HOKKAIDO UNIVERSITY

Title	Structural, super-resolution microscopy analysis of paraspeckle nuclear body organization
Author(s)	West, Jason A.; Mito, Mari; Kurosaka, Satoshi et al.
Citation	Journal of cell biology, 214(7), 817-830 https://doi.org/10.1083/jcb.201601071
Issue Date	2016-09-27
Doc URL	https://hdl.handle.net/2115/63798
Rights	©West et al., 2016. Originally published in The Journal of Cell Biology. doi:10.1083/jcb.201601071
Type	journal article
File Information	manuscript.pdf

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

Published September 19, 2016

This article has original data in the JCB Data Viewer
http://jcb-dataviewer.rupress.org/jcb/browse/12302

Structural, super-resolution microscopy analysis of paraspeckle nuclear body organization

Jason A. West,[1,2] Mari Mito,[3] Satoshi Kurosaka,[4] Toru Takumi,[4] Chiharu Tanegashima,[5] Takeshi Chujo,[6] Kaori Yanaka,[3] Robert E. Kingston,[1,2] Tetsuro Hirose,[6] Charles Bond,[7] Archa Fox,[8,9] and Shinichi Nakagawa[3,10]

[1]Department of Molecular Biology, Massachusetts General Hospital, Boston, MA 02114
[2]Department of Genetics, Harvard Medical School, Boston, MA 02115
[3]RNA Biology Laboratory, RIKEN, Wako 351-0198, Japan
[4]RIKEN Brain Science Institute, Wako, Saitama 351-0198, Japan
[5]Phyloinformatics Unit, RIKEN Center for Life Science Technologies, Chuo-ku, Kobe, Hyogo 650-0047, Japan
[6]Institute for Genetic Medicine, Hokkaido University, Sapporo 060-0815, Japan
[7]School of Chemistry and Biochemistry and [8]School of Anatomy, Physiology and Human Biology, The University of Western Australia, Crawley, Western Australia 6009, Australia
[9]Harry Perkins Institute of Medical Research, Queen Elizabeth II Medical Centre, Nedlands, Western Australia 6009, Australia
[10]RNA Biology Laboratory, Faculty of Pharmaceutical Sciences, Hokkaido University, Sapporo 060-0812, Japan

Paraspeckles are nuclear bodies built on the long noncoding RNA *Neat1*, which regulates a variety of physiological processes including cancer progression and corpus luteum formation. To obtain further insight into the molecular basis of the function of paraspeckles, we performed fine structural analyses of these nuclear bodies using structural illumination microscopy. Notably, paraspeckle proteins are found within different layers along the radially arranged bundles of *Neat1* transcripts, forming a characteristic core-shell spheroidal structure. In cells lacking the RNA binding protein Fus, paraspeckle spheroids are disassembled into smaller particles containing *Neat1*, which are diffusely distributed in the nucleoplasm. Sequencing analysis of RNAs purified from paraspeckles revealed that AG-rich transcripts associate with *Neat1*, which are distributed along the shell of the paraspeckle spheroids. We propose that paraspeckles sequester core components inside the spheroids, whereas the outer surface associates with other components in the nucleoplasm to fulfill their function.

Introduction

The nucleus is highly structured and organized into several nonmembranous nuclear bodies. These bodies contain discrete sets of proteins and nucleic acids that are involved in particular nuclear processes (Platani and Lamond, 2004). Paraspeckles were originally described as nuclear bodies that are enriched in the Drosophila behavior and human splicing (DBHS) family of RNA-binding proteins (Fox et al., 2002, 2005). Paraspeckles have since been found to be identical to interchromatin granule–associated zone, which are observed as electron-dense structures using electron microscopy (Cardinale et al., 2007; Bond and Fox, 2009). *Neat1* is a mammalian-specific, long noncoding RNA (lncRNA) and serves as an architectural component of paraspeckles. Depletion of *Neat1* leads to the disassembly of these bodies (Chen and Carmichael, 2009; Clemson et al., 2009; Sasaki et al., 2009; Sunwoo et al., 2009). Two isoforms of *Neat1* are made from a common promoter: the longer (20 kb in mice) isoform *Neat1_2* is required for the formation of paraspeckles, whereas the shorter (3.2 kb in mice) isoform *Neat1_1* is not necessary for its architectural function (Nakagawa et al., 2011; Naganuma et al., 2012). To date, >40 proteins are known to accumulate in paraspeckles. These proteins can be divided into three categories depending on the extent of paraspeckle disruption induced upon depletion of each protein (Naganuma et al., 2012). Category I proteins are essential for the structural maintenance of paraspeckles. They are further subdivided into category Ia proteins, which are required for the production or stabilization of *Neat1_2* (e.g., Sfpq, Nono, and Rbm14), and category Ib proteins, which do not affect the amount of *Neat1_2* (e.g., Fus/Tls and Brg1) (Sasaki et al., 2009; Naganuma et al., 2012; Hennig et al., 2015). The depletion of category II proteins (e.g., Tardbp) results in a substantial decrease in the number of paraspeckle-possessing cells. Category III proteins (e.g., Pspc1) do not have an apparent effect on

Correspondence to S. Nakagawa: nakagawas@pharm.hokudai.ac.jp
J.A. West's present address is Biogen, Cambridge, MA 02142.

Abbreviations used: CHART, capture hybridization analysis of RNA targets; CLIP, cross-linking immunoprecipitation; CLIP-seq, cross-linking immunoprecipitation sequencing; DBHS, Drosophila behavior and human splicing; KO, knockout; lncRNA, long noncoding RNA; MEF, mouse embryonic fibroblast; MEME, Multiple Em for Motif Elicitation; PrLD, prion-like domain; SIM, structured illumination microscopy; WT, wild-type.

© 2016 West et al. This article is distributed under the terms of an Attribution–Noncommercial–Share Alike–No Mirror Sites license for the first six months after the publication date (see http://www.rupress.org/terms). After six months it is available under a Creative Commons License (Attribution–Noncommercial–Share Alike 3.0 Unported license, as described at http://creativecommons.org/licenses/by-nc-sa/3.0/).

Supplemental Material can be found at:
/content/suppl/2016/09/07/jcb.201601071.DC1.html
Original image data can be found at:
http://jcb-dataviewer.rupress.org/jcb/browse/12302

The Rockefeller University Press $30.00
J. Cell Biol. Vol. 214 No. 7 817–830
www.jcb.org/cgi/doi/10.1083/jcb.201601071

THE JOURNAL OF CELL BIOLOGY

Downloaded from on December 11, 2016

Published September 19, 2016

paraspeckle formation (Naganuma et al., 2012). All paraspeckle proteins exhibit RNA-binding capacities but are not necessarily involved in common biological processes.

At the molecular level, paraspeckles have been proposed to sequester proteins or transcripts into the nuclear bodies, serving as molecular sponges that modulate the levels of active molecules in the nucleoplasm (Hirose et al., 2014; Imamura et al., 2014). Paraspeckles have been proposed to regulate a variety of cellular processes, including the nuclear retention of hyper A-to-I–edited mRNAs (Prasanth et al., 2005; Chen and Carmichael, 2009), the control of transcription via the sequestration of Sfpq (Hirose et al., 2014), and immune responses to polyinosinic-polycytidylic acid double-stranded nucleotides in particular cells (Imamura et al., 2014). In mice, *Neat1_1* is expressed in a wide variety of cell types, whereas *Neat1_2*, the architectural component of paraspeckles, is expressed only in a subpopulation of restricted cell types (Nakagawa et al., 2011). Accordingly, prominent paraspeckle formation is observed only in particular cell populations that abundantly express *Neat1_2*, including corpus luteal cells, which produce the steroid hormone progesterone that is essential for pregnancy (Nakagawa et al., 2011). Consistent with this expression pattern, the fertility of female *Neat1* knockout (KO) mice is severely impaired as the result of a lack of the formation of pregnant corpus luteum and a subsequent decrease in serum progesterone (Nakagawa et al., 2014). Paraspeckles have also been suggested to be involved in multiple physiological processes, including mammary gland development (Standaert et al., 2014) and prostate cancer progression (Chakravarty et al., 2014).

Previous observations using electron microscopy have revealed that the paraspeckles are usually detected as electron-dense, irregular sausage-like structures (Souquere et al., 2010). Interestingly, *Neat1_2* is arranged in an ordered manner in paraspeckles, with the 5′ and 3′ ends located in the periphery and the middle of *Neat1_2* found in the central paraspeckle region (Souquere et al., 2010). In addition, the length of the short axis of paraspeckles is constrained (∼360 nm in human cells), whereas the long axis is quite variable. These observations lead to the idea that Neat1_2 is radially arranged along the transverse plane of the sausage-like paraspeckles, providing a structural scaffold for the assembly of paraspeckle proteins. However, it remains unclear how protein components of paraspeckles are arranged in relation to the ordered architectural arrangement of *Neat1_2* transcripts and how sequestered molecules are retained within paraspeckles. Because the diameter of a paraspeckle is ∼300 nm (Souquere et al., 2010), i.e., close to the diffraction limit of light (∼200 nm), it is difficult to examine the fine internal structures of paraspeckles using conventional light microscopy or even confocal laser-scanning microscopy. To overcome this limitation, several super-resolution techniques based on different principles have recently become available, including structured illumination microscopy (SIM), stimulated emission depletion microscopy, and various localization microscopy techniques such as stochastic optical reconstruction microscopy and photoactivation localization microscopy (Schermelleh et al., 2010). SIM improves the resolution by a factor of two, achieving resolution near 100 nm in the xy axis (Gustafsson, 2000). SIM is advantageous for a wide range of fluorescent dyes that are used for simultaneous multicolor detection and has been successfully used to elucidate the spatial distribution of a lncRNA, *Xist*, and protein components involved in the formation of the inactive X chromosome (Cerase et al., 2014; Moindrot et al., 2015). These studies have demonstrated rather distinct distributions of polycomb complex 2 and *Xist*. SHARP, a transcription factor that has recently been shown to be essential for X inactivation (Chu et al., 2015; McHugh et al., 2015), largely overlaps with the distribution of *Xist* and provides crucial cell biological information that complements the proposed biochemical model of X chromosome inactivation (Cerase et al., 2014; Moindrot et al., 2015).

To obtain further insight into the molecular mechanism of paraspeckles, we performed fine structural analyses of these nuclear bodies using SIM. SIM observations revealed fine core-shell spheroidal structures and orderly distributions of proteins and RNA transcripts along the radially oriented *Neat1_2* transcripts. These observations reinforce the proposed sponge function of paraspeckles and exemplify the utility of super-resolution microscopy for fine structural analyses of submicron-sized nonmembranous cellular bodies.

Results

Paraspeckle components are arranged into a characteristic core-shell spheroidal structure

To gain insight into the molecular mechanism and function of paraspeckles, we examined their fine structure using SIM and compared the spatial relationship between different regions of *Neat1_2* (hereafter, *Neat1*) and paraspeckle proteins in detail. For this analysis, we used primary cultures of corpus luteal cells expressing luteal marker genes (Fig. S1), as the physiological function of paraspeckles in this cell type has been well documented in *Neat1* KO mice (Nakagawa et al., 2014). First, we performed FISH and simultaneously detected the middle and the 3′ regions of *Neat1* using probes that specifically detected each region (Fig. 1 A). The signals obtained using these probes largely overlapped when using a conventional epifluorescence microscope (Fig. 1, B and C). However, a single focus SIM observation clearly revealed a differential arrangement of the two regions of *Neat1*, with the centrally located middle region surrounded by the 3′ region located peripherally, forming a core-shell spheroidal structure (Fig. 1 C, Fig. 2 B, and Fig. S3). The characteristic core-shell organization of *Neat1* was consistent with previous electron microscopy observations (Souquere et al., 2010), indicating the validity of SIM for the observation and study of nuclear bodies. We also confirmed the core-shell structure using an inverse combination of fluorescent dyes (Fig. 1 D), suggesting that the layered organization of the two signals was not an artifact caused by the differential diffraction limits of the two different wavelengths of the light.

