibody will be useful when it is needed to study these two proteins separately. Co-IP experiments using antibodies developed during this study can identify protein and RNA components of paraspeckles, which will elucidate the cellular functions of this nuclear domain.

A recent report showed that NONO and SFPQ are also essential components of paraspeckles in addition to the *MEN ε/β* transcripts (Sasaki et al. 2009). Knock-down of these protein resulted in disruption of paraspeckles as well as degradation of *MEN β*, but not *MEN ε* transcripts in HeLa cells. However, knock-down of PSPC1 did not result in disruption of structures where NONO, SFPQ, and the *MEN ε/β* transcripts were colocalized together. This study raised a question about PSPC1 being a marker protein of paraspeckles. Another study showed that the ectopic expression of the *MEN ε* transcripts

induced an increase in numbers of paraspeckles, visualized by EYFP-PSPC1 (Clemson et al. 2009). Further study of how these three proteins and *MEN* ε/β transcripts establish paraspeckles will address which protein or transcripts should be considered as the best marker of paraspeckles. Two currently available monoclonal antibodies NC5 (anti-Nono) and FC23 (anti-Sfpq) are limited in their application in that they can recognize only mouse proteins. Development of more monoclonal antibodies to NONO and SFPQ is required to study paraspeckles in further detail.

.

Methods

Cloning of PSPC1 expression plasmids

PSPC1 sequence (Fox et al. 2002) was subcloned into pGEX4T-3 (GE Healthcare, Piscataway, NJ), pGEX6P-3 (GE Healthcare, Piscataway, NJ), pQE-30 (Qiagen, Valencia, CA), and pRSET-A (Promega, Madison, WI).

Induction test

PSPC1 expression plasmids were transformed into *E. coli* BL21 (DE3). Three hundred μL of overnight culture was seeded into 3 mL of LB with Ampicilin at 37°C. After cells were allowed to grow for 2 hours with vigorous shaking, IPTG was added to final concentration of 1 mM. 500 μL of bacteria was harvested every hour for 3 – 4.5 hours and analyzed by SDS-PAGE.

Large scale preparation of PSPC1

BL21 (DE3) transformed with pGEX 6P-PSPC1 was inoculated for overnight growth. Five mL of bacteria was seeded into 400 mL of LB/Amp in 2 L flask at 37°C with vigorous shaking. Approximately 2 hours later, OD_{595} was monitored. When OD_{595} reached to 0.5 – 0.6, IPTG was added to 0.1 mM to induce the expression of PSPC1 for 2 hours. After cells were harvested and resuspended in 20 mL of ice cold PBS, sonication was performed to break the bacterial cell wall. Insoluble fraction was removed by centrifugation at 20,000 g for 20 min. Supernatant was further filtered through 0.45 μm polyethersulfone filter (Nalgene, Rochester, NY). 1.5 mL (bedding volume) of GST beads (GE Healthcare, Piscataway, NJ) was prepared by washing twice in 10 mL of cold

PBS. GST-PSPC1 was allowed to bind to the GST beads at room temperature for 30 min with gentle rotation. After collected in column and washed with 25 mL of cold PBS, beads were equilibrated in 10 mL of cleavage buffer (50mM Tris pH 7.0, 150 mM NaCl, 1 mM EDTA, 1 mM DTT). One hundred units of Prescission protease was added and incubated at 4°C overnight. PSPC1 protein was eluted with 5 mL of cleavage buffer. Fractions that contain a significant amount of protein, measured by OD_{260}, were pooled and finally quantified using Bradford methods. PSPC1 in the cleavage buffer was directly injected into mice.

Establishment of hybridomas

Immune responses in three mice were induced as described elsewhere (Harlow and Lane 1988). After blood from each mouse tail was tested for its reactivity to PSPC1 protein by IB and IF, PSPC1 was injected one more time as a final boost. Two weeks later, the mouse was sacrificed and hybridomas were generated as described elsewhere (Harlow and Lane 1988). Two thousand hybridomas were screened by ELISA using PSPC1 protein. Several rounds of tests and clonal selection, 15 hybridomas were finally established.

Immunoblotting

Proteins are separated by 8 – 12% SDS-PAGE and transferred to a nitrocellulose membrane (GE Healthcare, Piscataway, NJ). After briefly rinsing in water, a membrane was blocked in 5% skim milk in 1X TBST (10 mM Tris (pH7.6), 150 mM NaCl, 0.05% Tween-20) at room temperature for 1 hour. Primary antibodies were added to appropriate

concentrations in 5% skim milk in 1X TBST and allowed to bind for 1 hour. After washing three times in 1X TBST, secondary antibodies were added in 5% skim milk in 1X TBST for 1 hour and washed three times in 1X TBST.

Immunofluorescence microscopy

Cells were grown on acid washed glass coverslips for 24 – 48 hours prior to any treatment. Cells were pre-extracted in CSK buffer (10 mM PIPES (pH 6.8), 100 mM NaCl, 3000 mM sucrose, 3 mM $MgCl_2$, 2 mM EDTA) containing 0.2% TX-100 on ice for 5 min and then fixed in 2 – 4% paraformaldehyde in PBS at room temperature for 10 min. After washing twice in PBS, nonspecific binding sites were blocked in 1% normal goat serum (NGS) in PBS for 20 min. Primary antibodies were incubated at 4°C for 1 hour. After washing three times in PBS/ 1% NGS, secondary antibodies were added at room temperature for 1 hour. Cells were washed three times in PBS and nuclei were counterstained with DAPI. Cells were examined using an Axioplan 2i fluorescence microscope (Carl Zeiss, Thornwood, NY) equipped with a 40X/1.3 N.A. and a 63X/1.4 N.A. objective lens and Chroma filters (Chroma Technology, Brattleboro, VT). OpenLab software (Improvision, Boston, MA) was used to collect digital images from an ORCA cooled charge-coupled device camera (Hamamatsu, Bridgewater, NJ).

Isotyping of hybridomas

Isotyping of each hybridoma was performed using Hybridoma subisotyping kit (Calbiochem, San Diego, CA) and IsoQuick strips (EnviroLogix, Portland, ME) as per manufacturer’s instruction.

Epitope mapping

Series of PSPC1 deletion mutants were cloned into pGEX 6P-3. GST fusion proteins were expressed on a small scale as in induction test (see above). Bacterial lysate was utilized in IB to determine epitopes for each monoclonal antibody.

Screening for Co-IP application

C2C12 cells were plated 2 days prior to the assays (70% confluency). After briefly rinsing in PBS, 1 mL of lysis buffer (50 mM Tris (pH 7.4), 150 mM NaCl, 5 mM EDTA, 50 mM NaF, 0.5 % NP-40, 1 mM Na_2VO_4) supplemented with Complete mini protease inhibitor (Roche Diagnostics, Indianapolis, IN) was added to each 10 cm plate. Cells were lysed on ice for 30 min with gentle rocking. Lysate was collected by centrifugation at 20,000g, 4°C for 15 min. Five hundred μL of supernatant was transferred to a new tube. Fifty μL of hybridoma supernatant was added. Binding occurred at 4°C for 1.5 hours with gentle rotation. 10 μL (bedding volume) of protein A sepharose beads (GE Healthcare, Piscataway, NJ) were added and incubated at 4°C for 1.5 hours with gentle rotation. Beads were collected by centrifugation at 500 g, 4°C for 1 min. Beads were washed three times in Lysis buffer. Proteins were eluted by boiling in 1X Laemmli buffer and analysed by SDS-PAGE.

To titrate the amount of 9-99 antibodies required to pull down the maximum amount of the Nono complex from C2C12 lysate, 300 μg of purified ascites 9-99 (see below) was conjugated to 1 mL of Dynabeads protein A (Invitrogen, Carlsbad, CA) as described elsewhere (Harlow and Lane 1988). C2C12 cells were briefly rinsed in PBS and lysed in 1mL of RIPA (50 mM Tris (pH 8.0), 150 mM NaCl, 1% NP-40, 0.5%

sodium deoxycholate, 0.1% SDS) supplemented with Complete mini protease inhibitor (Roche Diagnostics, Indianapolis, IN) on ice for 30 min. After cell debris was removed by centrifugation at 20000g, 4°C for 15 min, 10 – 100 uL of Dynabeads protein A that had been conjugated with purified 9-99 monoclonal antibodies were added. Binding occurs at 4°C for 2 hours with gentle rotation. Beads were washed four times in RIPA buffer. Proteins were eluted in 1X Laemmli buffer and analysed by 8% SDS-PAGE.

Purification of 9-99 antibody

Five mL of 9-99 ascites was centrifuged at 20,000g, 4°C for 3min. Supernatant were filtered through 0.45 μm polyethersulfone filter (Nalgene, Rochester, NY). Five mL of 100 mM Tris (pH 7.4) was added. Five mL of saturated ammonium sulfate was added drop by drop with stirring. Precipitation of contaminant proteins occurred at 4°C for 6 hours with gentle rotation. After centrifuged at 3000g, 4°C for 60 min, supernatant was transferred to a new tube. Five mL of saturated ammonium sulfate was added drop by drop with stirring. Precipitation of antibodies occurred at 4°C for overnight with gentle rotation. Antibodies were collected by centrifugation at 3000g, 4°C for 30 min and resuspended in 2 mL of PBS. After dialysis against 500 mL PBS at 4°C for 4 hours with vigorous shaking, antibodies were collected by brief centrifugation. Sodium azide was added to 0.02%.

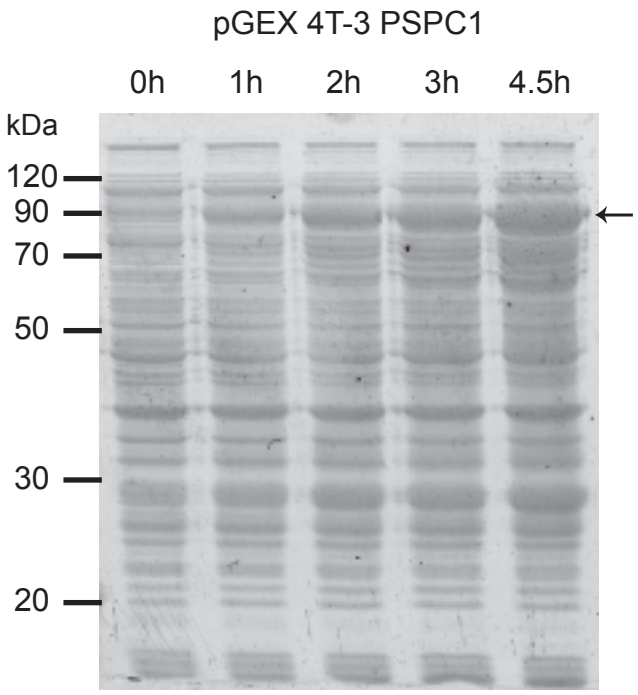
Further purification of 9-99 antibodies was performed using sepharose protein A beads (GE Healthcare, Piscataway, NJ) as described elsewhere (Harlow and Lane 1988). Briefly from 0.75 mL of antibodies prepared above, ~4 mg of antibodies were purififed.

Antibodies

For immunofluorescence, antibodies were used at the following concentrations; mouse blood at 1:50 and hybridoma supernatant at 1:1. For immunoblotting, antibodies were used at the following concentrations; mouse blood at 1:1,000, hybridoma supernatant at 1:1, NC5 (anti-Nono) at 1:50, FC23 (anti-Sfpq) at 1:50, anti-PSPC1 at 1:2,000 – 4,000, anti-β-Actin (Sigma, St. Louis, MO) at 1:5,000, HRP-DαM at 1:20,000, and HRP-GαR at 1:25,000.

Figure 3.1. The expression of PSPC1 fusion protein was tested on a small scale. (A) The expression of GST-PSPC1 became apparent at 2 hours post-induction by 0.1 mM IPTG. GST-PSP1 accumulated to a higher level after 4.5 hours of induction (arrow). **(B)** The expression of His-PSPC1 was not detected within 3 hours of induction by 0.1 mM IPTG.

A


pGEX 4T-3 PSPC1
0h
1h
2h
3h
4.5h
kDa
120
90
70
50
30
20

B

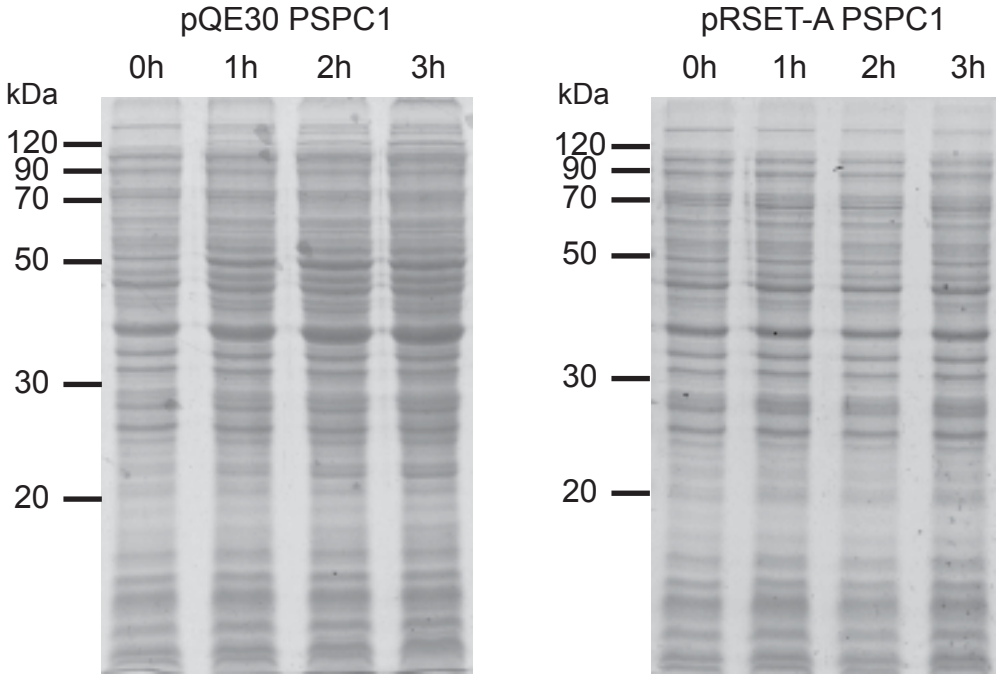

pQE30 PSPC1
0h
1h
2h
3h
kDa
120
90
70
50
30
20
pRSET-A PSPC1
0h
1h
2h
3h
kDa
120
90
70
50
30
20

Figure 3.2. PSPC1 protein was purified on a large scale. (A) PSPC1 protein was expressed in bacteria. After binding to GST beads and washing in PBS, GST-PSPC1 was eluted from half of the beads (lane 6, arrow). Another half of the beads were subject to Thrombin cleavage. While GST (lane 7, open arrowhead) remained bound to beads, PSPC1 was eluted. However, PSPC1 protein was lost after passing through a HiTrap Benzamidine column. Lane 1: pre-induction, lane 2: 3 hour induction, lane 3: unbound fraction to GST beads, lane 4: washed fraction in PBS, lane 5: GST beads post-elution, lane 6: eluted fraction, lane 7: proteins remained bound to GST beads after Thrombin cleavage, lane 8: eluted fraction after Thrombin cleavage, lane 9: Pass through HiTrap Benzamidine column. **(B)** The expression of GST-PSPC1 was induced by 0.1 mM IPTG for 3 hours (lane 2, arrow). After washing in PBS, Prescission protease cleaved PSPC1 away from the GST moiety (lane 6, open arrowhead) that still remained bound to the GST beads. PSPC1 was eluted after the cleavage reaction (lanes 7 – 9, arrowhead). Lane 1: pre-induction, lane 2: 3 hour induction, lane 3: unbound fraction to the GST beads, lane 4: washed fraction in PBS, lane 5: washed fraction in cleavage buffer, lane 6: proteins remained bound to beads after Prescission protease cleavage, lane 7 – 9: eluted fractions. **(C)** Immunoblotting using an anti-PSPC1 antibody was performed to track PSPC1 protein during purification and confirmed that the major protein in the eluted fraction is PSPC1 (lanes 7 – 9, arrowhead).