We then compared the distribution of three different regions of *Neat1* in various combinations to further investigate the organization of *Neat1* in a paraspeckle (Fig. 1 E and Fig. 2, A–D). To reveal the position of the transcription sites, we simultaneously detected nascent *Neat1* transcripts using a probe designed against the 3′ tail region of *Neat1* (Fig. 1 A), which produces unstable short transcripts containing a tRNA-like structure that served as a cleavage signal for *Neat1* (Sunwoo et al., 2009). We typically observed two to three dots/cells with the tail probe, suggesting that we could successfully visualize the putative transcription sites of Neat1 (Fig. S2). The FISH signals obtained with the 5′ or the 3′ regions of *Neat1* were always located surrounding the middle region of *Neat1*. However, the signals were not continuous and frequently interrupted, resulting in a dashed ring

Downloaded from on December 11, 2016

Published September 19, 2016

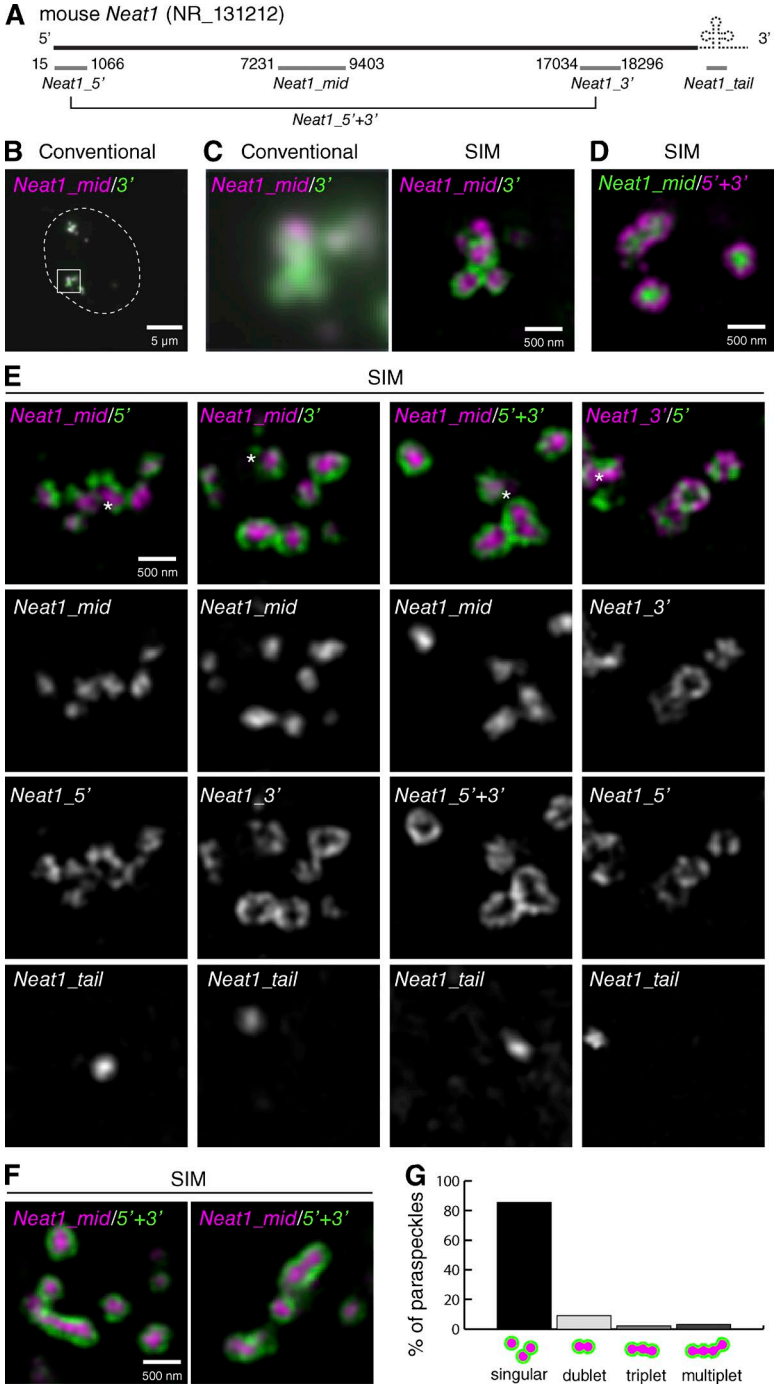

Figure 1. **Core-shell arrangement of *Neat1* in paraspeckle spheres I.** (A) Schematic diagrams of the positions of FISH probes that detect differential regions of *Neat1*. (B and C) Simultaneous detection of the middle and 3′ regions of *Neat1* using a conventional epifluorescence microscope (Conventional) and SIM. (D) The same FISH image detected with the converse combination of secondary antibodies as in C. (E) Comparisons of the differential distribution of each *Neat1* region in the paraspeckle spheres. Note that the middle region is located in the core of the paraspeckles, whereas the 5′ and the 3′ regions are located peripherally. Asterisks indicate the position of the putative transcription site detected with the *Neat1_tail* probe. (F) Paraspeckles with a sausage-like shape that were occasionally detected in the corpus luteal cells. (G) Histogram of paraspeckles with different shapes (n = 187). Bars: (B) 5 µm; (C–F) 500 nm.

(*Neat1_5′* and *Neat1_3′* in Fig. 1 E; Fig. 2, A and B; and Fig. S3). The FISH signals became fairly uninterrupted and formed a continuous circular ring when these two probes detecting the ends of *Neat1* were mixed and detected simultaneously using the same fluorescent dye (*Neat1_5′+3′* in Figs. 1 E, 2 C, and S3). This finding suggested that these two regions were separately assembled into distinct patches and not randomly mixed at the shell of the paraspeckle spheroids. When the 5′ and 3′ regions of *Neat1* were simultaneously detected using different fluorescent dyes, they made an alternate pattern along the surface of each spheroid (Figs. 1 E, 2 D, and S3). These observations suggested that the 5′ and 3′ region of *Neat1* are separately bundled together and radially arranged to form spheroidal structures. In some cases, the core-shell structure of paraspeckles was not prominent at the sites of transcription when visualized by the tail region of *Neat1*. This observation suggested that the paraspeckles were in the process of being assembled (Fig. 1 E), consistent with previous observations that paraspeckles are formed at transcription sites (Mao et al., 2011; Shevtsov and Dundr, 2011). Typically, paraspeckles were detected as separate spheroids, or aggregates of spheroids. However, they were occasionally fused to generate a long sausage-like structure (Fig. 1, F and G), as previously described (Souquere et al., 2010).

Paraspeckles contain >40 proteins that exhibit RNA-binding properties. We compared the FISH signals with the spatial distribution of seven of these proteins—Sfpq, Nono, Pspc1, Fus,

Downloaded from on December 11, 2016

Published September 19, 2016

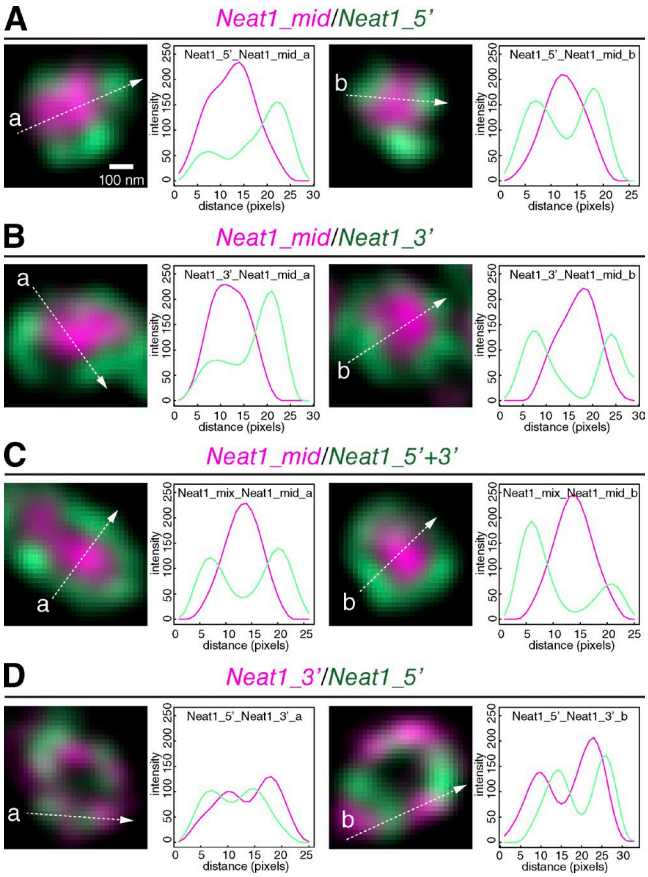

Figure 2. **Core-shell arrangement of *Neat1* in paraspeckle spheres II.** Higher magnification SIM images of two of the representative single paraspeckles stained with the *Neat1_mid* and the *Neat1_5′* probe (A), the *Neat1_mid* and the *Neat1_3′* probe (B), the *Neat1_mid* and the *Neat1_5′+3′* probe (C), and the *Neat1_3′* and the *Neat1_ 5′* probe (D). Intensity profiles along the dashed lines (a and b) are shown in the graphs. Note that the middle region of *Neat1* is centrally located, and the 5′ and the 3′ regions are distributed in a complementary manner along the shell of the paraspeckle spheres. Bar, 100 nm.

Rbm14, Brg1, and Tardbp, which were detected by immunohistochemistry after FISH (Figs. 3, 4, and S3). To optimize simultaneous detection of RNA and proteins, we omitted the proteinase K treatment that is commonly included in the FISH protocols using RNA probes, and this improved protocol well preserves the epitopes recognized by antibodies against paraspeckle proteins (Fig. S4). Notably, paraspeckle protein components can be categorized into three groups depending on their position in the paraspeckle spheroids: the core group, the patch group, and the shell group. The core group includes Sfpq, Nono, and Pspc1, all of which are members of the DBHS family of RNA-binding proteins (Dong et al., 1993; Shav-Tal and Zipori, 2002). The signals of the core group proteins largely coincided with the signals from the middle region of *Neat1*, which was surrounded by a continuous shell as revealed by the *Neat1_5′+3′* probe (Fig. 3 A; Fig. 4, A–C; and Fig. S3). Fus was also localized in the core of the spheroids, as detected by an mAb raised against the C-terminal region of the protein (Fig. 3 A; Fig. 4 D; and Fig. S3). Proteins in the second group, Rbm14 and Brg1, formed small patches that were primarily distributed in the core but also in the shell of the paraspeckle (Fig. 3 A; Fig. 4, E and F; and Fig. S3). The third group, consisting of only Tardbp, was predominantly localized at the shell of the paraspeckle. Weak but significant signals of Tardbp were also detected in the core of the paraspeckle (Figs. 3 A, Fig. 4 G; and Fig. S3). To computationally validate the arbitrary classification of the paraspeckle proteins, we used a pattern-recognition utility called wndchrm, which enabled the calculation of similarity distances between groups of images from a large (~2,700) set of features extracted from each image via a machine learning algorithm (Shamir et al., 2008). As expected, Sfpq, Nono, Pspc1, and Fus were grouped in a branch containing the middle region of *Neat1*, Rbm14 and Brg1 were closely related in a separate branch, and Tardbp was classified in a branch containing the 5′ and 3′ regions of *Neat1* (Fig. 3 B). Collectively, the SIM analyses revealed fine core-shell spheroidal structures of paraspeckles. Each paraspeckle component was distributed in a distinct position in an ordered manner (Fig. 3 C).