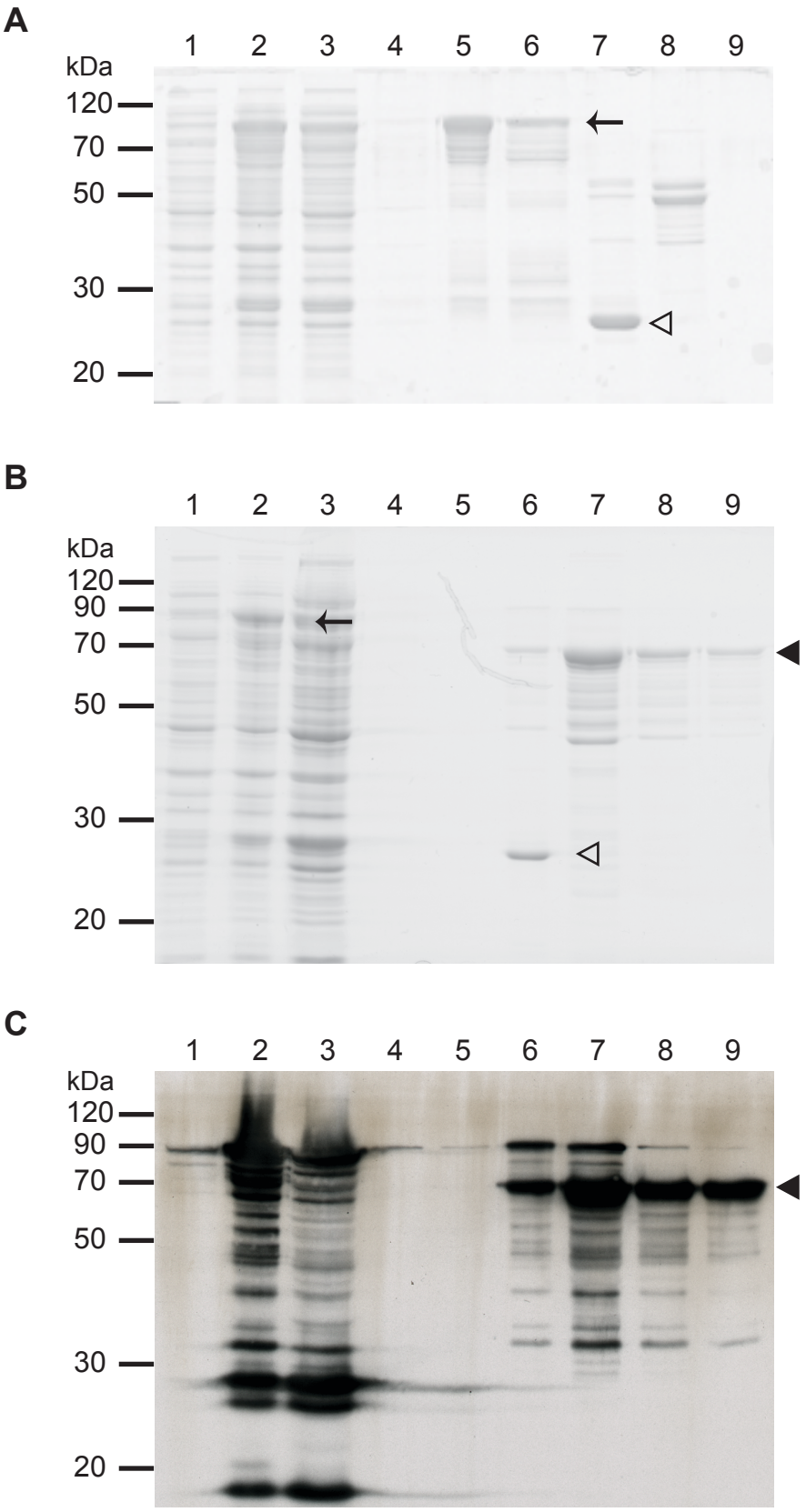
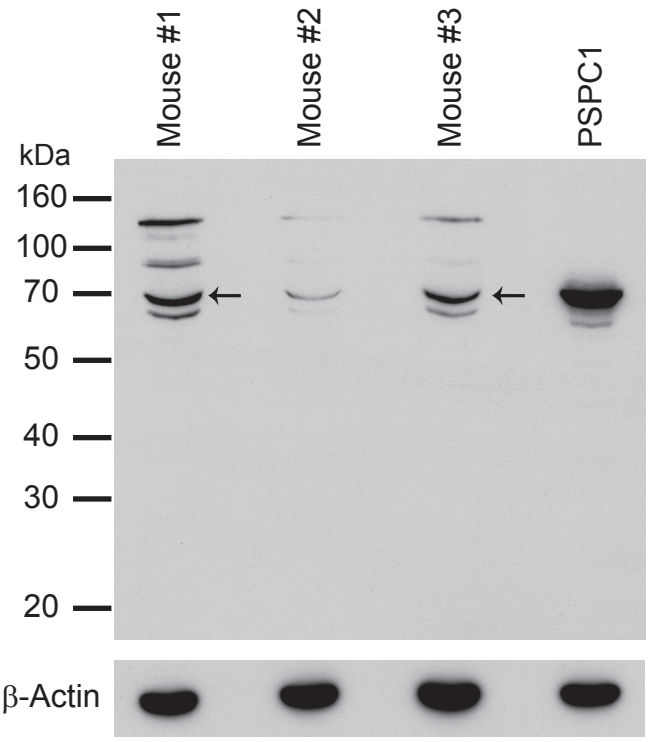

A
1 2 3 4 5 6 7 8 9
kDa
120
70
50
30
20
B
1 2 3 4 5 6 7 8 9
kDa
120
90
70
50
30
20
C
1 2 3 4 5 6 7 8 9
kDa
120
90
70
50
30
20

Figure 3.3. Mouse sera recognizing PSPC1. (A) Immunoblotting was performed using 40 μg of HeLa lysate to test whether antibodies to PSPC1 were raised in mice. Sera from mouse #1 and #3 exhibited a strong reactivity to PSPC1 (arrow) while additional proteins were also detected. Anti-PSPC1 antibody was used to confirm the migration rate of PSPC1. The membrane was stripped and reblotted using a β-Actin antibody. Beta-Actin serves as a loading control. **(B)** Immunofluorescence microscopy confirmed that antibodies generated in mouse were suitable for IF experiments using HeLa and C2C12 cells. Mouse #1 showed the best paraspeckle-like pattern while mouse #2 and #3 exhibited a weak reactivity. Scale bar, 10 μm.

A


Mouse #1
Mouse #2
Mouse #3
PSPC1
kDa
160
100
70
50
40
30
20
β-Actin

B

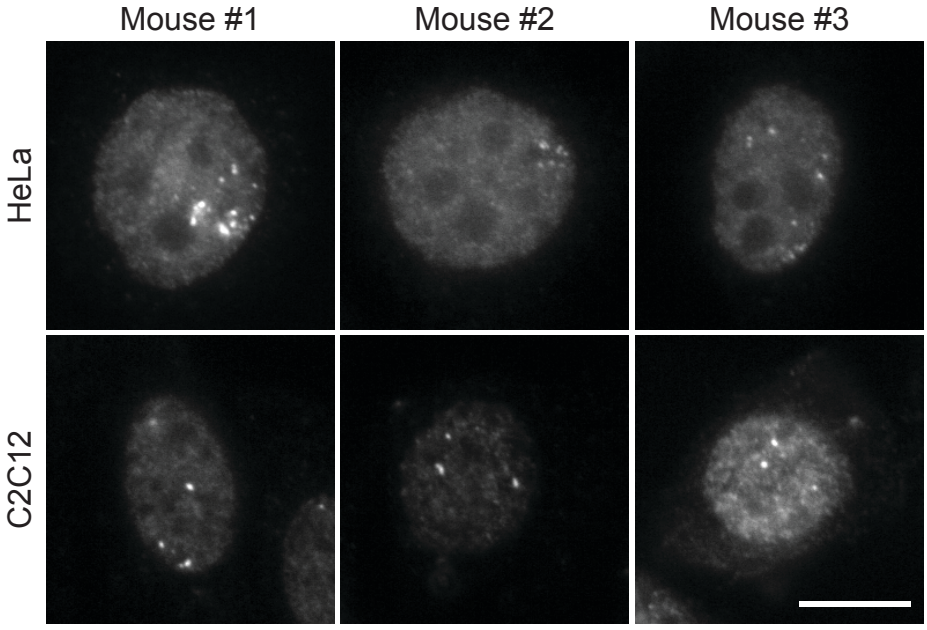

Mouse #1
Mouse #2
Mouse #3
HeLa
C2C12

Figure 3.4. Hybridomas were tested for their application in immunofluorescence microscopy. After screening ~2,000 hybridomas by ELISA, 96 hybridoma supernatants were utilized to perform IF using HeLa cells. Fifty one hybridomas (out of 96) could label a paraspeckle-like pattern in HeLa nuclei. Hybridoma 849 labeled a speckle-like pattern. Scale bar, 10 μm.

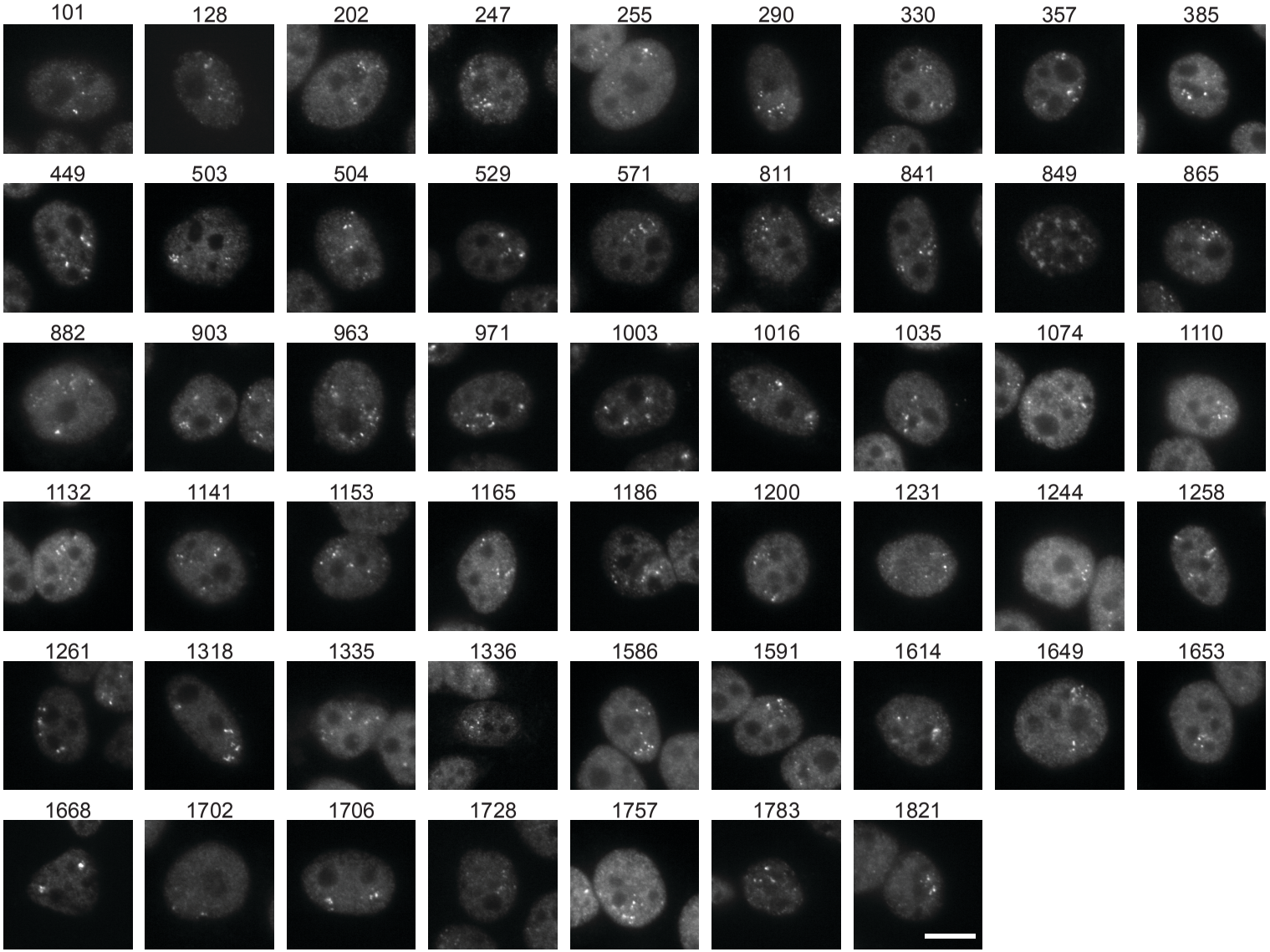

101
128
202
247
255
290
330
357
385
449
503
504
529
571
811
841
849
865
882
903
963
971
1003
1016
1035
1074
1110
1132
1141
1153
1165
1186
1200
1231
1244
1258
1261
1318
1335
1336
1586
1591
1614
1649
1653
1668
1702
1706
1728
1757
1783
1821

Figure 3.5. The final 15 Hybridomas can label paraspeckles in HeLa and/or C2C12 cell. Scale bar, 10 μm.

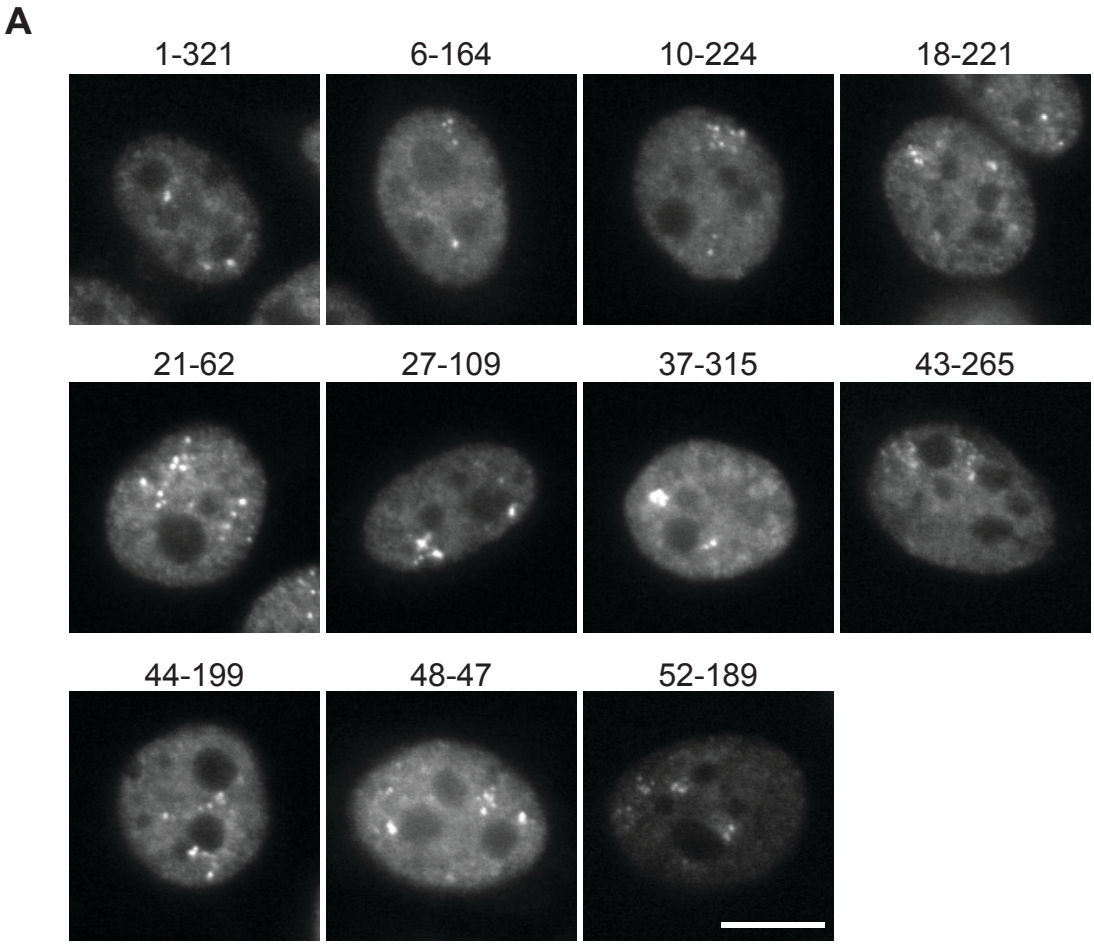

A
1-321
6-164
10-224
18-221
21-62
27-109
37-315
43-265
44-199
48-47
52-189
B
9-99
22-67
27-107
31-89
48-47

Figure 3.6. Majority of hybridomas exhibited specificities to PSPC1. (A) Immunoblotting using 11 different hybridomas was performed on HeLa lysates. Two PSPC1 antibodies (lanes 1 and 2, αPSPC1 and 1L4) served as a positive control. Most hybridomas specifically recognized PSPC1 protein (lanes 3 – 13). 1-321 detected SFPQ in addition to PSPC1 (lane 12, arrowhead). **(B)** Immunoblotting using 6 different hybridomas was performed on C2C12 lysates. Two PSPC1 antibodies (lanes 1 – 2), one Nono antibody (lane 9, NC5), and one Sfpq (lane 10, FC23) served as positive controls. Four hybridomas predominantly detected Pspc1 (lanes 2, 5, 6, and 8). 9-99 specifically recognized Nono instead of Pspc1 (lane 3). 1-321 detected other protein as well as Sfpq (lane 8, arrowhead).

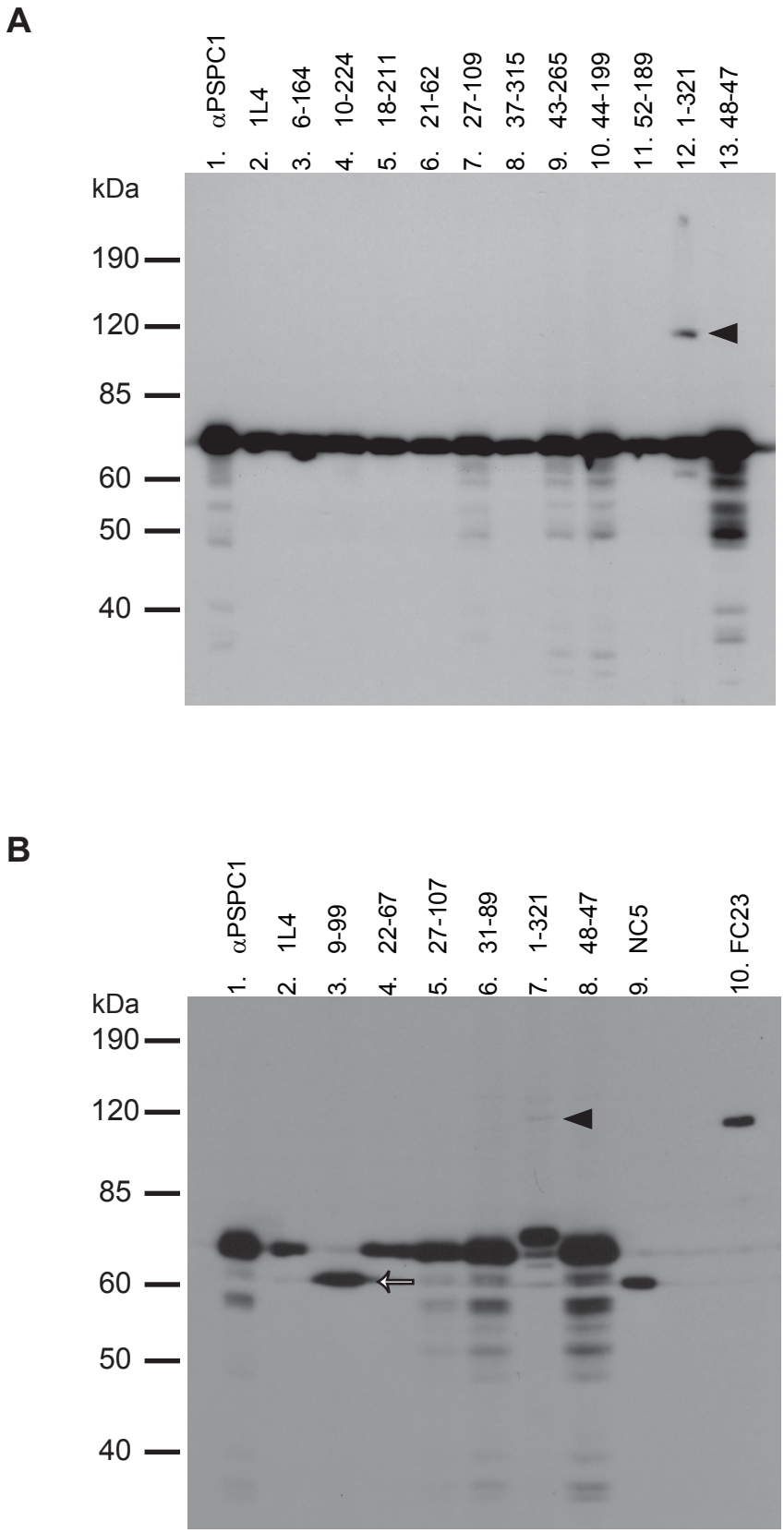

A
1. αPSPC1
2. 1L4
3. 6-164
4. 10-224
5. 18-211
6. 21-62
7. 27-109
8. 37-315
9. 43-265
10. 44-199
11. 52-189
12. 1-321
13. 48-47
kDa
190
120
85
60
50
40
B
1. αPSPC1
2. 1L4
3. 9-99
4. 22-67
5. 27-107
6. 31-89
7. 1-321
8. 48-47
9. NC5
10. FC23
kDa
190
120
85
60
50
40

Figure 3.7. The epitope recognized by each hybridoma was roughly mapped. (A) A series of PSPC1 deletion mutants were generated as GST fusion proteins. Numbers indicated positions of amino acids relative to full length PSPC1. RNA recognition motif 1 (RRM1) and NOPS were shown in the GST-full length PSPC1 construct. **(B)** SDS-PAGE analysis confirmed that GST-fusion deletion PSPC1 mutants were expressed in bacteria. Arrowheads depict series of PSPC1 proteins. **(C)** Immunoblotting was performed using bacterial lysate. Hybridomas can recognize either the N terminus or the C terminus of PSPC1 where the protein sequences are diverged from other proteins.

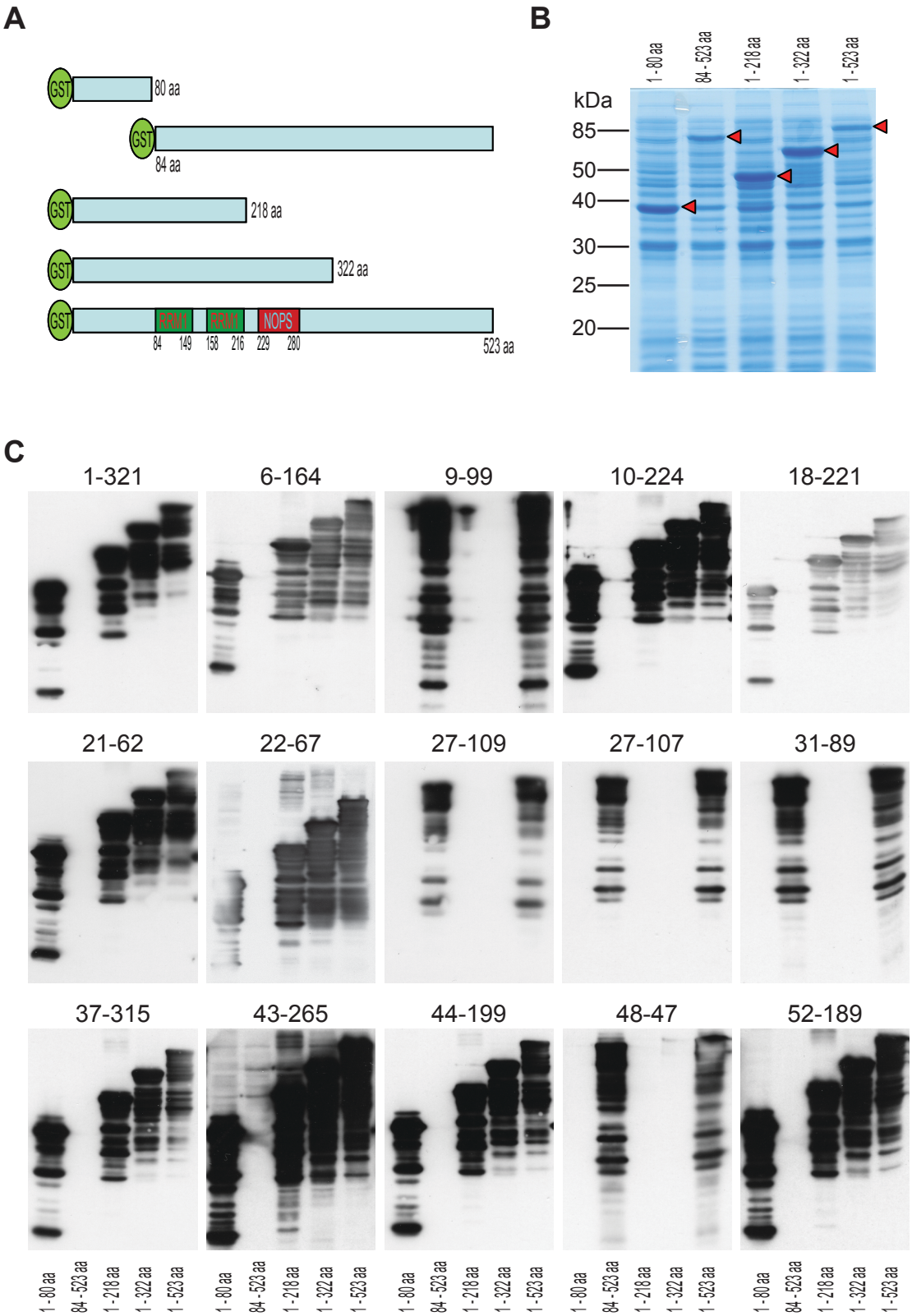
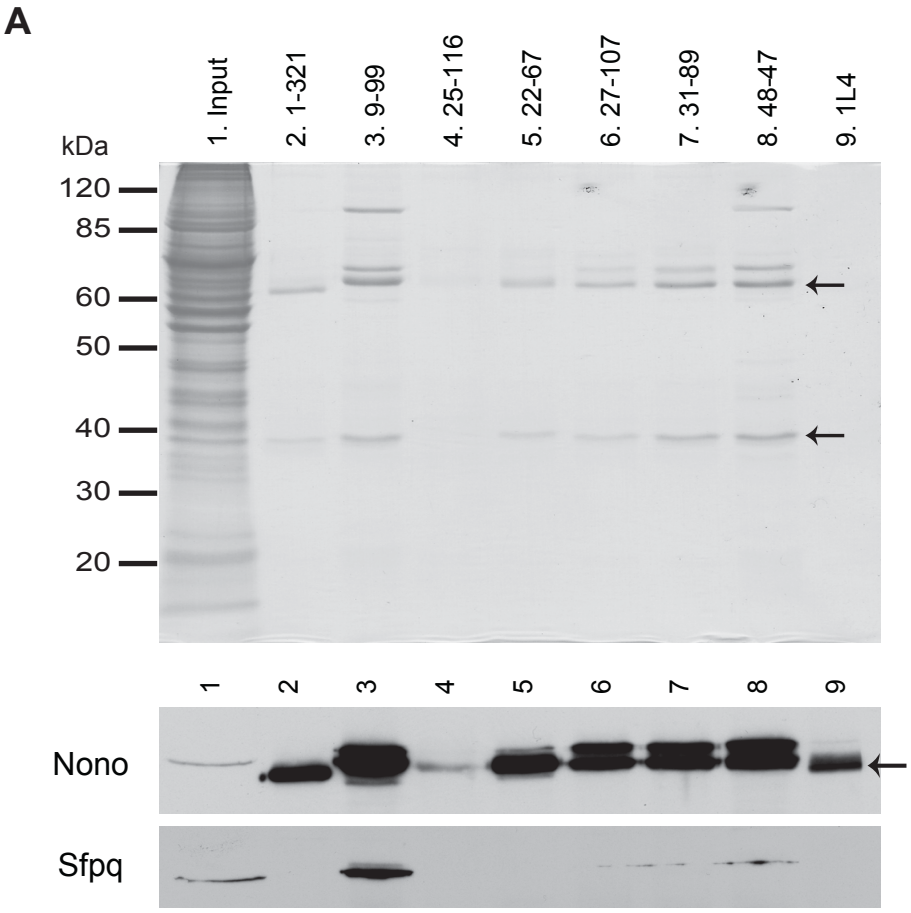
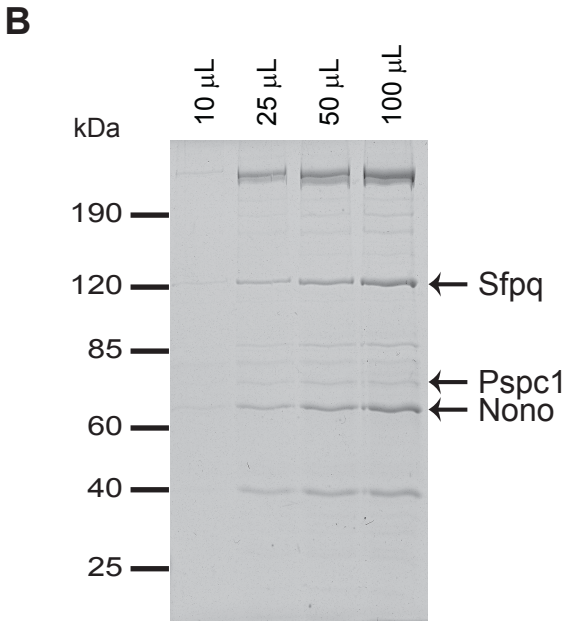