Fus regulates the assembly of *Neat1* ribonucleoprotein complex into large paraspeckles

Among the proteins that are essential for the formation of paraspeckles, the category Ib proteins, including Fus, are unique because the depletion of these proteins does not significantly affect the levels of *Neat1_2*, the architectural form of *Neat1* (Naganuma et al., 2012; Shelkovnikova et al., 2014). This is in sharp contrast to the depletion of category Ia protein (e.g., Sfpq or Nono), which leads to a dramatic decrease of *Neat1_2* (Naganuma et al., 2012). We thus investigated the structures formed by *Neat1_2* using mouse embryonic fibroblast (MEF) cells prepared from Fus KO mice (Hicks et al., 2000), which exhibit neonatal lethality caused by genomic instability. As previously reported (Prasanth et al., 2005; Nakagawa et al., 2011), distinct formation of paraspeckles was observed in MEFs prepared from wild-type (WT) embryos. Notably, we occasionally observed *Neat1*-positive, paraspeckle-like nuclear bodies prepared from Fus KO mice, revealed by a conventional epifluorescence microscope (Fig. 5, A and B). Thus, we investigated whether these bodies consisted of the core-shell structure we observed in the corpus luteal cells using SIM. FISH analyses using the region-specific *Neat1* probes revealed the characteristic core-shell spheroidal structures in the MEFs derived from WT mice. These structures were indistinguishable from the paraspeckles in the corpus luteal cells (Fig. 5 C). However, in the MEFs derived from Fus KO mice, *Neat1* accumulated at its putative transcription sites but never formed the core-shell structure (Fig. 5 C, asterisks indicating putative transcription sites). Instead, numerous *Neat1* FISH signals were observed throughout the nucleoplasm, and the signals for one region of *Neat1* were frequently accompanied by those of the other region (Fig. 5 C, arrowheads). These observations suggested that *Neat1* formed a primary unit, but failed to be assembled into paraspeckles, being released from the putative transcription sites in the Fus KO MEFs.

Notably, the FISH signals detected by the *Neat1_mid* probe were rarely flanked by the signals detected by the *Neat1_5′+3′* probe and instead were observed as neighboring signals (Fig. 5 D), suggesting that *Neat1* is folded in half, rather than forming a stretched rod, in these primary units (Fig. 5 D). Next, we measured the distances between each region of the *Neat1* in the primary units released from the putative transcription sites. The mean distances between the 5′–3′, the 5′–middle, and the middle–3′ were 86 ± 17, 113 ± 19, and 108 ± 21 nm,

Downloaded from on December 11, 2016

Published September 19, 2016

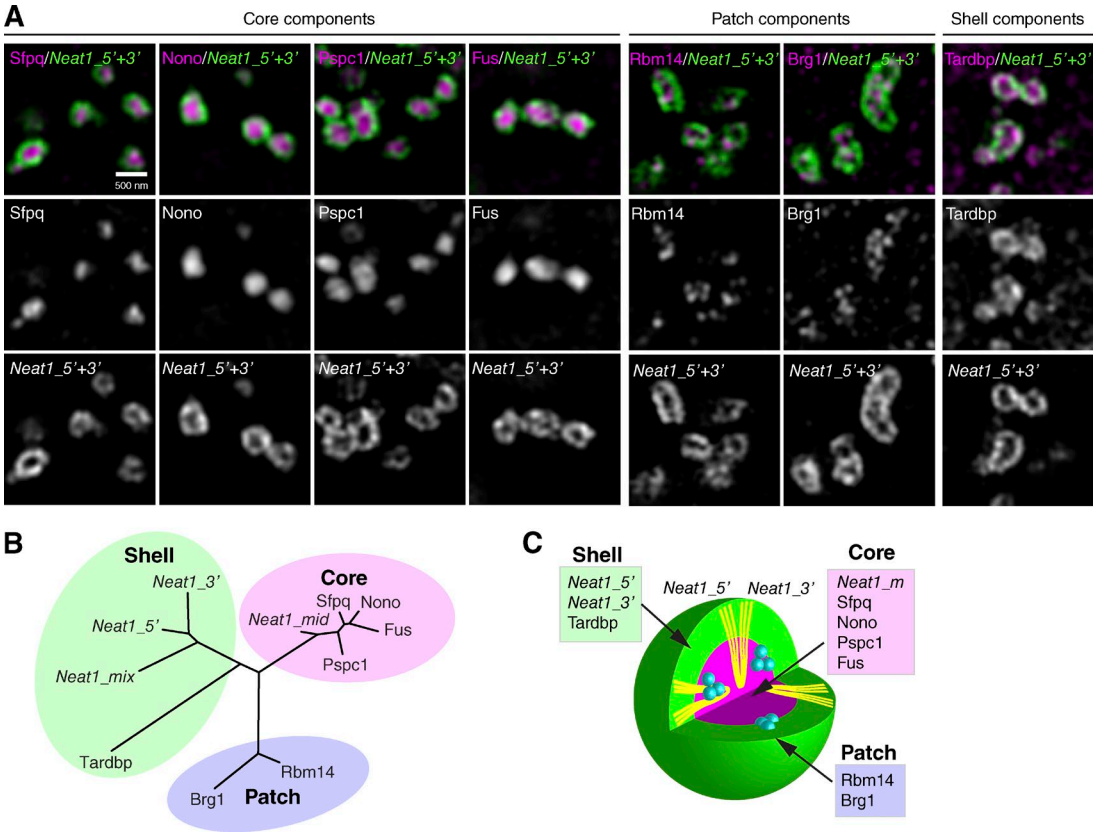

Figure 3. **Core-shell arrangement of protein components in paraspeckle spheres I.** (A) Simultaneous detection of *Neat1* and seven of the protein components of paraspeckles, including Sfpq, Nono, Pspc1, Fus, Rbm14, Brg1, and Tardbp in corpus luteal cells. Note that the paraspeckle proteins are grouped into the core, patch, and shell components depending on their distribution in the paraspeckles. (B) Dendrogram based on pairwise class-distance matrix generated using the machine-learning pattern-recognition tool wndchrm. The shell, core, and patch components are grouped into three distinct branches. (C) A model for the structure of paraspeckles. *Neat1* folds in half with the 5′ and the 3′ regions bundled independently and radially arranged to construct scaffolds of paraspeckles. Bar, 500 nm.

respectively (n = 50), whereas the mean distance between the midpoints of the spheroid rings was 353 ± 47 nm (n = 30; Fig. 5 E). These observations suggested that *Neat1* folded into a V-shape, and the 5′ and the 3′ regions were bundled separately and radially assembled into a larger spheroid by Fus (Fig. 5 F).

We next examined the localization of paraspeckle proteins in Fus KO MEFs (Fig. 6 A). The core group proteins Sfpq, Nono, and Pspc1 accumulated at the *Neat1* putative transcription site (Fig. 6 A), suggesting that they were tightly associated with *Neat1* even in the absence of Fus. Similar accumulations at putative transcription sites were also observed with Tardbp (Fig. 6 A). However, Brg1 and Rbm14, comprising the patch components, were not enriched at the putative transcription site (Fig. 6 A), suggesting that Fus stabilized the interaction of these proteins with nascent *Neat1* transcripts during the formation of paraspeckle spheres.

To confirm the function of Fus is conserved in human cells, we examined the organization of NEAT1 and NONO in paraspeckles using HAP1 cells that lack the expression of FUS. Similar to MEFs, the middle region of NEAT1 or NONO was located in the core of the paraspeckle spheres, surrounded by the 5′ and 3′ regions of NEAT1 located in the shell (Fig. 6 B). The core-shell structure was disrupted in the HAP1 cells deleted with FUS (ΔFUS; Fig. 6 B), suggesting that human FUS is also required for the highly ordered fine structure of paraspeckles.

To gain more insight into the properties of Fus, we reintroduced full-length or mutant forms of FUS (Fig. 5, C and D) into MEFs derived from Fus KO mice. As expected, full-length FUS restored the core-shell structure of paraspeckles, whereas this effect was not observed with mutant molecules that lack N′-located prion-like domain (PrLD) of FUS (NΔ; Fig. 6 E), which was consistent with a previous finding that PrLD is essential for the paraspeckle formation (Shelkovnikova et al., 2014). We also found that the C′ located RNA binding region including the arginine (R)-glycine-glycine domains and RNA recognition motifs is also required for the assembly of Neat1 RNPs into paraspeckle spheroids (CΔ; Fig. 5 E).

We then examined the localization of this protein using another antibody that specifically recognizes epitopes in the PrLD of Fus at the N-terminal region of this protein (Fig. 6, F and G). Interestingly, the signals obtained using this antibody largely differed from the signals detected using the mAb recognizing the C-terminal region of Fus. The N-terminal signals were observed as discrete dots distributed within and around the areas revealed by the antibody recognizing the C terminus of Fus (Fig. 6 C). These observations suggested that the N-terminal regions of Fus were pinned into small areas surrounded by the C-terminal regions of this protein. Alternatively, the access of the antibody to the epitope located in the N-terminal PrLD was prevented by the formation of a hydrogel, which has been proposed to play essential roles in the formation of RNA-containing nuclear bodies (Han et al., 2012; Kato et al., 2012).

Downloaded from on December 11, 2016

Published September 19, 2016

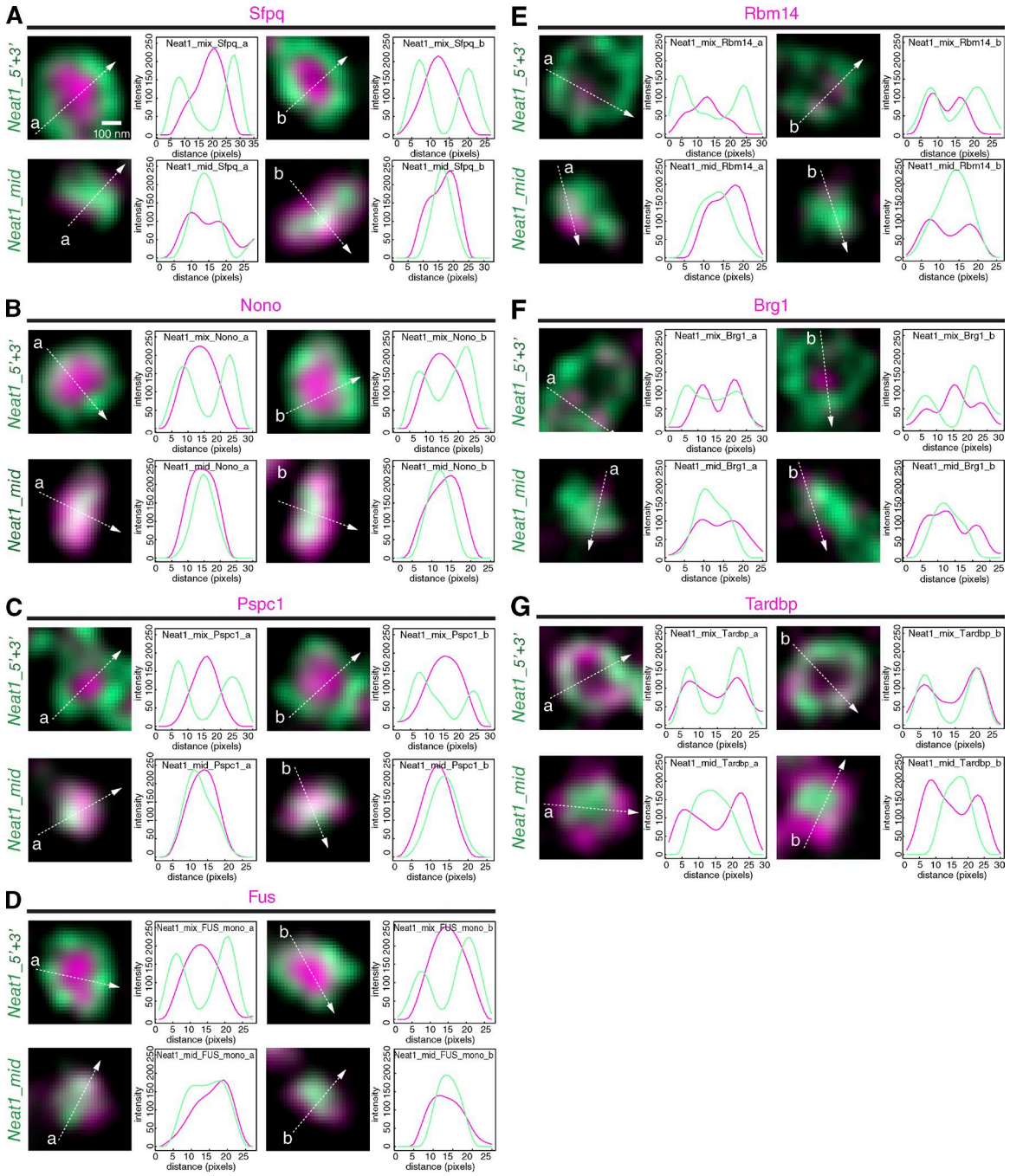

Figure 4. **Core-shell arrangement of protein components in paraspeckle spheres II.** Higher magnification SIM images of two of the representative single paraspeckles stained with the *Neat1 5′+3′* probe and Sfpq (A), Nono (B), Pspc1 (C), Fus (D), Rbm14 (E), Brg1 (F), and Tardbp (G). Intensity profiles along the dashed lines (a and b) are shown in the graphs next to the images. Bar, 100 nm.

Miscellaneous AG-rich RNAs accumulate on the surface of the paraspeckle sphere

Although the protein components of paraspeckles are well-characterized, limited information is available regarding their RNA components. To systematically identify RNA molecules that associated with paraspeckles, we purified the RNP complexes of *Neat1* using a method termed capture hybridization analysis of RNA targets (CHART; Fig. 7 A), which was originally developed to identify the genomic binding sites of particular lncRNAs using antisense oligonucleotides designed against particular lncRNAs (Simon et al., 2011, 2013; West et al., 2014). We used oligonucleotides designed against the 5′ region of *Neat1* (Fig. 7 B) because this region was the most sensitive to RNaseH digestion upon addition of the antisense oligonucleotides and was therefore expected to be accessible during the CHART purification (West et al., 2014). The *Neat1* RNPs were purified from primary cultures of corpus luteal cells using two different sets of antisense oligonucleotides. The co-purified RNAs were subsequently analyzed using a massively parallel sequencing (RNA sequencing) (Fig. 7 A). The 5′ region of *Neat1* was predominantly enriched by CHART purification (Fig. 7 B), suggesting that partial RNP fragments and not the

Downloaded from on December 11, 2016

Published September 19, 2016

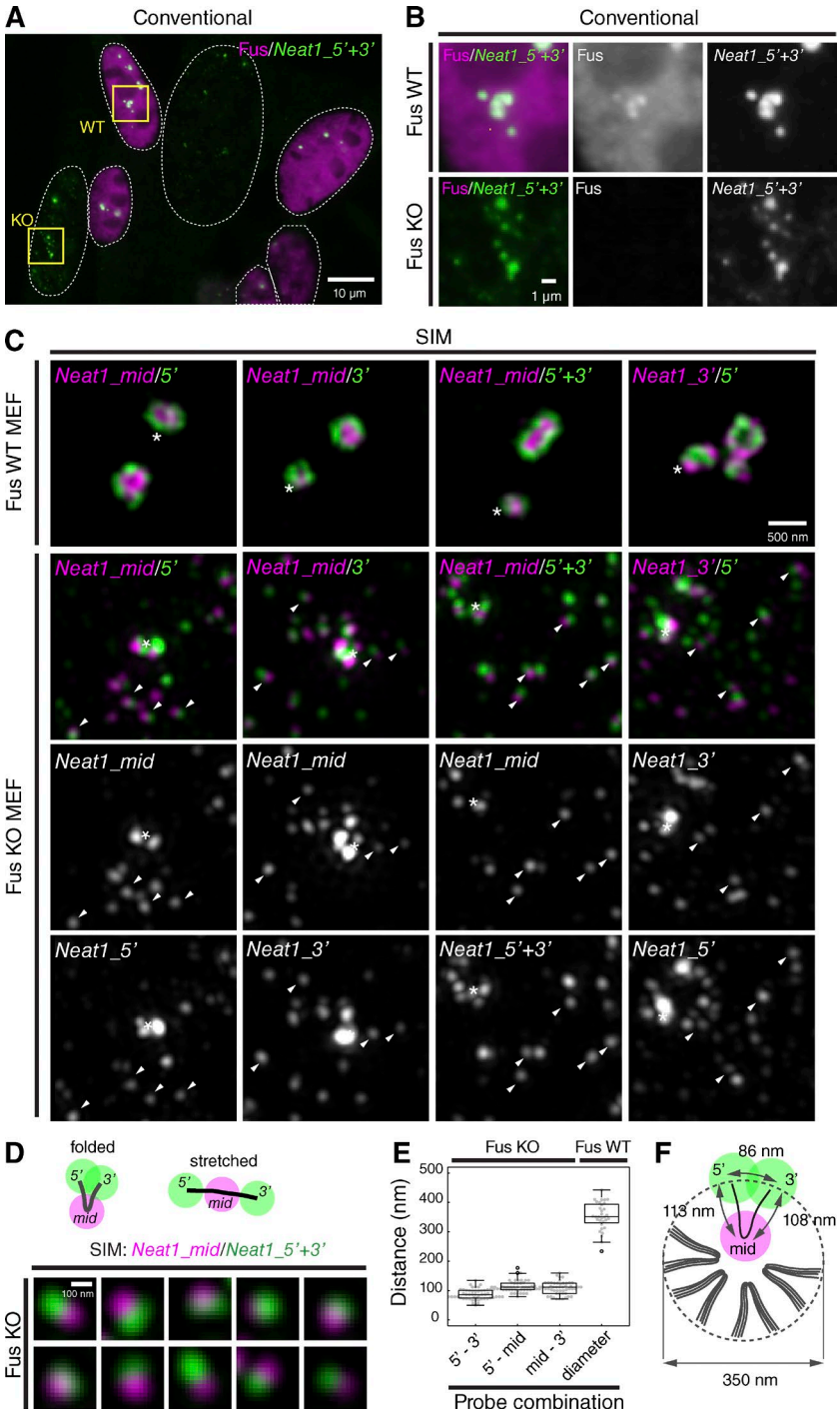

Figure 5. **Fus is required for the assembly of the core-shell structure of paraspeckles.** (A) Simultaneous detection of Fus and *Neat1* in a mixture of MEFs prepared from WT and Fus KO mice using a conventional epifluorescence microscope (Conventional). Yellow boxes indicate the areas shown at a higher magnification in B. Note that *Neat1* forms discrete nuclear body-like structures in Fus KO MEFs. (B) Higher magnification images shown in yellow boxes in A. (C) SIM observation of *Neat1* in WT and Fus KO MEFs using the region-specific probes. Note that the characteristic core-shell structure was not observed in the Fus KO MEFs. Asterisks indicate the position of the putative transcription sites detected with the *Neat1* tail probe. Arrowheads indicate *Neat1* primary units containing both of the detected regions of *Neat1*. (D) Models of the *Neat1* primary unit and higher magnification images of FISH signals obtained with the *Neat1_5′+3′* probe in Fus KO MEFs. Note the close association of the two signals. (E) Measurement of the distance between the two different regions of *Neat1* in Fus KO MEFs. Note that the distance between the 5′ and 3′ region of *Neat1* is shorter than the distance between the 5′ and the middle or the middle and the 3′ regions of *Neat1*. The median is indicated with a horizontal line in a box that represents the first and third quartiles. Outliers are indicated as circles, and the maximum and minimums are indicated at the end of the whiskers. Each dot represents each signal of *Neat1* particle. (F) A model of the organization of *Neat1* in paraspeckle spheres. Bars: (A) 10 µm; (B) 1 µm; (C) 500 nm; (D) 100 nm.

entire paraspeckle were recovered using this method. Interestingly, the 3′ region of *Neat1* was also enriched to some extent by the CHART purification (Fig. 7 B), which was consistent with the aforementioned observation that the 5′ and the 3′ regions of *Neat1* constitute the shell of the paraspeckle (Figs. 3 C and 5 F). We subsequently selected candidate RNA transcripts that were copurified with both of the two different sets of antisense oligonucleotide conditions (Fig. 7, C and D; Fig. S5; and Tables S1 and S2) to avoid artificial purification of specific RNA molecules via direct binding of the oligonucleotide to complementary sequences regardless of paraspeckles. These analyses revealed that two different types of RNA transcripts—spliced mRNAs, such as Trim44 and Numa1, and specific introns of particular genes, such as the third intron of Actr3 and the first intron of Prss35—copurified with *Neat1* (Fig. 7, C and D; and Fig. S5). Interestingly, Multiple Em for Motif Elicitation (MEME) analyses revealed that all CHART-enriched RNAs contained AG-rich sequence motifs that were arranged in tandem (Fig. 7, E and F). No other features, including exon–intron organization or chromosomal positions, were shared between the CHART-enriched RNAs. To confirm the paraspeckle localization of the first intron of Prss35, one of these candidate paraspeckle-enriched RNAs in corpus luteal cells, we performed FISH using probes designed to detect the sequences located outside the AG-rich sequence motifs. Conventional microscopic observation revealed that subpopulation of the Prss35 signals was overlapped with Neat1