A
GST
80 aa
84 aa
218 aa
322 aa
RRM1
RRM2
NOPS
84
149
158
216
229
280
523 aa
B
1 - 80 aa
84 - 523 aa
1 - 218 aa
1 - 322 aa
1 - 523 aa
kDa
85
50
40
30
25
20
C
1-321
6-164
9-99
10-224
18-221
21-62
22-67
27-109
27-107
31-89
37-315
43-265
44-199
48-47
52-189
1 - 80 aa
84 - 523 aa
1 - 218 aa
1 - 322 aa
1 - 523 aa

Figure 3.8. Hybridomas were tested for their application in Co-IP. (A) Eight different hybridoma supernatants were tested for their abilities to pull down the Pspc1 or Nono complexes from C2C12 lysates. Four antibodies (9-99, 27-107, 31-89, and 48-47) were able to pull down the protein complexes. Immunoblotting using NC5 and FC23 confirmed that Nono and Sfpq were pulled down, respectively. Arrows depict IgG heavy chain and light chain. **(B)** Increasing amount of 9-99 conjugated Dynabead protein A could pull down increasing amounts of the Nono complex.


A
1. Input
2. 1-321
3. 9-99
4. 25-116
5. 22-67
6. 27-107
7. 31-89
8. 48-47
9. 1L4
kDa
120
85
60
50
40
30
20
1
2
3
4
5
6
7
8
9
Nono
Sfpq


B
10 µL
25 µL
50 µL
100 µL
kDa
190
120
85
60
40
25
Sfpq
Pspc1
Nono

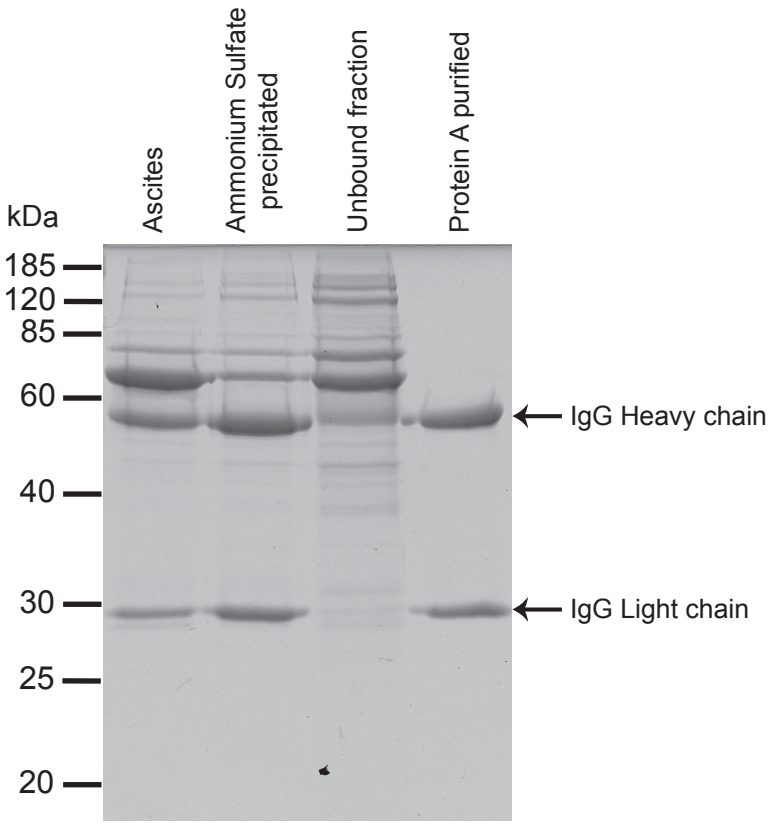


Figure 3.9. 9-99 antibodies were purified from ascites. First, ammonium sulfate precipitation was utilized to remove contaminants from mouse sera. 9-99 antibodies were further purified using sepharose protein A beads. Arrows depict IgG heavy chain and light chain

Hybridoma	Species (IF)	Epitope	Isotype
1-321	Human	1-80 aa	IgG1
6-164	Human	1-80 aa	IgG2b
10-224	Human	1-80 aa	IgG2b
18-221	Human	1-80 aa	IgG2b
21-62	Human	1-80 aa	IgG1
27-109	Human	323-523 aa	IgG2b
37-315	Human	1-80 aa	IgG1
43-265	Human	1-80 aa	IgG1
44-199	Human	1-80 aa	IgG1
52-189	Human	1-80 aa	IgG1
22-67	Mouse	1-80 aa	IgG1
27-107	Mouse	323-523 aa	IgG2b
31-89	Mouse	323-523 aa	IgG2b
9-99	Human and mouse	323-523 aa	IgG2a
48-47	Human and mouse	323-523 aa	IgG2b

Table 3.1. Summary of hybridomas. Epitope refers to amino acid position in PSPC1 that each hybridoma reacts with.

Conclusion and Perspective

Recent studies of mammalian transcriptomes revealed that pervasive transcription occurs throughout the genome producing not only mRNAs but also many ncRNAs (Bertone et al. 2004; Birney et al. 2007; Carninci et al. 2005; Cheng et al. 2005; Kapranov et al. 2007a). Although these non-coding transcripts were once considered as transcriptional noise, a hypothesis has been proposed that ncRNAs are functional in cells, providing another layer of complexity to organisms (for review, see Mattick 2004). In this study, a custom microarray identified 184 ncRNAs that exhibited more than 2-fold up- or down-regulation upon muscle differentiation. After RNA FISH was performed to study the localization of 14 ncRNA candidates, I focused on the *Men ε/β* locus. In this case, two ncRNA isoforms are produced from a single RNA polymerase II promoter, differing in the location of their 3' ends. *Men ε* is a 3.2-kb polyadenylated RNA, whereas *Men β* is a ~20-kb transcript containing a genomically encoded poly(A)-rich tract at its 3' end. The 3' end of *Men β* is generated by RNase P cleavage. The *Men ε/β* transcripts are localized to nuclear paraspeckles and directly interact with the Nono complex. I demonstrated that the knock-down of *MEN ε/β* expression results in the disruption of nuclear paraspeckles. Furthermore, the formation of paraspeckles, after release from transcriptional inhibition by DRB treatment, was suppressed in *MEN ε/β* depleted cells. These data indicate that the *MEN ε/β* ncRNAs are essential structural/organizational components of paraspeckles. Although *Xlsirts* ncRNA and *VegT* mRNA were shown to provide structural integrity to the cytoskeleton (Kloc et al. 2005), there has been no example of such ncRNAs for any nuclear structures. *MEN ε/β* ncRNAs provide the first example of ncRNAs that serve a critical role in nuclear organization.

In the process of characterizing of the *MEN ε/β* ncRNAs, I developed mouse monoclonal antibodies against PSPC1, a paraspeckle marker protein. PSPC1 was expressed as a GST-fusion protein in *E. coli* and purified. After immunizing three mice, one of them was utilized to establish hybridomas. After several rounds of screening, 15 hybridomas were characterized for their isotypes, epitopes, and applications including IF, IB, and Co-IP. Among them, antibody 9-99 that recognizes Nono, but not Pspc1 in C2C12 lysate, efficiently pulled down the Nono complex and was utilized extensively during this study. Other antibodies still remain for further characterization.

Two independent reports confirmed the structural role of the *MEN ε/β* ncRNAs in paraspeckles

Recently, two other groups reported very similar results supporting the conclusion of my study that the integrity of paraspeckles depends on the expression of *MEN ε/β* ncRNAs (Clemson et al. 2009; Sasaki et al. 2009). Although all three reports, including my own, agreed on the essential structural role of these ncRNAs in paraspeckles, there are some differences in the details of each study. Although Sasaki et al. observed the disruption of paraspeckles upon knock-down of *MEN ε/β* transcripts by ASOs, they also showed that knock-down of NONO or SFPQ by siRNAs resulted in disruption of paraspeckles as well as degradation of *MEN β*, but not *MEN ε* (Sasaki et al. 2009). Therefore, it was argued that the *MEN β*, but not *MEN ε*, is the structural RNA for paraspeckles based on evidence that only *MEN β* interacts with the NONO and SFPQ complex (Sasaki et al. 2009). Although it is clear that NONO and SFPQ are also essential components of paraspeckles, the loss of paraspeckles may be a secondary effect from the

degradation of the *MEN β* transcripts. One critical caveat was that ASOs in their experiments target the shared region of *MEN ε* and *MEN β*, resulting in knock-down of both isoforms at the same time. Therefore, their argument that only *MEN β* is essential for paraspeckles is not supported by direct evidence, but only by indirect evidence that NONO and SFPQ specifically maintain the stability of the *MEN β* transcripts. In contrast, I utilized a *MEN β*-specific ASO to deplete only *MEN β*, but not *MEN ε*. Although knock-down of *MEN β* expression caused most cells to lose paraspeckles, paraspeckles still remained colocalized only with the *MEN ε* transcripts in some population of *MEN β*-depleted cells (Figure 2.19). In addition, Clemson et al. showed that over-expression of *Men ε* can increase the number of paraspeckles per nuclei (Clemson et al. 2009). These data suggest that *MEN ε* is, at least partly, sufficient to maintain paraspeckles. However, Clemson et al. were negligent of the *MEN β* transcripts throughout their study. Considering all of these data, it is unlikely that *MEN β* is the only structural RNA in paraspeckles.

The second interesting point is to reveal which isoform among *MEN ε* and *MEN β* interacts with paraspeckle proteins. I demonstrated that the *Men ε* and *Men β* transcripts interact with the Nono complex in mouse cells (Figure 2.16). Clemson et al. showed that *MEN ε* interacts with PSPC1 *in vivo* and with PSPC1/NONO *in vitro* (Clemson et al. 2009). However, Sasaki et al. argued that *MEN β*, but not *MEN ε*, interacts with the NONO and SFPQ complex (Sasaki et al. 2009). These discrepancies may be due to simple variations in experimental material or conditions. Three different antibodies were used in the three separate Co-IP experiments. While mouse cells were utilized in my

study, HeLa cells were utilized in the two other studies. While Sasaki et al. utilized less stringent conditions (0.5% Triton X-100), Clemson et al. and I adopted more stringent conditions (RIPA buffer) during the Co-IP experiment. Another technical difficulty arises because the primers used in the Q-PCR analysis cannot distinguish *MEN ε* from *MEN β* since the *MEN β* transcript contains the entire *MEN ε* sequence. To address this problem, I performed Northern blot analysis to confirm that *Men ε* transcripts unambiguously exist in the Nono complex (Figure 2.16). The same approach will be essential to rule out any misinterpretation of results from HeLa cells.

The last discrepancy is in regards to the existence of mitotic paraspeckles. It was previously shown that foci of PSPC1 are maintained throughout the cell cycle although the number of foci decreases upon entry into mitosis (Fox et al. 2002). Clemson et al. argued that mitotic accumulation of PSPC1 and NONO are not paraspeckles since the *MEN ε* transcripts were not detected in such foci of PSPC1 and NONO (Clemson et al. 2009). However, the *Men ε/β* transcripts exhibit punctate distribution throughout mitosis (Figure 2.3). In addition, I observed that the *MEN ε/β* transcripts were colocalized with PSPC1/NONO foci during M phase (Figure 2.17). These data indicate that the observed mitotic accumulation of PSPC1 and NONO are indeed paraspeckles. Future studies will address these issues and confirm the importance of each RNA isoform.

Paraspeckles and nuclear organization

The mammalian nucleus is a highly compartmentalized organelle harboring various specialized nuclear domains such as chromosome territories, interchromatin granules (also known as speckles), nucleoli, Cajal bodies, PML bodies, and paraspeckles,

to name a few (for review, see Spector 2001; Spector 2006). To understand the functions of these nuclear domains, several studies have tried to systematically identify protein components of a few nuclear domains including interchromatin granules (Mintz et al. 1999; Saitoh et al. 2004), nucleoli (Andersen et al. 2002; Scherl et al. 2002), and Cajal bodies (Lam et al. 2002). Since paraspeckles were identified less than a decade ago (Fox et al. 2002), only a few proteins and RNAs are thus far known to localize to paraspeckles; PSPC1 (Fox et al. 2002), NONO (Fox et al. 2002), SFPQ (Fox et al. 2005), CPSF6 (also known as Cleavage Factor I_m68, Cardinale et al. 2007), RBM14 (also known as CoAA/PSP2, Fox et al. 2002), BCL6 (Liu et al. 2006), BCL11A (Liu et al. 2006), *CTN-RNA* (Prasanth et al. 2005), and *MEN* ε/β ncRNAs (Clemson et al. 2009; Sasaki et al. 2009; Sunwoo et al. 2009). However, a successful biochemical purification of paraspeckles has not yet been reported. Since only limited information about paraspeckle composition is available, the function of paraspeckles still remains elusive. There is only one report that proposed paraspeckles as a RNA storage depot; *CTN-RNA* is transiently stored at paraspeckles and leaves paraspeckles when a stress signal induces a cleavage of its long 3' UTR (Prasanth et al. 2005). In fact, paraspeckles are RNA-rich nuclear bodies (Prasanth et al. 2005). All known paraspeckle proteins except BCL6 and BCL11A harbor one or two RNA recognition motifs (RRMs). Interestingly, these RRMs are required for paraspeckle localization of PSPC1 (Fox et al. 2005) and CSPF-6 (Cardinale et al. 2007), suggesting that interacting RNAs can be recruited to and potentially stored in paraspeckles. Considering these data, the proteomic composition of paraspeckles may not be sufficient to provide us with a complete picture of these nuclear domains. It is also necessary to obtain information with regard to the RNA composition of these nuclear

bodies.