Downloaded from on December 11, 2016

Published September 19, 2016

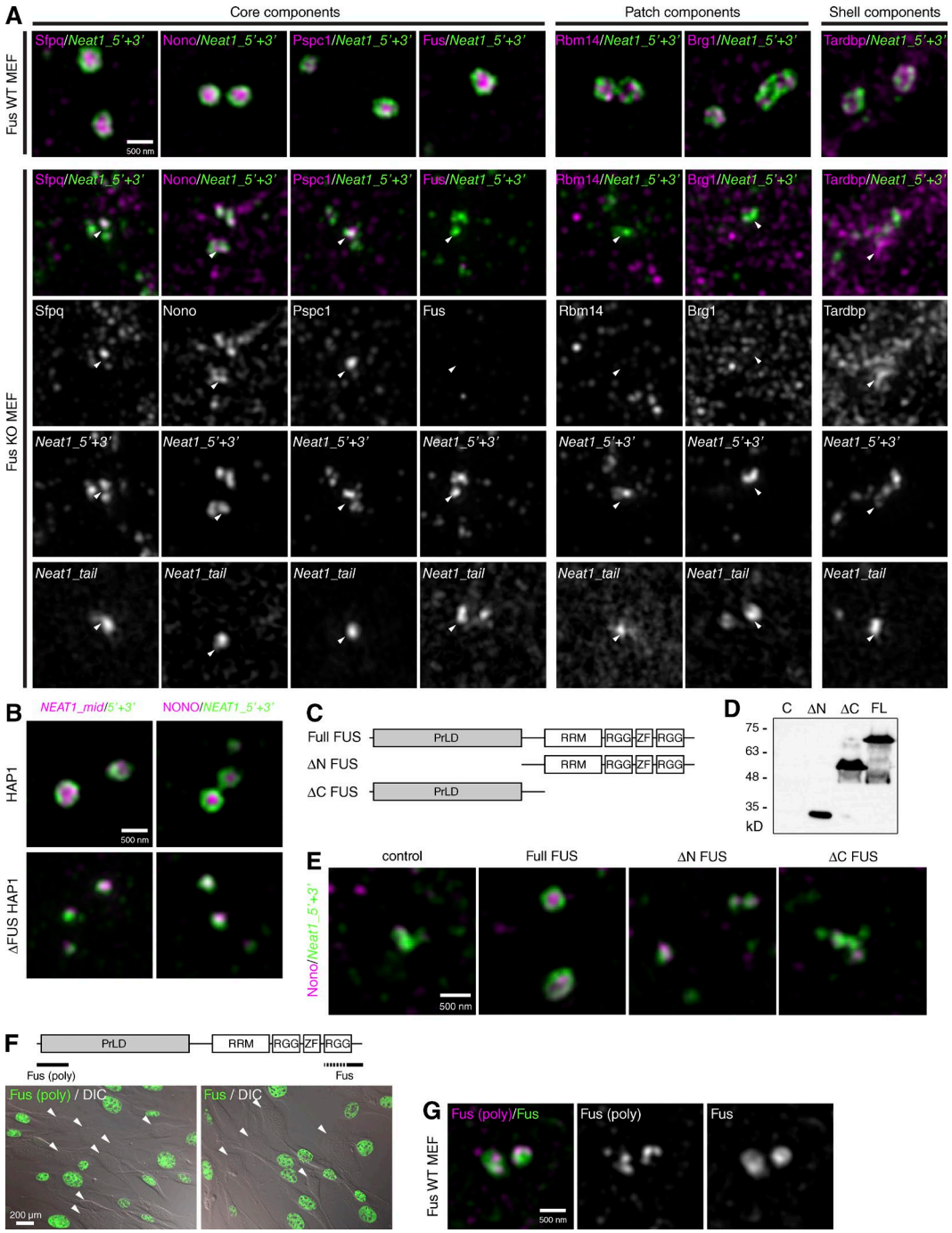

Figure 6. **Fus-independent and dependent recruitment of paraspeckle proteins.** (A) Simultaneous detection of *Neat1* and seven of the protein components of paraspeckles, including Sfpq, Nono, Pspc1, Fus, Rbm14, Brg1, and Tardbp, in MEFs derived from WT and Fus KO mice. Note that DBHS family proteins (Sfpq, Nono, and Pspc1) and Tardbp, but not Rbm14 and Brg1, are recruited to the putative transcription site in the absence of Fus. Arrowheads indicate paraspeckle-like nuclear bodies formed at the putative *Neat1* transcription site in Fus KO MEFs. (B) Simultaneous detection of various forms of NEAT1 and NONO in HAP1 cells and FUS-deleted HAP1 cells (ΔFUS HAP1). Probes used to detect NEAT1 are shown in the top boxes. (C) Schematic drawing of full-length and mutant FUS protein exogenously expressed by lentiviruses. ΔN FUS lack the PrLD and ΔC FUS lack the RNA binding domains including RNA recognition motifs (RRM) and arginine (R)-glycine-glycine domain (RGG) as well as zinc finger domain (ZF). (D) Western blot analyses of lysate from the cells infected with control EGFP (C), full-length FUS (FL), ΔN FUS (ΔN), and ΔC FUS (ΔC). Note that migration of FL and ΔN are much slower than predicted molecular mass (57 and 35 kD, respectively), probably because of the presence of PrLD in these molecules. (E) Simultaneous detection of *Neat1 5′+3′* and Nono in Fus KO MEFs expressing various forms of FUS protein. Note that the core-shell structure of paraspeckles was rescued with FL FUS, but not with mutant molecules that lack either PrLD or RNA binding domains. (F) Confirmation of the specificity of polyclonal [Fus (poly)] and monoclonal (Fus) antibodies against Fus. Mixtures of MEFs derived from WT and KO mice of Fus were stained with each antibody. Note the complete absence of signals in the Fus KO MEFs (arrowheads). The positions of the epitope of these antibodies are shown in the schematic drawing of the domain structure of Fus. (G) Simultaneous detection of Fus using polyclonal antibodies and mAbs that recognize the N- and C-terminal region of the protein, respectively. Bars: (A, B, E, and G) 500 nm; (F) 200 μm.

Downloaded from on December 11, 2016

Published September 19, 2016

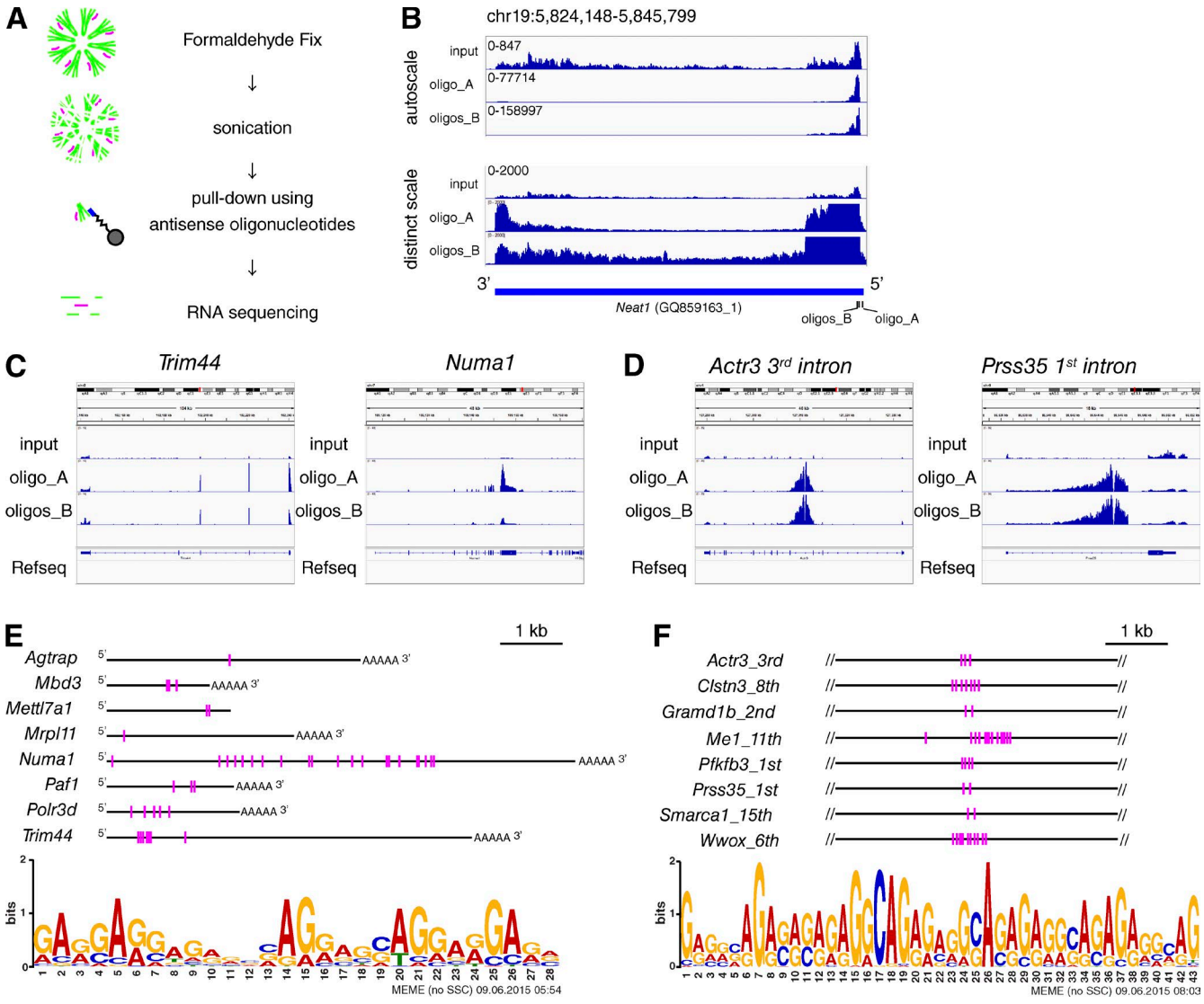

Figure 7. **Identification of novel paraspeckle RNA components by CHART RNAseq.** (A) Schematic cartoon showing the CHART purification of paraspeckle fragments. The *Neat1* complexes were purified using antisense oligonucleotides, and copurified RNAs were analyzed by RNAseq. (B) Schematic of the *Neat1* locus showing the position of the oligonucleotide sets (oligos_A and oligos_B) used for the CHART purification and mapped reads of the input and CHART-purified RNAs. The scales are automatically adjusted in the top panel and adjusted to a distinct value (0–2,000) in the bottom panel. Note that the 5′ region of *Neat1* is predominantly enriched by the CHART purification, whereas the 3′ region is also moderately enriched by both oligonucleotide sets. Mapping of CHART-enriched RNAseq reads at the genomic loci of *Trim44, Numa1, Actr3,* and *Prss35.* Note that the reads are mapped to exons in *Trim44* and *Numa1* (C), whereas they are mapped to the third and the first intron of *Actr3* and *Prss35,* respectively (D). MEME-identified AG-rich sequence motifs and their distribution along the exon-enriched (E) and intron-enriched (F) genes. Partial regions of each intron containing the AG-rich motifs are shown in F. Bar, 500 nm.

(Fig. 8, A and B, left), suggesting that they were indeed enriched in the paraspeckles. In contrast, the signals obtained with probes that detect exons of Prss35 did not coincide with Neat1 (Fig. 8, A and B, right), suggesting that spliced introns, but not pre-mRNAs, were enriched in the paraspeckles. Subsequently, we analyzed the distribution of the AG-rich motif-containing transcripts in corpus luteal cells using SIM. Notably, all of the AG-rich RNAs localized at the shell of the paraspeckle spheres, as revealed by the *Neat1_5′+3′* probe (Fig. 8, C and D; and Fig. S5). The signals of the AG-rich RNAs were discontinuous, observed as dots and aligned along the surface of the paraspeckles (Fig. 8, C and D; and Fig. S5). As shown in Fig. 8 A, we noticed only a subpopulation of AG-rich RNAs was colocalized to paraspeckles, suggesting that the paraspeckles did not entirely sequester these target RNAs but rather trapped them when they were encountered in the nucleoplasm. Consistent with this hypothesis, we could not detect significant changes in the amount of CHART-enriched AG-rich RNAs in the nuclear or cytoplasmic fractions of the corpus luteal cells prepared from *Neat1* KO mice (Fig. 8 E).

Discussion

We have demonstrated that paraspeckles consist of core-shell structures in which protein and RNA components are regularly arranged in characteristic spheroidal structures. These findings represent a significant extension to previous electron microscopy observations (Souquere et al., 2010). We newly found that: (a) paraspeckles consist of stretches or aggregates of spheroids

Downloaded from on December 11, 2016

Published September 19, 2016

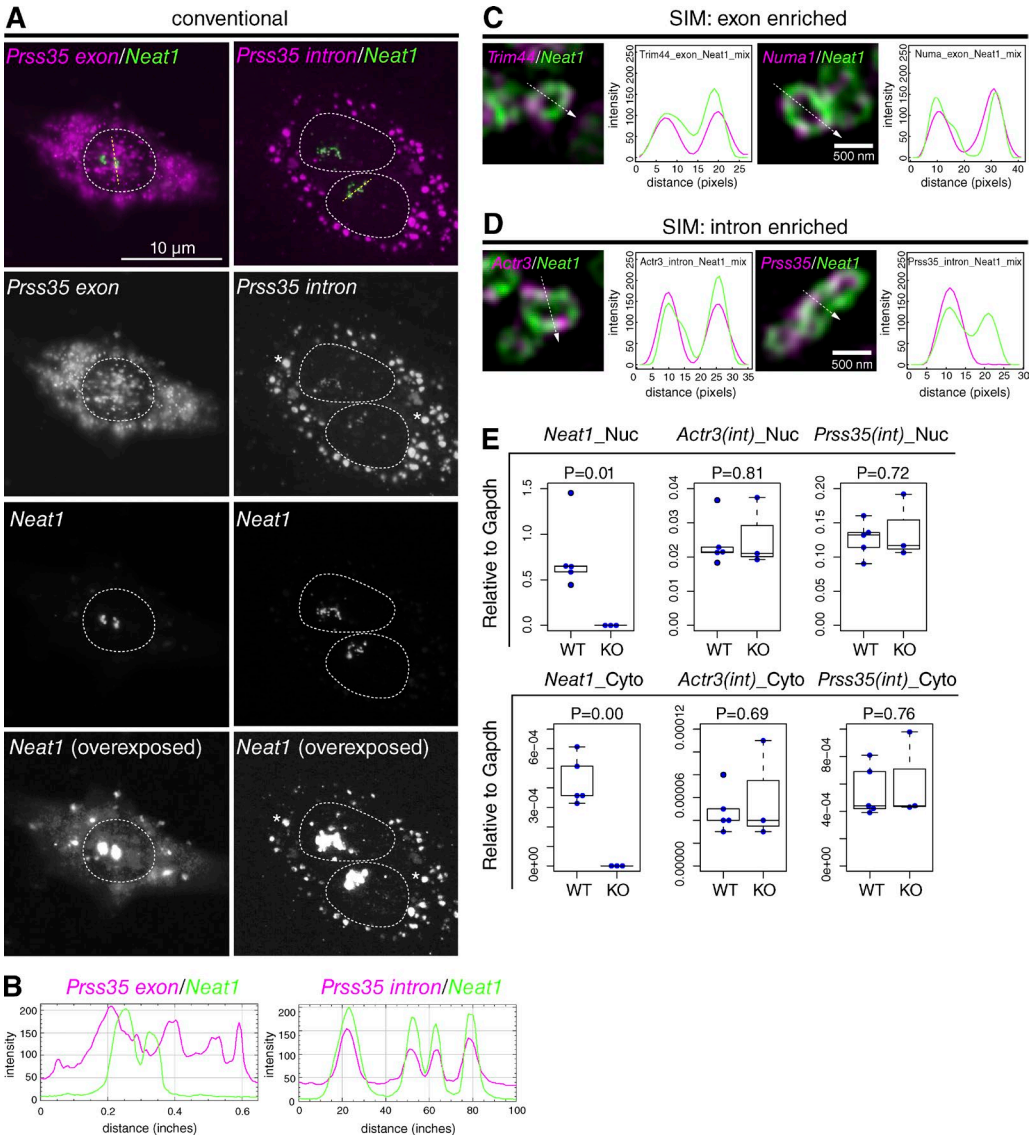

Figure 8. **FISH analyses of the localization of paraspeckle-enriched AG-rich RNAs.** (A and B) Simultaneous detection of Neat1 and the exons and the first intron of *Prss35* in corpus luteal cells using confocal microscopy. A single optical section image is shown. Note that subpopulation of *Prss35* intron signals colocalized with Neat1-positive paraspeckles, whereas exon signals were mostly observed in the cytoplasm and did not coincide with the Neat1 signals. Intensity profiles along the yellow dashed line are shown in B. Dashed white curving lines indicate position of the nucleus. Note that the bright round signals in the cytoplasm are derived from autofluorescence of lipid droplets, some of which are shown by asterisks and are clearly identifiable in a different channel overexposed for Neat1 signals. Bar, 10 µm. (C and D) Simultaneous detection of AG-rich RNA and *Neat1_5′+3′* in corpus luteal cells using SIM. Intensity profiles along the dashed line are shown in the graphs adjacent to the images. Bars, 500 nm. (E) Box and whisker plots showing the expression of *Neat1* and AG-rich RNAs in the cytoplasm (cyto) or nucleus (nuc) of corpus luteal cells from WT and *Neat1* KO mice. The median is indicated with a horizontal line in a box that represents the first and the third quartiles. Outliers are indicated as circles, and the maximum and the minimums are indicated at the end of the whiskers. Each blue dot represents a sample from an individual mouse.

that occasionally fuse to form sausage-like structures; (b) the 5′ and 3′ regions of *Neat1* are distinct in paraspeckle shells, suggesting bundles of Neat1 RNAs; and (c) paraspeckle spheroid cores consisting of DBHS family proteins are separated from the nucleoplasm by paraspeckle shells containing the 5′ and 3′ regions of *Neat1* and Tardbp, occasionally bridged by patches of Rbm14 and Brg1. Importantly, paraspeckles have been proposed to function as a molecular sponge to indirectly regulate target gene expression; this is achieved by sequestering Sfpq that serves as a negative or positive regulator of transcription in different contexts (Hirose et al., 2014; Imamura et al., 2014). In this study, the observed internal localization of Sfpq in paraspeckle spheres is consistent with the proposed sponge function of paraspeckles. In addition, proteins that are essential for the structural maintenance of paraspeckles (categories Ia and Ib in Naganuma et al., 2012) are localized to the core or patches, but not the shell of the paraspeckles, suggesting that the former components play architectural roles, whereas the shell components associates with nucleoplasmic components to fulfill their function. Given that the core-shell structure of the paraspeckle sphere is functionally important, the ordered structure might be used as a marker of functional paraspeckles. We observed a disrupted organization of *Neat1* in the MEFs derived from the Fus KO mice. However, the Neat1 transcripts were observed to form aggregates containing DBHS family proteins. It would be interesting to determine whether paraspeckles could

Downloaded from on December 11, 2016

Published September 19, 2016

preserve the characteristic core-shell structures in certain abnormal conditions, such as in cancerous cells.

Recent advances in sequencing technology, such as high-throughput sequencing of RNA isolated by cross-linking immunoprecipitation (CLIP), has enabled the identification of RNA sequences that associate with particular proteins of interest. The genome-wide CLIP-sequencing (CLIP-seq) data for certain paraspeckle proteins are available in public databases, including those for Tardbp and Fus. Interestingly, the highest peaks for the Tardbp binding sites have been found at the 5′ and 3′ regions of *Neat1* (Polymenidou et al., 2011; Tollervey et al., 2011), consistent with the strong Tardbp signals in paraspeckle sphere shells observed by SIM. In the case of Fus, the CLIP-seq signals have been rather uniformly observed throughout the *Neat1_1* transcript with a bias for the 5′ region of *Neat1* (Hoell et al., 2011; Lagier-Tourenne et al., 2012; Rogelj et al., 2012). However, these data were obtained from brain tissues, which do not express high levels of *Neat1_2* and thus lack paraspeckles (Nakagawa et al., 2011). Because Fus is recruited to RNA polymerase II during transcription (Schwartz et al., 2012), these CLIP-seq reads may have been derived from nascent *Neat1* in the high-throughput sequencing of RNA isolated by CLIP analyses. Alternatively, based on our recent observations that *Neat1* is extremely insoluble even in highly denaturing solution used for RNA extraction, such as TRIzol (Thermo Fisher Scientific; unpublished data), the Fus-bound *Neat1* transcripts embedded in the core of the paraspeckles may not be solubilized in the CLIP buffer and thus are not represented in the CLIP data. Regardless of the mechanism, it would be informative to compare the results obtained using biochemical approaches, such as CLIP-seq, with the spatial information obtained by SIM observations to validate our model in the same cell type.

Although we have found that miscellaneous AG-rich transcripts associate with paraspeckles, their physiological significance (e.g., their architectural role in the formation of paraspeckle spheres) remains unknown. We have confirmed the paraspeckle localization of the 16 highest CHART-enriched AG-rich transcripts, all of which, excluding Mrpl11, are transcribed in trans from chromosomes without *Neat1*. Because paraspeckles are constructed at the transcription site of *Neat1*, the AG-rich RNAs transcribed from genomic loci distinct from the *Neat1* locus are not likely involved in the active formation of paraspeckles. We also failed to detect any significant changes in the subcellular distribution of the AG-rich transcripts in *Neat1* KO corpus luteal cells that lack paraspeckles, at least under normal culture conditions. It would be intriguing to examine whether certain environmental stresses affect the fate of AG-rich transcripts in a manner dependent on the formation of paraspeckles.

Previously, paraspeckles have been proposed to be enriched in hyper A-to-I–edited transcripts containing inverted repeat insertions (Prasanth et al., 2005; Clemson et al., 2009). However, these transcripts were not enriched by our CHART purification, and indeed are not highly expressed in these murine corpus luteal cells (unpublished data). Because we designed the oligonucleotides against the 5′ regions of *Neat1*, it is also possible that the hyper A-to-I edited RNA was associated with the core of the paraspeckles but not with the shell component of the paraspeckles (including the 5′ region of *Neat1*). Indeed, the 5′ region of *Neat1* was predominantly enriched by the CHART purification, whereas the central region of *Neat1* was rarely recovered. CHART purification using different antisense oligonucleotides designed against various regions of *Neat1* would likely further reveal the subdomain organization of paraspeckles, similarly to the elucidation of the module structure of *roX1* and *roX2* using ChIRP, a comparable technique (Quinn et al., 2014). To further clarify these points, future studies should be designed to develop new methods to isolate entire paraspeckles and not partial fragments of *Neat1*.

Several nonmembranous cellular bodies were initially described by electron microscopy and have been subsequently confirmed by the localization of specific proteins or nucleic acids (Spector, 2006). The sizes of these cellular bodies are typically at submicron levels, and hence it is difficult to investigate fine internal structures using conventional light microscopy. The emergence of super-resolution microscopy has enabled the rapid observations of internal structures of cellular bodies by simultaneous detection and comparison of the signals of each component within the bodies. Recently, super-resolution observations of nuclear speckles have revealed ordered internal structures containing an lncRNA, *Malat1*, and Srsf1 protein (Prasanth, K.V., personal communication), further confirming the usefulness of this technology for fine structural analyses of cellular bodies. Notably, many of these cellular bodies contain specific sets of RNA molecules (Spector, 2006). Because the visualization of different regions of RNA molecules is feasible using region-specific FISH probes, the combination of FISH detection and super-resolution microscopy will provide an extremely useful tool for the structural analyses of cellular bodies as long as certain RNA molecules are regularly arranged, as is the case for paraspeckles. These techniques can be applied for the observation of other RNA-containing bodies, including polycomb bodies, Cajal bodies, P-bodies, and Nuage in germ cells.