The localization study of PSPC1 led to the identification of paraspeckles (Fox et al. 2002). Since then, PSPC1 has been considered a paraspeckle marker (Fox et al. 2005; Fox et al. 2002). However, the localization of endogenous PSPC1 is not limited to paraspeckles, but also to the nucleoplasm (Figure 3.5). Strikingly, even when PSPC1 was depleted by siRNAs, cells still maintain paraspeckle-like structures where NONO, SFPQ, and the *MEN ε/β* transcripts co-localize (Sasaki et al. 2009). In contrast, knock-down of NONO or SFPQ results in disruption of paraspeckles, suggesting NONO and SFPQ are also essential components of paraspeckles, in addition to *MEN ε/β* ncRNAs (Sasaki et al. 2009). On the other hand, *MEN β* ncRNA is exclusively localized to paraspeckles while *MEN ε* exhibits additional speckle localization depending on the cell lines examined (Figures 2.2, 2.14, and 2.15). *CTN-RNA* is localized to paraspeckles only in a subset of cells (Prasanth et al. 2005). Knock-down of *MEN ε/β* expression abolishes the integrity of paraspeckles without altering the protein level of PSPC1, NONO, and SFPQ (Figure 2.18; Sasaki et al. 2009). Over-expression of *Men ε* induces an increase in number of paraspeckles (Clemson et al. 2009). Together, these data indicate that *MEN ε/β* ncRNAs are sufficient and necessary components of paraspeckles. I propose that the *MEN ε/β* ncRNAs are the best markers of paraspeckles. When RNA FISH cannot be performed, NONO and SFPQ can serve as alternative markers to label paraspeckles.

Biogenesis of nuclear bodies

Nuclear domains have unique structures since their periphery is not defined by a lipid membrane like cytoplasmic organelles (for review, see Spector 2001). One

fundamental question is: how are nuclear domains organized in cells? A recent study suggests that nuclear bodies are self-organizing structures through protein-protein interactions (Kaiser et al. 2008). Each of the Cajal body components was independently tethered to a DNA locus through DNA-protein or DNA-protein-RNA interaction in the nucleus to determine whether a Cajal body could be established *de novo* at that locus (Kaiser et al. 2008). Many protein components as well as small Cajal body associated RNA were shown to recruit other Cajal body components, including proteins as well as RNAs, resulting in the establishment of a Cajal body (Kaiser et al. 2008).

The studies discussed above suggest that the establishment of paraspeckles may depend on RNA-protein interactions. First, paraspeckles are localized nearby the *MEN ε/β* locus in nuclei (Clemson et al. 2009). Second, paraspeckles are sensitive to the activity of RNA polymerase II (Figures 2.20 – 21). Third, depletion of *MEN ε/β* ncRNAs results in disruption of paraspeckles (Clemson et al. 2009; Sasaki et al. 2009; Sunwoo et al. 2009). Lastly, over-expression of *Men ε* induce more paraspeckles per nucleus, suggesting that *MEN ε/β* ncRNAs are limiting factors in the cell to determine the number of paraspeckles (Clemson et al. 2009). Considering these data, it is reasonable to propose the following model. When *MEN ε/β* ncRNAs are transcribed, paraspeckle proteins including PSPC1, NONO, and SFPQ are recruited to the nascent *MEN ε/β* transcripts, thereby assembling into RNPs. These RNPs are clustered into a bigger structure, developing into paraspeckles. This model of paraspeckle biogenesis at the *Men ε/β* locus may also explain how *MEN ε/β* ncRNAs are retained in the nucleus without A-to-I editing. RNA A-to-I hyperediting within inverted repeats has been suggested as a nuclear retention mechanism of RNA through binding to NONO (Chen et al. 2008; Kumar and

Carmichael 1997; Prasanth et al. 2005; Zhang and Carmichael 2001). Interestingly, Nono can bind to the *Men ε/β* transcripts that are not subject to RNA editing, suggesting that Nono may be generally involved in nuclear retention of RNA, not limited to that of hyperedited RNAs.

The model suggested above provides a great advantage to visualize the process of "*de novo*" paraspeckle biogenesis. Several years ago, Janicki et al. developed a gene expression system to elegantly visualize DNA, RNA, and proteins at the same time in a live cell (Janicki et al. 2004). The system consists of multiple components; Lac operator to mark DNA locus by LacI, Tet-inducible promoter to control expression of CFP-SKL, MS2 coat protein to trace RNA through binding to MS2 repeats in the RNA transcript (Janicki et al. 2004). A few modifications of this approach may allow live cell imaging of paraspeckle biogenesis; replacement of CFP-SKL with the *Men ε/β* sequences and placing MS2 repeats at the 5' end of the *Men ε/β* sequences. The integrated locus can be visualized by LacI fused with fluorescent protein. Upon addition of doxycycline, rtTA protein induces the expression of *Men ε/β* transcript, thereby recruiting paraspeckle proteins fused with the second color of fluorescent protein. The assembly of these RNPs would then develop into paraspeckles. At the same time, the MS2 coat protein fused with a third color of fluorescent protein would allow one to visualize the *Men ε/β* transcripts containing MS2 repeats. Using this system, one can monitor the biogenesis of paraspeckles and the kinetics of recruitment of paraspeckle proteins.

The previous study examining the biogenesis of Cajal bodies claimed that each Cajal body component can seed *de novo* formation of a Cajal body when its concentration reaches a certain level by tethering to the locus (Kaiser et al. 2008). However, this study

did not exclude the possibility that these proteins formed a seed complex in the cell prior to their association with the DNA locus. Therefore, tethering of a protein could have recruited a seeding complex of protein and/or RNA rather than a protein. In contrast, one can monitor the real-time paraspeckle biogenesis by directly modulating the expression of *Men ε/β* ncRNAs. In addition, tethering of paraspeckle proteins will address whether paraspeckles are another self-organizing structure like Cajal body or assemble according to a determined order. If this system is developed, paraspeckles can serve as a model system of how nuclear domains are established and help the field of nuclear organization to advance to the next level.

Future directions

My study has revealed the structural role of the *MEN ε/β* ncRNAs in paraspeckles. Interestingly, I observed that *Men ε/β* ncRNAs are up-regulated upon muscle differentiation in C2C12 cells (Figure 2.4) and expressed only in certain types of cells in mouse tissues (figure 2.5). Similarly, the bovine ortholog of *Men ε* ncRNA was also shown to be 6.8-fold up-regulated during the late stages of muscle development (Lehnert et al. 2007). Considering the sufficient and necessary role of *MEN ε/β* ncRNAs in paraspeckles, some cells may require more or less paraspeckles. Although a question of why cells need more or less *Men ε/β* ncRNAs or paraspeckles is fundamentally important to understand the function of paraspeckles, it was not pursued during my study because ASOs failed to deplete the expression of *Men ε/β* in C2C12 cells (data not shown).

In addition to ASOs, siRNAs have also been utilized to knock-down the expression of target genes (for review, see Crooke 2004). Since it has been generally

accepted that the effector machinery of the RNAi pathway functions only in the cytoplasm (Zeng and Cullen 2002), the siRNA approach has been not considered suitable to knock-down nuclear retained RNAs. A few years ago, a report showed that a functional RNAi machinery exists in the nucleus (Robb et al. 2005). More recently, siRNAs were shown to knock-down the nuclear retained *MEN ε* transcripts in HeLa cells (Clemson et al. 2009). Although it is still unclear how siRNAs can knock-down nuclear retained *MEN ε/β* ncRNAs, this approach may allow one to modulate the expression *Men ε/β* expression in C2C12 cells, thereby addressing whether these ncRNAs are required for myoblast differentiation into myotubes. Ultimately, a knock-out mouse model can address the importance of *Men ε/β* ncRNAs at the organism level and also presumably that of paraspeckles.

I have successfully developed multiple monoclonal antibodies to PSPC1 (Chapter 3). One can utilize these antibodies in immunoprecipitation experiments to pull down paraspeckles along with the PSPC1 complex and identify additional RNA species in paraspeckles or the PSPC1 complex by high thruput deep sequencing technology or microarray analysis. One caveat here is that PSPC1 localizes to paraspeckle as well as to the nucleoplasm, indicating inevitable contamination of the soluble PSPC1 complexes from the nucleoplasm. However, until a procedure to biochemically purify paraspeckles is developed, the immunoaffinity purification may allow a very quick and efficient way to identify paraspeckle-localizing RNAs and proteins as a similar approach confirmed the existence of *Men ε/β* in the Nono complex (Figure 2.16). Characterization of such RNAs may validate or disprove the suggested function of paraspeckles as a RNA storage depot.

In the course of studying *Men ε/β,* it was found that the 3’ end of *Men β* is

generated by RNase P (Figure 2.11), unlike most RNA polymerase II transcripts that are subject to classical cleavage/polyadenylation (for review, see Danckwardt et al. 2008). Recently, the 3' end of another ncRNA *MALAT1* was also shown to be processed by the same mechanism (Wilusz et al. 2008). Although there are only two examples of such RNA polymerase II transcripts that are subject to tRNA processing machinery, it has not been asked whether this 3' end processing occurs more generally to other transcripts. The secondary structure of immediate downstream sequences is predicted to fold into tRNA-like structure (Figure 2.13; Wilusz et al. 2008), presumably providing signals for the tRNA processing machinery. Although a search for tRNA-like secondary structure in the genome may identify others, the prediction algorithms create a lot of false positive and fail to identify other candidates considering that a single RNA sequence may be predicted to have very different structures. On the other hand, the tRNA processing machinery adds CCA trinucleotides to the 3' end of tRNA and tRNA-like small RNAs. Searching for this signature in the databases of small RNA libraries by deep sequencing may be a straight forward strategy to identify other tRNA-like small RNAs. We have just begun to appreciate such novel 3' end processing mechanisms. Further identification of additional small RNAs in this class will provide a means to start functional characterization of these interesting RNAs.

Perspective

It is undeniable that long ncRNAs can serve various functional roles in cells. Their paradigms are not only limited to the well known function of ncRNAs as precursors to small RNAs, but also extended to roles that were once considered to belong only to

proteins, including regulating transcriptional activity, modulating protein activities, serving structural or organizational roles, and altering RNA processing events (Figure 1.1). With an increasing number of ncRNAs being identified from genome wide studies (Birney et al. 2007; Guttman et al. 2009), thus far, only a few ncRNAs have been characterized, mainly in a cellular context (for review, see Mercer et al. 2009; Prasanth and Spector 2007).

Interestingly, many ncRNAs seem to function around their loci, for example *Airn*, *Kcnq1ot1*, $ncRNA_{CCND1}$, and so on, regulating expression of nearby protein-coding genes. Insertion of a polyadenylation signal that can result in premature termination of transcription has been utilized to distinguish whether transcription or the ncRNA product plays a key role in regulation (Hirota et al. 2008; Sleutels et al. 2002; Uhler et al. 2007). This approach has frequently led to a conclusion that transcription itself is enough for regulation (Hirota et al. 2008; Uhler et al. 2007). However, a premature conclusion might have been drawn since the length of transcription may be important as in the cases of *Kcnq1ot1* or *DHFR* ncRNAs (Kanduri et al. 2006; Martianov et al. 2007). Since an ectopic expression of cis-acting RNAs cannot recapitulate their activity, there is still no definite way to discriminate whether RNA products or the act of transcription is functional. Provided that RNA product is an inevitable result of transcription, this might be just another chicken and egg question.

The *MEN* ε/β transcripts are one of a few trans-acting ncRNAs. They are highly concentrated in paraspeckles (Figure 2.15) and essential structural components of paraspeckles (Figure 2.l8). PSPC1 directly binds to the *MEN* ε sequences (Clemson et al. 2009) and exhibits a very slow exchange (half-life of 31s) between paraspeckles and the

nucleoplasm (Fox et al. 2005). Although these data suggest that these ncRNAs function mainly in paraspeckles as a structural platform, it cannot be ruled out that the *MEN ε/β* RNP complex may also function in the nucleoplasm. As expected from a proteomic study of nucleoli, PSPC1 traffics between the nucleolus and paraspeckles (Fox et al. 2002). Provided that RNase P and other members of tRNA processing machinery are primarily localized to the nucleolus (Bertrand et al. 1998), the *MEN β* transcript may transiently be localized to nucleolus to have its 3' end processed and return to paraspeckles. However, it is equally possible that the processing machinery may leave the nucleolus and be recruited to the *MEN ε/β* transcription site for processing.

Several nuclear structures such as PML bodies (Muratani et al. 2002) and Cajal bodies (Platani et al. 2002) exhibit mobility in the interphase nucleus. In contrast, paraspeckles are rather static nuclear bodies (Fox et al. 2005) that localize close to the *MEN ε/β* locus (Clemson et al. 2009). These data raise an interesting question: how paraspeckles can keep their relative positions in the nucleus? Is there an anchoring point like a certain DNA locus or nuclear matrix? Another interesting feature of paraspeckles is that they persist throughout the cell cycle unlike other nuclear domains including nucleoli (Leung et al. 2004) and Cajal bodies (Carmo-Fonseca et al. 1993) that dissociate upon entry to mitosis. Paraspeckles are sensitive to RNA polymerase II activity (Fox et al. 2002) and the expression of *MEN ε/β* ncRNAs (Figure 2.18). However, their association in mitosis may represent a means of segregating their components to daughter cells.

One suggested role for paraspeckles is that nuclear retained RNAs are stored in them (Prasanth et al. 2005). Interestingly, knock-down of *MEN ε/β* expression did not alter the level of paraspeckle proteins but only disrupted the integrity of paraspeckles

(Figure 2.18; Sasaki et al. 2009). If paraspeckle are essential for nuclear retention of certain transcripts such as *CTN-RNA*, knock-down of *MEN ε/β* will result in export of *CTN-RNA* into the cytoplasm. Future studies will address the questions asked here and allow us to understand how the mammalian nucleus is organized and how its organization regulates its cellular functions.

Finally, we are just beginning to see emerging evidence that ncRNAs provides complex organisms with various levels of regulation from gene expression, modulating protein activity, to serving a structural role. There are still an overwhelming number of ncRNAs that need to be studied in terms of their biology. In the near future, I expect that the field of long ncRNA will explode and explain why complex organisms maintain pervasive transcription producing ncRNAs.