Materials and methods

All experiments using animals and recombinant DNA were approved by the safety division of RIKEN. Nucleotide sequences for primers and oligonucleotides are shown in Table S3.

Cell culture

To prepare a primary culture of corpus luteal cells, female mice between the ages of 3 to 4 wk were injected with 5 IU PMSG. The ovaries of mice sacrificed by cervical dislocation were dissected 48 h after injection. The granulosa cells were recovered by squeezing the ovaries through a cell strainer (100-µm mesh size; Falcon; Corning) using the plunger of a 1-ml syringe in culture medium (1:1 mixture of DMEM and Ham's F12 supplemented with penicillin/streptomycin, 10% FBS, and B27; Gibco) seeded onto a 12-well plate. We typically plated granulosa cells from one individual mouse (two ovaries) into 4 wells of the 12-well plate. After 48 h, forskolin was added at a concentration of 10 µg/ml to induce the differentiation of corpus luteal cells. The culture was maintained for 48 h before fixation. MEFs were prepared from WT or Fus KO embryos (embryonic day 14.5; Hicks et al., 2000) and cultured in a 1:1 mixture of DMEM and Ham's F12 supplemented with penicillin/streptomycin. HAP1 cells and mutant HAP1 cells that lack the expression of FUS are obtained from Horizon Genomics, and they were cultured in Iscove's Modified Dulbecco's Medium in the presence of 10% FCS and penicillin/streptomycin.

Generation of FUS-expressing retroviruses

Retrovirus vectors expressing full-length or mutant FUS molecules that lack the PrLD or RNA-binding regions were generated using ViraPower Lentiviral Expression System (Invitrogen) according to the

Downloaded from on December 11, 2016

Published September 19, 2016

manufacturer's instructions. In brief, full-length FUS with N-terminal–tagged FLAG sequences were amplified by PCR from plasmid vector containing FLAG-FUS and subcloned into pENTR (Invitrogen) to generate pENTR-FUS. Mutant molecules were generated using primer sequences that are designed to delete 3–267 and 285–500 of FUS to generate pENTR ΔN FUS and pENTR ΔC FUS, respectively. After cloning into pLenti6/V5-DEST, the expression vector together with helper plasmids were transfected into 293 cells using Fugene (Promega). Culture supernatants containing the virus were collected 72 h after the transfection. To infect MEF cells, cells (0.5×10^4) were cultured in 0.5 ml undiluted virus solution for 9 h and further cultured for 48 h in a fresh culture medium. Basically, all of the cells expressed the tagged FUS proteins under this condition.

FISH

FISH was performed as previously described (Mito et al., 2016). In brief, 0.17-mm–thick coverslips were washed in detergent using an ultrasonic washer and coated with 0.5 mg/ml poly-L-lysine overnight at 4°C. After washing three times with distilled water, the coverslips were coated with 0.1% gelatin for 5 min at room temperature, washed once with distilled water, and then placed into 12-well plates before seeding with cells. Cells on the coverslips were fixed in 4% PFA in a Ca^{2+}- and Mg^{2+}-free saline buffered with Hepes (HCMF; pH 7.4) at room temperature for 10 min, washed twice with PBS, and permeabilized in 0.1% Triton X-100 (35501-15; Nacalai Tesque) in PBS for 10 min. After washing with PBS, the cells were incubated in a prehybridization buffer for 2 h and hybridized with digoxigenin (DIG), FITC, or biotin-labeled RNA probes diluted in hybridization buffer at 5–10 µg/ml overnight at 55°C. After hybridization, the cells were washed twice with 55% formamide/2× SSC for 30 min, treated with 1 µg/ml RNaseA in buffer (500 mM NaCl, 10 mM Tris [pH 8], and 1 mM EDTA) for 1 h at 37°C, washed twice with 2× SSC at 55°C for 30 min, and washed twice with 0.2× SSC at 55°C for 30 min. The hybridized probes were immunohistochemically detected using primary antibodies against DIG (anti-DIG mouse monoclonal [21H8] antibody; 420; Abcam), FITC (anti-FITC rabbit polyclonal antibody; ab19491; Abcam), and secondary antibodies (Cy3-conjugated anti-mouse IgG, AP124C; Merck Millipore; and Cy2-conjugated anti-mouse IgG, ab6944; Abcam). Biotin-labeled probes were directly detected using Cy5-labeled streptavidin (PA45001; GE Healthcare). For the simultaneous detection of paraspeckle proteins, the following antibodies were used: mouse mAb against Sfpq (clone B92; Abcam), mouse mAb against Nono (Souquere et al., 2010), mouse mAb against Pspc1 (clone 1L4; Sigma-Aldrich), mouse mAb against Fus (clone 4H11; Santa Cruz Biotechnology, Inc.), rabbit polyclonal antibody against Fus (ab84078; Abcam), rabbit polyclonal antibody against Brg1 (A300-813A; Bethyl Laboratories, Inc.), rabbit polyclonal antibody against Tardbp (10782-2-AP; Proteintech), and rabbit polyclonal antibody against Rbm14 (A300-311A; Bethyl Laboratories, Inc.). The stained samples were postfixed in 4% PFA in HCMF for 10 min at room temperature, washed with PBS, and mounted in 97% 2,2'-thiodiethanol containing 2% 1,4-diazabicyclo[2.2.2]octane. For the calibration of multicolor signals, TetraSpeck beads (T7280; Thermo Fisher Scientific) were added at a ratio of 1:100 in the mounting medium.

SIM observations and image processing

The SIM images were obtained using an Elyra system with 100× objective lens (NA 1.46; ZEISS) as previously described (Mito et al., 2016). To observe paraspeckles, 20 of the Z-series images were obtained at 100-nm intervals, and the SIM images were calculated using default settings with theoretically predicted point spread function parameters. To align the multicolor images, an alignment file was generated for each sample (e.g., each glass slide). The SIM images were discarded when one of the channels of the multicolor images was obviously shifted in one direction even after channel alignment. To classify the *Neat1* signals, 20 equivalently sized (30 × 30 pixels) paraspeckle sphere images were cropped from single-focus Z sections and analyzed using wndchrm. A distance tree was drawn according to the similarity distance matrix calculated by wndchrm.

CHART purification and RNA sequencing

3% formaldehyde cross-linked and sonicated nuclear extracts were prepared as previously described (Simon et al., 2011; Davis and West, 2015). The extracts were then incubated with *Neat1* or control capture oligonucleotide cocktails and hybridized overnight. The hybridized material was captured with magnetic streptavidin resin (Invitrogen). Bound materials were washed and eluted with RNase H (New England Biolabs, Inc.) as previously described (West et al., 2014; Davis and West, 2015). To prepare RNA from the CHART-enriched material, two consecutive phenol/chloroform/isoamyl alcohol washes followed by two chloroform/isoamyl alcohol rinses were performed. Subsequently, RNA was ethanol precipitated and resuspended in 100 µl water. RNA was further rinsed and concentrated using an RNA Clean and Concentration kit according to the manufacturer's instructions (Zymo Research). Resuspended RNA was subsequently used for downstream analyses. After purification with RNA Clean XP (Beckman Coulter), they were quantified with Qubit RNA HS Assay kit (Thermo Fisher Scientific) on a Qubit Fluorometer (Thermo Fisher Scientific). These enriched RNAs (22.2 ng for each, based on the measurement with Qubit), together with nonenriched input RNA (222 ng), were individually subject to library preparation with the TruSeq RNA Sample Prep kit v2 (Illumina). Library preparation was processed following the manufacturer's instruction until adapter ligation, except that the initial step for poly-A selection was skipped. After the adapter ligation, the optimal number of PCR cycles for each library was estimated using an aliquot (3 µl) of the product from the previous step with the Real-Time Library Amplification kit (Kapa Biosystems, Inc.). The rest of the adapter-ligated DNA was amplified with seven PCR cycles for the enriched samples and five cycles for the input sample. The amplification products were sequenced in a single lane on a HiSeq 1000 (Illumina) in the High Output mode with the proportion of 1:1:2 in molar quantity for oligo_A–enriched, oligo_B–enriched, and input samples, respectively. The sequencing was performed using TruSeq SR Cluster kit v3-cBot-HS (Illumina) and TruSeq SBS kit v3-HS (50 cycle; Illumina) with 51 SBS cycles to produce single reads. Image analysis and base calling were processed with the standard Illumina software consisting of HiSeq Control Software version 1.5.15.1 and Real-Time Analysis version 1.13.48.

Data analyses of CHART RNA sequencing

Low-quality reads were removed using FASTQ Quality Filter (80% of bases are above quality 25), and unique reads were mapped onto the mouse genome assembly mm9 using TopHat version 2.0.4 using GTF and bowtie index files downloaded from iGenome (http://support.illumina.com/sequencing/sequencing_software/igenome.html). The BED file for intron sequences were obtained using the table browser of the University of California Santa Cruz Genome Bioinformatics site (http://genome.ucsc.edu/index.html). The read counts were calculated using Cufflinks for Refseq genes and HTseq for intron regions. To select candidate introns, genes were filtered by the number of read counts (>2,000) and the fold change compared with input sample (>2.5), for both oligo_A- and oligo_B-purified samples. Among 69 genes that satisfied these criteria, 8 genes were arbitrarily selected and used for subsequent FISH analyses (Table S1, sheet Selected candidates). To

Downloaded from on December 11, 2016

Published September 19, 2016

select candidate Refseq genes, genes were filtered by the fold change compared with input samples (>9.9), and 8 genes were randomly selected for subsequent FISH analyses among the top 100 genes that were most highly enriched (Table S2, sheet Selected mRNAs). To identify enriched motifs, MEME analyses (http://meme-suite.org/tools/meme) were performed using full-length cDNA sequences for exon-enriched genes and sequences of each intron containing the peak of the mapped reads for intron-enriched genes.

The sequencing data have been deposited in the DNA Data Bank of Japan under accession no. DRA004262.

Online supplemental material

Fig. S1 shows marker expression in primary cultures of corpus luteal cells. Fig. S2 shows detection of putative transcription sites by *Neat1* tail probe. Fig. S3 shows series of optical sections of paraspeckles. Fig. S4 shows immunohistochemical detection of paraspeckle proteins before and after the FISH treatment. Fig. S5 shows mapping of RNA sequencing reads and the shell-distribution of CHART-enriched AG-rich RNAs. Table S1 is a list of the number of CHART RNA sequencing (RNAseq) reads mapped to introns of Refseq genes. Table S2 is a list of the number of CHART RNAseq reads mapped to mRNAs of Refseq genes. Table S3 is a list of primers and oligonucleotides used in this study. Online supplemental material is available at http://www.jcb.org/cgi/content/full/jcb.201601071/DC1. Additional data are available in the JCB DataViewer at http://dx.doi.org/10.1083/jcb.201601071.dv.

Acknowledgments

We thank Ms. Chieko Nashiki for maintenance of laboratory research environment and technical assistance. We also thank Dr. Takeya Kasukawa, Dr. Osamu Nishimura, and Dr. Rei Yoshimoto for assistance on deep sequencing data handling, Dr. Shigehiro Kuraku for coordinating the deep sequencing analyses, and Dr. Tomohiro Yamazaki and Dr. Christopher Davis for discussions.

This work was supported by Grants-in-Aid for Scientific Research on Innovative Areas from the Ministry of Education, Culture, Sports, Science, and Technology of Japan (26113005 and 23111005).