List of References

Allen, T.A., S. Von Kaenel, J.A. Goodrich, and J.F. Kugel. 2004. The SINE-encoded mouse B2 RNA represses mRNA transcription in response to heat shock. *Nat Struct Mol Biol* **11:** 816-821.

Andersen, J.S., C.E. Lyon, A.H. Fox, A.K. Leung, Y.W. Lam, H. Steen, M. Mann, and A.I. Lamond. 2002. Directed proteomic analysis of the human nucleolus. *Curr Biol* **12:** 1-11.

Aravin, A.A., G.J. Hannon, and J. Brennecke. 2007. The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. *Science* **318:** 761-764.

Aravin, A.A., M. Lagos-Quintana, A. Yalcin, M. Zavolan, D. Marks, B. Snyder, T. Gaasterland, J. Meyer, and T. Tuschl. 2003. The small RNA profile during Drosophila melanogaster development. *Dev Cell* **5:** 337-350.

Bae, E., V.C. Calhoun, M. Levine, E.B. Lewis, and R.A. Drewell. 2002. Characterization of the intergenic RNA profile at abdominal-A and Abdominal-B in the Drosophila bithorax complex. *Proc Natl Acad Sci U S A* **99:** 16847-16852.

Beltran, M., I. Puig, C. Pena, J.M. Garcia, A.B. Alvarez, R. Pena, F. Bonilla, and A.G. de Herreros. 2008. A natural antisense transcript regulates Zeb2/Sip1 gene expression during Snail1-induced epithelial-mesenchymal transition. *Genes Dev* **22:** 756-769.

Berezney, R. and D.S. Coffey. 1974. Identification of a nuclear protein matrix. *Biochem Biophys Res Commun* **60:** 1410-1417.

Berretta, J., M. Pinskaya, and A. Morillon. 2008. A cryptic unstable transcript mediates transcriptional trans-silencing of the Ty1 retrotransposon in S. cerevisiae. *Genes Dev* **22:** 615-626.

Bertone, P., V. Stolc, T.E. Royce, J.S. Rozowsky, A.E. Urban, X. Zhu, J.L. Rinn, W. Tongprasit, M. Samanta, S. Weissman et al. 2004. Global identification of human transcribed sequences with genome tiling arrays. *Science* **306:** 2242-2246.

Bertrand, E., F. Houser-Scott, A. Kendall, R.H. Singer, and D.R. Engelke. 1998. Nucleolar localization of early tRNA processing. *Genes Dev* **12:** 2463-2468.

Birney, E. J.A. Stamatoyannopoulos A. Dutta R. Guigo T.R. Gingeras E.H. Margulies Z. Weng M. Snyder E.T. Dermitzakis R.E. Thurman et al. 2007. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* **447:** 799-816.

Blackshaw, S., S. Harpavat, J. Trimarchi, L. Cai, H. Huang, W.P. Kuo, G. Weber, K. Lee, R.E. Fraioli, S.H. Cho et al. 2004. Genomic analysis of mouse retinal development. *PLoS Biol* **2:** E247.

Blower, M.D., M. Nachury, R. Heald, and K. Weis. 2005. A Rae1-containing ribonucleoprotein complex is required for mitotic spindle assembly. *Cell* **121:** 223-234.

Blume, S.W., Z. Meng, K. Shrestha, R.C. Snyder, and P.D. Emanuel. 2003. The 5'-untranslated RNA of the human dhfr minor transcript alters transcription pre-initiation complex assembly at the major (core) promoter. *J Cell Biochem* **88:** 165-180.

Brawerman, G. and J. Diez. 1975. Metabolism of the polyadenylate sequence of nuclear RNA and messenger RNA in mammalian cells. *Cell* **5:** 271-280.

Bussemakers, M.J., A. van Bokhoven, G.W. Verhaegh, F.P. Smit, H.F. Karthaus, J.A. Schalken, F.M. Debruyne, N. Ru, and W.B. Isaacs. 1999. DD3: a new prostate-specific gene, highly overexpressed in prostate cancer. *Cancer Res* **59:** 5975-5979.

Cai, X., C.H. Hagedorn, and B.R. Cullen. 2004. Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *Rna* **10:** 1957-1966.

Cardinale, S., B. Cisterna, P. Bonetti, C. Aringhieri, M. Biggiogera, and S.M. Barabino. 2007. Subnuclear localization and dynamics of the Pre-mRNA 3' end processing factor mammalian cleavage factor I 68-kDa subunit. *Mol Biol Cell* **18:** 1282-1292.

Carmo-Fonseca, M., J. Ferreira, and A.I. Lamond. 1993. Assembly of snRNP-containing coiled bodies is regulated in interphase and mitosis--evidence that the coiled body is a kinetic nuclear structure. *J Cell Biol* **120:** 841-852.

Carninci, P. T. Kasukawa S. Katayama J. Gough M.C. Frith N. Maeda R. Oyama T. Ravasi B. Lenhard C. Wells et al. 2005. The transcriptional landscape of the

mammalian genome. *Science* **309:** 1559-1563.

Carrozza, M.J., B. Li, L. Florens, T. Suganuma, S.K. Swanson, K.K. Lee, W.J. Shia, S. Anderson, J. Yates, M.P. Washburn et al. 2005. Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* **123:** 581-592.

Chen, L.L., J.N. DeCerbo, and G.G. Carmichael. 2008. Alu element-mediated gene silencing. *Embo J* **27:** 1694-1705.

Cheng, J., P. Kapranov, J. Drenkow, S. Dike, S. Brubaker, S. Patel, J. Long, D. Stern, H. Tammana, G. Helt et al. 2005. Transcriptional maps of 10 human chromosomes at 5-nucleotide resolution. *Science* **308:** 1149-1154.

Clemson, C.M., J.N. Hutchinson, S.A. Sara, A.W. Ensminger, A.H. Fox, A. Chess, and J.B. Lawrence. 2009. An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. *Mol Cell* **33:** 717-726.

Clemson, C.M., J.A. McNeil, H.F. Willard, and J.B. Lawrence. 1996. XIST RNA paints the inactive X chromosome at interphase: evidence for a novel RNA involved in nuclear/chromosome structure. *J Cell Biol* **132:** 259-275.

Core, L.J., J.J. Waterfall, and J.T. Lis. 2008. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* **322:** 1845-1848.

Costa, F.F. 2005. Non-coding RNAs: new players in eukaryotic biology. *Gene* **357:** 83-94.

Crooke, S.T. 2004. Progress in antisense technology. *Annu Rev Med* **55:** 61-95.

Czech, B., C.D. Malone, R. Zhou, A. Stark, C. Schlingeheyde, M. Dus, N. Perrimon, M. Kellis, J.A. Wohlschlegel, R. Sachidanandam et al. 2008. An endogenous small interfering RNA pathway in Drosophila. *Nature* **453:** 798-802.

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

Davis, C.A. and M. Ares, Jr. 2006. Accumulation of unstable promoter-associated transcripts upon loss of the nuclear exosome subunit Rrp6p in Saccharomyces

cerevisiae. *Proc Natl Acad Sci U S A* **103:** 3262-3267.

de Kok, J.B., G.W. Verhaegh, R.W. Roelofs, D. Hessels, L.A. Kiemeney, T.W. Aalders, D.W. Swinkels, and J.A. Schalken. 2002. DD3(PCA3), a very sensitive and specific marker to detect prostate tumors. *Cancer Res* **62:** 2695-2698.

Dinger, M.E., P.P. Amaral, T.R. Mercer, K.C. Pang, S.J. Bruce, B.B. Gardiner, M.E. Askarian-Amiri, K. Ru, G. Solda, C. Simons et al. 2008. Long noncoding RNAs in mouse embryonic stem cell pluripotency and differentiation. *Genome Res* **18:** 1433-1445.

Dyck, J.A., G.G. Maul, W.H. Miller, Jr., J.D. Chen, A. Kakizuka, and R.M. Evans. 1994. A novel macromolecular structure is a target of the promyelocyte-retinoic acid receptor oncoprotein. *Cell* **76:** 333-343.

Ebert, M.S., J.R. Neilson, and P.A. Sharp. 2007. MicroRNA sponges: competitive inhibitors of small RNAs in mammalian cells. *Nat Methods* **4:** 721-726.

Ebisuya, M., T. Yamamoto, M. Nakajima, and E. Nishida. 2008. Ripples from neighbouring transcription. *Nat Cell Biol* **10:** 1106-1113.

Erwin, J.A. and J.T. Lee. 2008. New twists in X-chromosome inactivation. *Current Opinion in Cell Biology* **20:** 349-355.

Espinoza, C.A., T.A. Allen, A.R. Hieb, J.F. Kugel, and J.A. Goodrich. 2004. B2 RNA binds directly to RNA polymerase II to repress transcript synthesis. *Nat Struct Mol Biol* **11:** 822-829.

Faghihi, M.A., F. Modarresi, A.M. Khalil, D.E. Wood, B.G. Sahagan, T.E. Morgan, C.E. Finch, G. St Laurent, 3rd, P.J. Kenny, and C. Wahlestedt. 2008. Expression of a noncoding RNA is elevated in Alzheimer's disease and drives rapid feed-forward regulation of beta-secretase. *Nat Med* **14:** 723-730.

Fejes-Toth, K., V. Sotirova, R. Sachidanandam, G. Assaf, G.J. Hannon, P. Kapranov, S. Foissac, A.T. Willingham, R. Duttagupta, E. Dumais et al. 2009. Post-transcriptional processing generates a diversity of 5'-modified long and short RNAs. *Nature* **457:** 1028-1032.

Feng, J., C. Bi, B.S. Clark, R. Mady, P. Shah, and J.D. Kohtz. 2006. The Evf-2 noncoding

RNA is transcribed from the Dlx-5/6 ultraconserved region and functions as a Dlx-2 transcriptional coactivator. *Genes Dev* **20:** 1470-1484.

Fey, E.G., G. Krochmalnic, and S. Penman. 1986. The nonchromatin substructures of the nucleus: the ribonucleoprotein (RNP)-containing and RNP-depleted matrices analyzed by sequential fractionation and resinless section electron microscopy. *J Cell Biol* **102:** 1654-1665.

Fitzpatrick, G.V., P.D. Soloway, and M.J. Higgins. 2002. Regional loss of imprinting and growth deficiency in mice with a targeted deletion of KvDMR1. *Nat Genet* **32:** 426-431.

Fox, A.H., C.S. Bond, and A.I. Lamond. 2005. P54nrb forms a heterodimer with PSP1 that localizes to paraspeckles in an RNA-dependent manner. *Mol Biol Cell* **16:** 5304-5315.

Fox, A.H., Y.W. Lam, A.K. Leung, C.E. Lyon, J. Andersen, M. Mann, and A.I. Lamond. 2002. Paraspeckles: a novel nuclear domain. *Curr Biol* **12:** 13-25.

Galindo, M.I., J.I. Pueyo, S. Fouix, S.A. Bishop, and J.P. Couso. 2007. Peptides encoded by short ORFs control development and define a new eukaryotic gene family. *PLoS Biol* **5:** e106.

Ganesan, G. and S.M. Rao. 2008. A novel noncoding RNA processed by Drosha is restricted to nucleus in mouse. *Rna* **14:** 1399-1410.

Geirsson, A., R.J. Lynch, I. Paliwal, A.L. Bothwell, and G.L. Hammond. 2003a. Human trophoblast noncoding RNA suppresses CIITA promoter III activity in murine B-lymphocytes. *Biochem Biophys Res Commun* **301:** 718-724.

Geirsson, A., I. Paliwal, R.J. Lynch, A.L. Bothwell, and G.L. Hammond. 2003b. Class II transactivator promoter activity is suppressed through regulation by a trophoblast noncoding RNA. *Transplantation* **76:** 387-394.

Ghildiyal, M., H. Seitz, M.D. Horwich, C. Li, T. Du, S. Lee, J. Xu, E.L. Kittler, M.L. Zapp, Z. Weng et al. 2008. Endogenous siRNAs derived from transposons and mRNAs in Drosophila somatic cells. *Science* **320:** 1077-1081.

Guru, S.C., S.K. Agarwal, P. Manickam, S.E. Olufemi, J.S. Crabtree, J.M. Weisemann,

M.B. Kester, Y.S. Kim, Y. Wang, M.R. Emmert-Buck et al. 1997. A transcript map for the 2.8-Mb region containing the multiple endocrine neoplasia type 1 locus. *Genome Res* **7:** 725-735.

Guttman, M., I. Amit, M. Garber, C. French, M.F. Lin, D. Feldser, M. Huarte, O. Zuk, B.W. Carey, J.P. Cassady et al. 2009. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **458:** 223-227.

Harlow, E. and D. Lane. 1988. *Antibodies : a laboratory manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

He, Y., B. Vogelstein, V.E. Velculescu, N. Papadopoulos, and K.W. Kinzler. 2008. The antisense transcriptomes of human cells. *Science* **322:** 1855-1857.

Heard, E. and C.M. Disteche. 2006. Dosage compensation in mammals: fine-tuning the expression of the X chromosome. *Genes Dev* **20:** 1848-1867.

Heasman, J., O. Wessely, R. Langland, E.J. Craig, and D.S. Kessler. 2001. Vegetal localization of maternal mRNAs is disrupted by VegT depletion. *Dev Biol* **240:** 377-386.

Hellwig, S. and B.L. Bass. 2008. A starvation-induced noncoding RNA modulates expression of Dicer-regulated genes. *Proc Natl Acad Sci U S A* **105:** 12897-12902.

Herman, R.C., J.G. Williams, and S. Penman. 1976. Message and non-message sequences adjacent to poly(A) in steady state heterogeneous nuclear RNA of HeLa cells. *Cell* **7:** 429-437.

Hirota, K., T. Miyoshi, K. Kugou, C.S. Hoffman, T. Shibata, and K. Ohta. 2008. Stepwise chromatin remodelling by a cascade of transcription initiation of non-coding RNAs. *Nature* **456:** 130-134.

Hirota, K. and K. Ohta. 2009. Cascade transcription of mRNA-type long non-coding RNAs (mlonRNAs) and local chromatin remodeling. *Epigenetics* **4:** 5-7.

Hogan, P.G., L. Chen, J. Nardone, and A. Rao. 2003. Transcriptional regulation by calcium, calcineurin, and NFAT. *Genes Dev* **17:** 2205-2232.

Hogga, I. and F. Karch. 2002. Transcription through the iab-7 cis-regulatory domain of

the bithorax complex interferes with maintenance of Polycomb-mediated silencing. *Development* **129:** 4915-4922.

Hogness, D.S., H.D. Lipshitz, P.A. Beachy, D.A. Peattie, R.B. Saint, M. Goldschmidt-Clermont, P.J. Harte, E.R. Gavis, and S.L. Helfand. 1985. Regulation and products of the Ubx domain of the bithorax complex. *Cold Spring Harb Symp Quant Biol* **50:** 181-194.

Hongay, C.F., P.L. Grisafi, T. Galitski, and G.R. Fink. 2006. Antisense transcription controls cell fate in Saccharomyces cerevisiae. *Cell* **127:** 735-745.

Houseley, J., L. Rubbi, M. Grunstein, D. Tollervey, and M. Vogelauer. 2008. A ncRNA modulates histone modification and mRNA induction in the yeast GAL gene cluster. *Mol Cell* **32:** 685-695.

Hu, J.F., K.A. Balaguru, R.D. Ivaturi, H. Oruganti, T. Li, B.T. Nguyen, T.H. Vu, and A.R. Hoffman. 1999. Lack of reciprocal genomic imprinting of sense and antisense RNA of mouse insulin-like growth factor II receptor in the central nervous system. *Biochem Biophys Res Commun* **257:** 604-608.

Hutchinson, J.N., A.W. Ensminger, C.M. Clemson, C.R. Lynch, J.B. Lawrence, and A. Chess. 2007. A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains. *BMC Genomics* **8:** 39.

International Human Genome Sequencing Consortium. 2004. Finishing the euchromatic sequence of the human genome. *Nature* **431:** 931-945.

Janicki, S.M., T. Tsukamoto, S.E. Salghetti, W.P. Tansey, R. Sachidanandam, K.V. Prasanth, T. Ried, Y. Shav-Tal, E. Bertrand, R.H. Singer et al. 2004. From silencing to gene expression: real-time analysis in single cells. *Cell* **116:** 683-698.

Jaskiewicz, L. and W. Filipowicz. 2008. Role of Dicer in posttranscriptional RNA silencing. *Curr Top Microbiol Immunol* **320:** 77-97.

Ji, P., S. Diederichs, W. Wang, S. Boing, R. Metzger, P.M. Schneider, N. Tidow, B. Brandt, H. Buerger, E. Bulk et al. 2003. MALAT-1, a novel noncoding RNA, and thymosin beta4 predict metastasis and survival in early-stage non-small cell lung cancer. *Oncogene* **22:** 8031-8041.

Kaiser, T.E., R.V. Intine, and M. Dundr. 2008. De novo formation of a subnuclear body. *Science* **322:** 1713-1717.

Kanduri, C., N. Thakur, and R.R. Pandey. 2006. The length of the transcript encoded from the Kcnq1ot1 antisense promoter determines the degree of silencing. *Embo J* **25:** 2096-2106.

Kapranov, P., J. Cheng, S. Dike, D.A. Nix, R. Duttagupta, A.T. Willingham, P.F. Stadler, J. Hertel, J. Hackermuller, I.L. Hofacker et al. 2007a. RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science* **316:** 1484-1488.

Kapranov, P., A.T. Willingham, and T.R. Gingeras. 2007b. Genome-wide transcription and the implications for genomic organization. *Nat Rev Genet* **8:** 413-423.

Karch, F., B. Weiffenbach, M. Peifer, W. Bender, I. Duncan, S. Celniker, M. Crosby, and E.B. Lewis. 1985. The abdominal region of the bithorax complex. *Cell* **43:** 81-96.

Katayama, S., Y. Tomaru, T. Kasukawa, K. Waki, M. Nakanishi, M. Nakamura, H. Nishida, C.C. Yap, M. Suzuki, J. Kawai et al. 2005. Antisense transcription in the mammalian transcriptome. *Science* **309:** 1564-1566.

Kent, W.J. 2002. BLAT--the BLAST-like alignment tool. *Genome Res* **12:** 656-664.

Kloc, M., S. Bilinski, and M.T. Dougherty. 2007. Organization of cytokeratin cytoskeleton and germ plasm in the vegetal cortex of Xenopus laevis oocytes depends on coding and non-coding RNAs: three-dimensional and ultrastructural analysis. *Exp Cell Res* **313:** 1639-1651.

Kloc, M., K. Wilk, D. Vargas, Y. Shirato, S. Bilinski, and L.D. Etkin. 2005. Potential structural role of non-coding and coding RNAs in the organization of the cytoskeleton at the vegetal cortex of Xenopus oocytes. *Development* **132:** 3445-3457.

Klymenko, T. and J. Muller. 2004. The histone methyltransferases Trithorax and Ash1 prevent transcriptional silencing by Polycomb group proteins. *EMBO Rep* **5:** 373-377.

Kondo, T., Y. Hashimoto, K. Kato, S. Inagaki, S. Hayashi, and Y. Kageyama. 2007. Small

peptide regulators of actin-based cell morphogenesis encoded by a polycistronic mRNA. *Nat Cell Biol* **9:** 660-665.

Krystal, G.W., B.C. Armstrong, and J.F. Battey. 1990. N-myc mRNA forms an RNA-RNA duplex with endogenous antisense transcripts. *Mol Cell Biol* **10:** 4180-4191.

Kumar, M. and G.G. Carmichael. 1997. Nuclear antisense RNA induces extensive adenosine modifications and nuclear retention of target transcripts. *Proc Natl Acad Sci U S A* **94:** 3542-3547.

Kuroda, M., J. Sok, L. Webb, H. Baechtold, F. Urano, Y. Yin, P. Chung, D.G. de Rooij, A. Akhmedov, T. Ashley et al. 2000. Male sterility and enhanced radiation sensitivity in TLS(-/-) mice. *Embo J* **19:** 453-462.

Kuwahara, S., A. Ikei, Y. Taguchi, Y. Tabuchi, N. Fujimoto, M. Obinata, S. Uesugi, and Y. Kurihara. 2006. PSPC1, NONO, and SFPQ are expressed in mouse Sertoli cells and may function as coregulators of androgen receptor-mediated transcription. *Biol Reprod* **75:** 352-359.

Lam, Y.W., C.E. Lyon, and A.I. Lamond. 2002. Large-scale isolation of Cajal bodies from HeLa cells. *Mol Biol Cell* **13:** 2461-2473.

Lanz, R.B., N.J. McKenna, S.A. Onate, U. Albrecht, J. Wong, S.Y. Tsai, M.J. Tsai, and B.W. O'Malley. 1999. A steroid receptor coactivator, SRA, functions as an RNA and is present in an SRC-1 complex. *Cell* **97:** 17-27.

Lecuyer, E., H. Yoshida, N. Parthasarathy, C. Alm, T. Babak, T. Cerovina, T.R. Hughes, P. Tomancak, and H.M. Krause. 2007. Global analysis of mRNA localization reveals a prominent role in organizing cellular architecture and function. *Cell* **131:** 174-187.

Lee, J.T., L.S. Davidow, and D. Warshawsky. 1999a. Tsix, a gene antisense to Xist at the X-inactivation centre. *Nat Genet* **21:** 400-404.

Lee, M.P., M.R. DeBaun, K. Mitsuya, H.L. Galonek, S. Brandenburg, M. Oshimura, and A.P. Feinberg. 1999b. Loss of imprinting of a paternally expressed transcript, with antisense orientation to KVLQT1, occurs frequently in Beckwith-Wiedemann syndrome and is independent of insulin-like growth factor II imprinting. *Proc Natl Acad Sci U S A* **96:** 5203-5208.

Lee, Y., J. Han, K.H. Yeom, H. Jin, and V.N. Kim. 2006. Drosha in primary microRNA processing. *Cold Spring Harb Symp Quant Biol* **71:** 51-57.

Lee, Y., M. Kim, J. Han, K.H. Yeom, S. Lee, S.H. Baek, and V.N. Kim. 2004. MicroRNA genes are transcribed by RNA polymerase II. *Embo J* **23:** 4051-4060.

Lehnert, S., A. Reverter, K. Byrne, Y. Wang, G. Nattrass, N. Hudson, and P. Greenwood. 2007. Gene expression studies of developing bovine longissimus muscle from two different beef cattle breeds. *BMC Developmental Biology* **7:** 95.

Lempradl, A. and L. Ringrose. 2008. How does noncoding transcription regulate Hox genes? *Bioessays* **30:** 110-121.

Leung, A.K., D. Gerlich, G. Miller, C. Lyon, Y.W. Lam, D. Lleres, N. Daigle, J. Zomerdijk, J. Ellenberg, and A.I. Lamond. 2004. Quantitative kinetic analysis of nucleolar breakdown and reassembly during mitosis in live human cells. *J Cell Biol* **166:** 787-800.

Leygue, E. 2007. Steroid receptor RNA activator (SRA1): unusual bifaceted gene products with suspected relevance to breast cancer. *Nucl Recept Signal* **5:** e006.

Liang, F., I. Holt, G. Pertea, S. Karamycheva, S.L. Salzberg, and J. Quackenbush. 2000. Gene index analysis of the human genome estimates approximately 120,000 genes. *Nat Genet* **25:** 239-240.

Lin, R., S. Maeda, C. Liu, M. Karin, and T.S. Edgington. 2007. A large noncoding RNA is a marker for murine hepatocellular carcinomas and a spectrum of human carcinomas. *Oncogene* **26:** 851-858.

Lipshitz, H.D., D.A. Peattie, and D.S. Hogness. 1987. Novel transcripts from the Ultrabithorax domain of the bithorax complex. *Genes Dev* **1:** 307-322.

Liu, H., G.C. Ippolito, J.K. Wall, T. Niu, L. Probst, B.S. Lee, K. Pulford, A.H. Banham, L. Stockwin, A.L. Shaffer et al. 2006. Functional studies of BCL11A: characterization of the conserved BCL11A-XL splice variant and its interaction with BCL6 in nuclear paraspeckles of germinal center B cells. *Mol Cancer* **5:** 18.

Luikenhuis, S., A. Wutz, and R. Jaenisch. 2001. Antisense transcription through the Xist locus mediates Tsix function in embryonic stem cells. *Mol Cell Biol* **21:** 8512-

8520.

Lyle, R., D. Watanabe, D. te Vruchte, W. Lerchner, O.W. Smrzka, A. Wutz, J. Schageman, L. Hahner, C. Davies, and D.P. Barlow. 2000. The imprinted antisense RNA at the Igf2r locus overlaps but does not imprint Mas1. *Nat Genet* **25:** 19-21.

Maeda, R.K. and F. Karch. 2006. The ABC of the BX-C: the bithorax complex explained. *Development* **133:** 1413-1422.

Mancini-Dinardo, D., S.J. Steele, J.M. Levorse, R.S. Ingram, and S.M. Tilghman. 2006. Elongation of the Kcnq1ot1 transcript is required for genomic imprinting of neighboring genes. *Genes Dev* **20:** 1268-1282.

Marahrens, Y., B. Panning, J. Dausman, W. Strauss, and R. Jaenisch. 1997. Xist-deficient mice are defective in dosage compensation but not spermatogenesis. *Genes Dev* **11:** 156-166.

Mariner, P.D., R.D. Walters, C.A. Espinoza, L.F. Drullinger, S.D. Wagner, J.F. Kugel, and J.A. Goodrich. 2008. Human Alu RNA is a modular transacting repressor of mRNA transcription during heat shock. *Mol Cell* **29:** 499-509.

Martens, J.A., L. Laprade, and F. Winston. 2004. Intergenic transcription is required to repress the Saccharomyces cerevisiae SER3 gene. *Nature* **429:** 571-574.

Martianov, I., A. Ramadass, A. Serra Barros, N. Chow, and A. Akoulitchev. 2007. Repression of the human dihydrofolate reductase gene by a non-coding interfering transcript. *Nature* **445:** 666-670.

Masters, J.N. and G. Attardi. 1985. Discrete human dihydrofolate reductase gene transcripts present in polysomal RNA map with their 5' ends several hundred nucleotides upstream of the main mRNA start site. *Mol Cell Biol* **5:** 493-500.

Mattick, J.S. 2004. RNA regulation: a new genetics? *Nat Rev Genet* **5:** 316-323.

Mattick, J.S. and I.V. Makunin. 2006. Non-coding RNA. *Hum Mol Genet* **15 Spec No 1:** R17-29.

Mercer, T.R., M.E. Dinger, and J.S. Mattick. 2009. Long non-coding RNAs: insights into functions. *Nat Rev Genet* **10:** 155-159.

Mercer, T.R., M.E. Dinger, S.M. Sunkin, M.F. Mehler, and J.S. Mattick. 2008. Specific expression of long noncoding RNAs in the mouse brain. *Proc Natl Acad Sci U S A* **105:** 716-721.

Mikkelsen, T.S., M. Ku, D.B. Jaffe, B. Issac, E. Lieberman, G. Giannoukos, P. Alvarez, W. Brockman, T.K. Kim, R.P. Koche et al. 2007. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448:** 553-560.

Mintz, P.J., S.D. Patterson, A.F. Neuwald, C.S. Spahr, and D.L. Spector. 1999. Purification and biochemical characterization of interchromatin granule clusters. *Embo J* **18:** 4308-4320.

Mohammad, F., R.R. Pandey, T. Nagano, L. Chakalova, T. Mondal, P. Fraser, and C. Kanduri. 2008. Kcnq1ot1/Lit1 noncoding RNA mediates transcriptional silencing by targeting to the perinucleolar region. *Mol Cell Biol* **28:** 3713-3728.

Munroe, S.H. and M.A. Lazar. 1991. Inhibition of c-erbA mRNA splicing by a naturally occurring antisense RNA. *J Biol Chem* **266:** 22083-22086.

Murakami, K., M. Oshimura, and H. Kugoh. 2007. Suggestive evidence for chromosomal localization of non-coding RNA from imprinted LIT1. *J Hum Genet* **52:** 926-933.

Muratani, M., D. Gerlich, S.M. Janicki, M. Gebhard, R. Eils, and D.L. Spector. 2002. Metabolic-energy-dependent movement of PML bodies within the mammalian cell nucleus. *Nat Cell Biol* **4:** 106-110.

Nagano, T., J.A. Mitchell, L.A. Sanz, F.M. Pauler, A.C. Ferguson-Smith, R. Feil, and P. Fraser. 2008. The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. *Science* **322:** 1717-1720.

Neil, H., C. Malabat, Y. d'Aubenton-Carafa, Z. Xu, L.M. Steinmetz, and A. Jacquier. 2009. Widespread bidirectional promoters are the major source of cryptic transcripts in yeast. *Nature* **457:** 1038-1042.

O'Neill, M.J. 2005. The influence of non-coding RNAs on allele-specific gene expression in mammals. *Hum Mol Genet* **14 Spec No 1:** R113-120.

Ogawa, Y., B.K. Sun, and J.T. Lee. 2008. Intersection of the RNA interference and X-inactivation pathways. *Science* **320:** 1336-1341.

Okamura, K., W.J. Chung, J.G. Ruby, H. Guo, D.P. Bartel, and E.C. Lai. 2008. The Drosophila hairpin RNA pathway generates endogenous short interfering RNAs. *Nature* **453:** 803-806.

Pandey, R.R., T. Mondal, F. Mohammad, S. Enroth, L. Redrup, J. Komorowski, T. Nagano, D. Mancini-Dinardo, and C. Kanduri. 2008. Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. *Mol Cell* **32:** 232-246.

Penny, G.D., G.F. Kay, S.A. Sheardown, S. Rastan, and N. Brockdorff. 1996. Requirement for Xist in X chromosome inactivation. *Nature* **379:** 131-137.

Perry, R.P., D.E. Kelley, and J. LaTorre. 1974. Synthesis and turnover of nuclear and cytoplasmic polyadenylic acid in mouse L cells. *J Mol Biol* **82:** 315-331.

Peters, J. and J.E. Robson. 2008. Imprinted noncoding RNAs. *Mamm Genome* **19:** 493-502.

Petruk, S., Y. Sedkov, K.M. Riley, J. Hodgson, F. Schweisguth, S. Hirose, J.B. Jaynes, H.W. Brock, and A. Mazo. 2006. Transcription of bxd noncoding RNAs promoted by trithorax represses Ubx in cis by transcriptional interference. *Cell* **127:** 1209-1221.

Peyman, J.A. 1999. Repression of major histocompatibility complex genes by a human trophoblast ribonucleic acid. *Biol Reprod* **60:** 23-31.

Pheasant, M. and J.S. Mattick. 2007. Raising the estimate of functional human sequences. *Genome Res* **17:** 1245-1253.

Platani, M., I. Goldberg, A.I. Lamond, and J.R. Swedlow. 2002. Cajal body dynamics and association with chromatin are ATP-dependent. *Nat Cell Biol* **4:** 502-508.

Plath, K., S. Mlynarczyk-Evans, D.A. Nusinow, and B. Panning. 2002. Xist RNA and the mechanism of X chromosome inactivation. *Annu Rev Genet* **36:** 233-278.

Pollard, K.S., S.R. Salama, N. Lambert, M.A. Lambot, S. Coppens, J.S. Pedersen, S. Katzman, B. King, C. Onodera, A. Siepel et al. 2006. An RNA gene expressed during cortical development evolved rapidly in humans. *Nature* **443:** 167-172.

Ponjavic, J., C.P. Ponting, and G. Lunter. 2007. Functionality or transcriptional noise? Evidence for selection within long noncoding RNAs. *Genome Res* **17:** 556-565.

Prasanth, K.V., S.G. Prasanth, Z. Xuan, S. Hearn, S.M. Freier, C.F. Bennett, M.Q. Zhang, and D.L. Spector. 2005. Regulating gene expression through RNA nuclear retention. *Cell* **123:** 249-263.

Prasanth, K.V. and D.L. Spector. 2007. Eukaryotic regulatory RNAs: an answer to the 'genome complexity' conundrum. *Genes Dev* **21:** 11-42.

Preker, P., J. Nielsen, S. Kammler, S. Lykke-Andersen, M.S. Christensen, C.K. Mapendano, M.H. Schierup, and T.H. Jensen. 2008. RNA exosome depletion reveals transcription upstream of active human promoters. *Science* **322:** 1851-1854.

Pueyo, J.I. and J.P. Couso. 2008. The 11-aminoacid long Tarsal-less peptides trigger a cell signal in Drosophila leg development. *Dev Biol* **324:** 192-201.

Rank, G., M. Prestel, and R. Paro. 2002. Transcription through intergenic chromosomal memory elements of the Drosophila bithorax complex correlates with an epigenetic switch. *Mol Cell Biol* **22:** 8026-8034.

Ravasi, T., H. Suzuki, K.C. Pang, S. Katayama, M. Furuno, R. Okunishi, S. Fukuda, K. Ru, M.C. Frith, M.M. Gongora et al. 2006. Experimental validation of the regulated expression of large numbers of non-coding RNAs from the mouse genome. *Genome Res* **16:** 11-19.

Rinn, J.L., M. Kertesz, J.K. Wang, S.L. Squazzo, X. Xu, S.A. Brugmann, L.H. Goodnough, J.A. Helms, P.J. Farnham, E. Segal et al. 2007. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* **129:** 1311-1323.

Robb, G.B., K.M. Brown, J. Khurana, and T.M. Rana. 2005. Specific and potent RNAi in the nucleus of human cells. *Nat Struct Mol Biol* **12:** 133-137.

Ronshaugen, M., F. Biemar, J. Piel, M. Levine, and E.C. Lai. 2005. The Drosophila microRNA iab-4 causes a dominant homeotic transformation of halteres to wings. *Genes Dev* **19:** 2947-2952.

Saha, S., S. Murthy, and P.N. Rangarajan. 2006. Identification and characterization of a virus-inducible non-coding RNA in mouse brain. *J Gen Virol* **87:** 1991-1995.

Saitoh, N., C.S. Spahr, S.D. Patterson, P. Bubulya, A.F. Neuwald, and D.L. Spector. 2004. Proteomic analysis of interchromatin granule clusters. *Mol Biol Cell* **15:** 3876-3890.

Sanchez-Elsner, T., D. Gou, E. Kremmer, and F. Sauer. 2006. Noncoding RNAs of trithorax response elements recruit Drosophila Ash1 to Ultrabithorax. *Science* **311:** 1118-1123.

Sasaki, Y.T., T. Ideue, M. Sano, T. Mituyama, and T. Hirose. 2009. MENepsilon/beta noncoding RNAs are essential for structural integrity of nuclear paraspeckles. *Proc Natl Acad Sci U S A* **106:** 2525-2530.

Scherl, A., Y. Coute, C. Deon, A. Calle, K. Kindbeiter, J.C. Sanchez, A. Greco, D. Hochstrasser, and J.J. Diaz. 2002. Functional proteomic analysis of human nucleolus. *Mol Biol Cell* **13:** 4100-4109.

Schmitt, S., M. Prestel, and R. Paro. 2005. Intergenic transcription through a polycomb group response element counteracts silencing. *Genes Dev* **19:** 697-708.

Seidl, C.I., S.H. Stricker, and D.P. Barlow. 2006. The imprinted Air ncRNA is an atypical RNAPII transcript that evades splicing and escapes nuclear export. *Embo J* **25:** 3565-3575.

Seila, A.C., J.M. Calabrese, S.S. Levine, G.W. Yeo, P.B. Rahl, R.A. Flynn, R.A. Young, and P.A. Sharp. 2008. Divergent transcription from active promoters. *Science* **322:** 1849-1851.

Sessa, L., A. Breiling, G. Lavorgna, L. Silvestri, G. Casari, and V. Orlando. 2007. Noncoding RNA synthesis and loss of Polycomb group repression accompanies the colinear activation of the human HOXA cluster. *Rna* **13:** 223-239.

Sha, K. 2008. A mechanistic view of genomic imprinting. *Annu Rev Genomics Hum Genet* **9:** 197-216.

Shamovsky, I., M. Ivannikov, E.S. Kandel, D. Gershon, and E. Nudler. 2006. RNA-mediated response to heat shock in mammalian cells. *Nature* **440:** 556-560.

Shen, X., J.M. Collier, M. Hlaing, L. Zhang, E.H. Delshad, J. Bristow, and H.S. Bernstein. 2003. Genome-wide examination of myoblast cell cycle withdrawal during differentiation. *Dev Dyn* **226:** 128-138.

Sleutels, F., R. Zwart, and D.P. Barlow. 2002. The non-coding Air RNA is required for silencing autosomal imprinted genes. *Nature* **415:** 810-813.

Smetana, K., W.J. Steele, and H. Busch. 1963. A nuclear ribonucleoprotein network. *Experimental Cell Research* **31:** 198-201.

Smilinich, N.J., C.D. Day, G.V. Fitzpatrick, G.M. Caldwell, A.C. Lossie, P.R. Cooper, A.C. Smallwood, J.A. Joyce, P.N. Schofield, W. Reik et al. 1999. A maternally methylated CpG island in KvLQT1 is associated with an antisense paternal transcript and loss of imprinting in Beckwith-Wiedemann syndrome. *Proc Natl Acad Sci U S A* **96:** 8064-8069.

Smyth, G.K. 2004. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* **3:** Article3.

Smyth, G.K. and T. Speed. 2003. Normalization of cDNA microarray data. *Methods* **31:** 265-273.

Sone, M., T. Hayashi, H. Tarui, K. Agata, M. Takeichi, and S. Nakagawa. 2007. The mRNA-like noncoding RNA Gomafu constitutes a novel nuclear domain in a subset of neurons. *J Cell Sci* **120:** 2498-2506.

Spector, D.L. 2001. Nuclear domains. *J Cell Sci* **114:** 2891-2893.

Spector, D.L. 2006. SnapShot: Cellular bodies. *Cell* **127:** 1071.

Stavropoulos, N., N. Lu, and J.T. Lee. 2001. A functional role for Tsix transcription in blocking Xist RNA accumulation but not in X-chromosome choice. *Proc Natl Acad Sci U S A* **98:** 10232-10237.

Stavropoulos, N., R.K. Rowntree, and J.T. Lee. 2005. Identification of developmentally specific enhancers for Tsix in the regulation of X chromosome inactivation. *Mol Cell Biol* **25:** 2757-2769.

Struhl, K. 2007. Transcriptional noise and the fidelity of initiation by RNA polymerase II.

Nat Struct Mol Biol **14:** 103-105.

Sunwoo, H., M.E. Dinger, J.E. Wilusz, P.P. Amaral, J.S. Mattick, and D.L. Spector. 2009. MEN {varepsilon}/{beta} nuclear-retained non-coding RNAs are up-regulated upon muscle differentiation and are essential components of paraspeckles. *Genome Res* **19:** 347-359.

Tam, O.H., A.A. Aravin, P. Stein, A. Girard, E.P. Murchison, S. Cheloufi, E. Hodges, M. Anger, R. Sachidanandam, R.M. Schultz et al. 2008. Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. *Nature* **453:** 534-538.

Topisirovic, I., B. Culjkovic, N. Cohen, J.M. Perez, L. Skrabanek, and K.L. Borden. 2003. The proline-rich homeodomain protein, PRH, is a tissue-specific inhibitor of eIF4E-dependent cyclin D1 mRNA transport and growth. *Embo J* **22:** 689-703.

Uhler, J.P., C. Hertel, and J.Q. Svejstrup. 2007. A role for noncoding transcription in activation of the yeast PHO5 gene. *Proc Natl Acad Sci U S A* **104:** 8011-8016.

Uranishi, H., T. Tetsuka, M. Yamashita, K. Asamitsu, M. Shimizu, M. Itoh, and T. Okamoto. 2001. Involvement of the pro-oncoprotein TLS (translocated in liposarcoma) in nuclear factor-kappa B p65-mediated transcription as a coactivator. *J Biol Chem* **276:** 13395-13401.

Wang, J., J. Zhang, H. Zheng, J. Li, D. Liu, H. Li, R. Samudrala, J. Yu, and G.K. Wong. 2004. Mouse transcriptome: neutral evolution of 'non-coding' complementary DNAs. *Nature* **431:** 1 p following 757; discussion following 757.

Wang, X., S. Arai, X. Song, D. Reichart, K. Du, G. Pascual, P. Tempst, M.G. Rosenfeld, C.K. Glass, and R. Kurokawa. 2008. Induced ncRNAs allosterically modify RNA-binding proteins in cis to inhibit transcription. *Nature* **454:** 126-130.

Watanabe, T., Y. Totoki, A. Toyoda, M. Kaneda, S. Kuramochi-Miyagawa, Y. Obata, H. Chiba, Y. Kohara, T. Kono, T. Nakano et al. 2008. Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature* **453:** 539-543.

Waterston, R.H. K. Lindblad-Toh E. Birney J. Rogers J.F. Abril P. Agarwal R. Agarwala R. Ainscough M. Alexandersson P. An et al. 2002. Initial sequencing and

comparative analysis of the mouse genome. *Nature* **420:** 520-562.

Willingham, A.T., A.P. Orth, S. Batalov, E.C. Peters, B.G. Wen, P. Aza-Blanc, J.B. Hogenesch, and P.G. Schultz. 2005. A strategy for probing the function of noncoding RNAs finds a repressor of NFAT. *Science* **309:** 1570-1573.

Wilusz, J.E., S.M. Freier, and D.L. Spector. 2008. 3' End Processing of a Long Nuclear-Retained Noncoding RNA Yields a tRNA-like Cytoplasmic RNA. *Cell* **135:** 919-932.

Wilusz, J.E., H. Sunwoo, and D.L. Spector. 2009. Long non-coding RNAs: Functional surprises from the RNA world. *Genes Dev. in press*.

Wyers, F., M. Rougemaille, G. Badis, J.C. Rousselle, M.E. Dufour, J. Boulay, B. Regnault, F. Devaux, A. Namane, B. Seraphin et al. 2005. Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. *Cell* **121:** 725-737.

Xu, Z., W. Wei, J. Gagneur, F. Perocchi, S. Clauder-Munster, J. Camblong, E. Guffanti, F. Stutz, W. Huber, and L.M. Steinmetz. 2009. Bidirectional promoters generate pervasive transcription in yeast. *Nature* **457:** 1033-1037.

Yamada, K., J. Kano, H. Tsunoda, H. Yoshikawa, C. Okubo, T. Ishiyama, and M. Noguchi. 2006. Phenotypic characterization of endometrial stromal sarcoma of the uterus. *Cancer Sci* **97:** 106-112.

Yan, M.D., C.C. Hong, G.M. Lai, A.L. Cheng, Y.W. Lin, and S.E. Chuang. 2005. Identification and characterization of a novel gene Saf transcribed from the opposite strand of Fas. *Hum Mol Genet* **14:** 1465-1474.

Yang, L., L.J. Embree, S. Tsai, and D.D. Hickstein. 1998. Oncoprotein TLS interacts with serine-arginine proteins involved in RNA splicing. *J Biol Chem* **273:** 27761-27764.

Yu, W., D. Gius, P. Onyango, K. Muldoon-Jacobs, J. Karp, A.P. Feinberg, and H. Cui. 2008. Epigenetic silencing of tumour suppressor gene p15 by its antisense RNA. *Nature* **451:** 202-206.

Zeng, Y. and B.R. Cullen. 2002. RNA interference in human cells is restricted to the

cytoplasm. *Rna* **8:** 855-860.

Zhang, L.F., K.D. Huynh, and J.T. Lee. 2007. Perinucleolar targeting of the inactive X during S phase: evidence for a role in the maintenance of silencing. *Cell* **129:** 693-706.

Zhang, Z. and G.G. Carmichael. 2001. The fate of dsRNA in the nucleus: a p54(nrb)-containing complex mediates the nuclear retention of promiscuously A-to-I edited RNAs. *Cell* **106:** 465-475.

Zhao, J., B.K. Sun, J.A. Erwin, J.J. Song, and J.T. Lee. 2008. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science* **322:** 750-756.

Zwart, R., F. Sleutels, A. Wutz, A.H. Schinkel, and D.P. Barlow. 2001. Bidirectional action of the Igf2r imprint control element on upstream and downstream imprinted genes. *Genes Dev* **15:** 2361-2366.