The authors declare no competing financial interests.

Submitted: 21 January 2016
Accepted: 24 August 2016

References

Bond, C.S., and A.H. Fox. 2009. Paraspeckles: nuclear bodies built on long noncoding RNA. *J. Cell Biol.* 186:637–644. http://dx.doi.org/10.1083/jcb.200906113

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. http://dx.doi.org/10.1091/mbc.E06-09-0846

Cerase, A., D. Smeets, Y.A. Tang, M. Gdula, F. Kraus, M. Spivakov, B. Moindrot, M. Leleu, A. Tattermusch, J. Demmerle, et al. 2014. Spatial separation of Xist RNA and polycomb proteins revealed by superresolution microscopy. *Proc. Natl. Acad. Sci. USA.* 111:2235–2240. http://dx.doi.org/10.1073/pnas.1312951111

Chakravarty, D., A. Sboner, S.S. Nair, E. Giannopoulou, R. Li, S. Hennig, J.M. Mosquera, J. Pauwels, K. Park, M. Kossai, et al. 2014. The oestrogen receptor alpha-regulated lncRNA NEAT1 is a critical modulator of prostate cancer. *Nat. Commun.* 5:5383. http://dx.doi.org/10.1038/ncomms6383

Chen, L.L., and G.G. Carmichael. 2009. Altered nuclear retention of mRNAs containing inverted repeats in human embryonic stem cells: functional role of a nuclear noncoding RNA. *Mol. Cell.* 35:467–478. http://dx.doi.org/10.1016/j.molcel.2009.06.027

Chu, C., Q.C. Zhang, S.T. da Rocha, R.A. Flynn, M. Bharadwaj, J.M. Calabrese, T. Magnuson, E. Heard, and H.Y. Chang. 2015. Systematic discovery of Xist RNA binding proteins. *Cell.* 161:404–416. http://dx.doi.org/10.1016/j.cell.2015.03.025

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. http://dx.doi.org/10.1016/j.molcel.2009.01.026

Davis, C.P., and J.A. West. 2015. Purification of specific chromatin regions using oligonucleotides: capture hybridization analysis of RNA targets (CHART). *Methods Mol. Biol.* 1262:167–182. http://dx.doi.org/10.1007/978-1-4939-2253-6_10

Dong, B., D.S. Horowitz, R. Kobayashi, and A.R. Krainer. 1993. Purification and cDNA cloning of HeLa cell p54nrb, a nuclear protein with two RNA recognition motifs and extensive homology to human splicing factor PSF and *Drosophila* NONA/BJ6. *Nucleic Acids Res.* 21:4085–4092. http://dx.doi.org/10.1093/nar/21.17.4085

Fox, A.H., Y.W. Lam, A.K.L. Leung, C.E. Lyon, J. Andersen, M. Mann, and A.I. Lamond. 2002. Paraspeckles: a novel nuclear domain. *Curr. Biol.* 12:13–25. http://dx.doi.org/10.1016/S0960-9822(01)00632-7

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. http://dx.doi.org/10.1091/mbc.E05-06-0587

Gustafsson, M.G. 2000. Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy. *J. Microsc.* 198:82–87. http://dx.doi.org/10.1046/j.1365-2818.2000.00710.x

Han, T.W., M. Kato, S. Xie, L.C. Wu, H. Mirzaei, J. Pei, M. Chen, Y. Xie, J. Allen, G. Xiao, and S.L. McKnight. 2012. Cell-free formation of RNA granules: bound RNAs identify features and components of cellular assemblies. *Cell.* 149:768–779. http://dx.doi.org/10.1016/j.cell.2012.04.016

Hennig, S., G. Kong, T. Mannen, A. Sadowska, S. Kobelke, A. Blythe, G.J. Knott, K.S. Iyer, D. Ho, E.A. Newcombe, et al. 2015. Prion-like domains in RNA binding proteins are essential for building subnuclear paraspeckles. *J. Cell Biol.* 210:529–539. http://dx.doi.org/10.1083/jcb.201504117

Hicks, G.G., N. Singh, A. Nashabi, S. Mai, G. Bozek, L. Klewes, D. Arapovic, E.K. White, M.J. Koury, E.M. Oltz, et al. 2000. Fus deficiency in mice results in defective B-lymphocyte development and activation, high levels of chromosomal instability and perinatal death. *Nat. Genet.* 24:175–179. http://dx.doi.org/10.1038/72842

Hirose, T., G. Virnicchi, A. Tanigawa, T. Naganuma, R. Li, H. Kimura, T. Yokoi, S. Nakagawa, M. Bénard, A.H. Fox, and G. Pierron. 2014. NEAT1 long noncoding RNA regulates transcription via protein sequestration within subnuclear bodies. *Mol. Biol. Cell.* 25:169–183. http://dx.doi.org/10.1091/mbc.E13-09-0558

Hoell, J.I., E. Larsson, S. Runge, J.D. Nusbaum, S. Duggimpudi, T.A. Farazi, M. Hafner, A. Borkhardt, C. Sander, and T. Tuschl. 2011. RNA targets of wild-type and mutant FET family proteins. *Nat. Struct. Mol. Biol.* 18:1428–1431. http://dx.doi.org/10.1038/nsmb.2163

Imamura, K., N. Imamachi, G. Akizuki, M. Kumakura, A. Kawaguchi, K. Nagata, A. Kato, Y. Kawaguchi, H. Sato, M. Yoneda, et al. 2014. Long noncoding RNA NEAT1-dependent SFPQ relocation from promoter region to paraspeckle mediates IL8 expression upon immune stimuli. *Mol. Cell.* 53:393–406. (published erratum appears in *Mol. Cell.* 2014. 540:1055) http://dx.doi.org/10.1016/j.molcel.2014.01.009

Kato, M., T.W. Han, S. Xie, K. Shi, X. Du, L.C. Wu, H. Mirzaei, E.J. Goldsmith, J. Longgood, J. Pei, et al. 2012. Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell.* 149:753–767. http://dx.doi.org/10.1016/j.cell.2012.04.017

Lagier-Tourenne, C., M. Polymenidou, K.R. Hutt, A.Q. Vu, M. Baughn, S.C. Huelga, K.M. Clutario, S.C. Ling, T.Y. Liang, C. Mazur, et al. 2012. Divergent roles of ALS-linked proteins FUS/TLS and TDP-43 intersect in processing long pre-mRNAs. *Nat. Neurosci.* 15:1488–1497. http://dx.doi.org/10.1038/nn.3230

Mao, Y.S., H. Sunwoo, B. Zhang, and D.L. Spector. 2011. Direct visualization of the co-transcriptional assembly of a nuclear body by noncoding RNAs. *Nat. Cell Biol.* 13:95–101. http://dx.doi.org/10.1038/ncb2140

McHugh, C.A., C.K. Chen, A. Chow, C.F. Surka, C. Tran, P. McDonel, A. Pandya-Jones, M. Blanco, C. Burghard, A. Moradian, et al. 2015. The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature.* 521:232–236. http://dx.doi.org/10.1038/nature14443

Mito, M., T. Kawaguchi, T. Hirose, and S. Nakagawa. 2016. Simultaneous multicolor detection of RNA and proteins using super-resolution

Downloaded from on December 11, 2016

Published September 19, 2016

microscopy. *Methods.* 98:158–165. http://dx.doi.org/10.1016/j.ymeth.2015.11.007

Moindrot, B., A. Cerase, H. Coker, O. Masui, A. Grijzenhout, G. Pintacuda, L. Schermelleh, T.B. Nesterova, and N. Brockdorff. 2015. A Pooled shRNA Screen Identifies Rbm15, Spen, and Wtap as Factors Required for Xist RNA-Mediated Silencing. *Cell Reports.* 12:562–572. http://dx.doi.org/10.1016/j.celrep.2015.06.053

Naganuma, T., S. Nakagawa, A. Tanigawa, Y.F. Sasaki, N. Goshima, and T. Hirose. 2012. Alternative 3′-end processing of long noncoding RNA initiates construction of nuclear paraspeckles. *EMBO J.* 31:4020–4034. http://dx.doi.org/10.1038/emboj.2012.251

Nakagawa, S., T. Naganuma, G. Shioi, and T. Hirose. 2011. Paraspeckles are subpopulation-specific nuclear bodies that are not essential in mice. *J. Cell Biol.* 193:31–39. http://dx.doi.org/10.1083/jcb.201011110

Nakagawa, S., M. Shimada, K. Yanaka, M. Mito, T. Arai, E. Takahashi, Y. Fujita, T. Fujimori, L. Standaert, J.C. Marine, and T. Hirose. 2014. The lncRNA Neat1 is required for corpus luteum formation and the establishment of pregnancy in a subpopulation of mice. *Development.* 141:4618–4627. http://dx.doi.org/10.1242/dev.110544

Platani, M., and A. Lamond. 2004. Nuclear organisation and subnuclear bodies. In RNA Trafficking and Nuclear Structure Dynamics. Vol. 35. P. Jeanteur, editor. Springer, Berlin. 1–22.

Polymenidou, M., C. Lagier-Tourenne, K.R. Hutt, S.C. Huelga, J. Moran, T.Y. Liang, S.C. Ling, E. Sun, E. Wancewicz, C. Mazur, et al. 2011. Long pre-mRNA depletion and RNA missplicing contribute to neuronal vulnerability from loss of TDP-43. *Nat. Neurosci.* 14:459–468. http://dx.doi.org/10.1038/nn.2779

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. http://dx.doi.org/10.1016/j.cell.2005.08.033

Quinn, J.J., I.A. Ilik, K. Qu, P. Georgiev, C. Chu, A. Akhtar, and H.Y. Chang. 2014. Revealing long noncoding RNA architecture and functions using domain-specific chromatin isolation by RNA purification. *Nat. Biotechnol.* 32:933–940. http://dx.doi.org/10.1038/nbt.2943

Rogelj, B., L.E. Easton, G.K. Bogu, L.W. Stanton, G. Rot, T. Curk, B. Zupan, Y. Sugimoto, M. Modic, N. Haberman, et al. 2012. Widespread binding of FUS along nascent RNA regulates alternative splicing in the brain. *Sci. Rep.* 2:603. http://dx.doi.org/10.1038/srep00603

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. USA.* 106:2525–2530. http://dx.doi.org/10.1073/pnas.0807899106

Schermelleh, L., R. Heintzmann, and H. Leonhardt. 2010. A guide to super-resolution fluorescence microscopy. *J. Cell Biol.* 190:165–175. http://dx.doi.org/10.1083/jcb.201002018

Schwartz, J.C., C.C. Ebmeier, E.R. Podell, J. Heimiller, D.J. Taatjes, and T.R. Cech. 2012. FUS binds the CTD of RNA polymerase II and regulates its phosphorylation at Ser2. *Genes Dev.* 26:2690–2695. http://dx.doi.org/10.1101/gad.204602.112

Shamir, L., N. Orlov, D.M. Eckley, T. Macura, J. Johnston, and I.G. Goldberg. 2008. Wndchrm - an open source utility for biological image analysis. *Source Code Biol. Med.* 3:13. http://dx.doi.org/10.1186/1751-0473-3-13

Shav-Tal, Y., and D. Zipori. 2002. PSF and p54(nrb)/NonO--multi-functional nuclear proteins. *FEBS Lett.* 531:109–114. http://dx.doi.org/10.1016/S0014-5793(02)03447-6

Shelkovnikova, T.A., H.K. Robinson, C. Troakes, N. Ninkina, and V.L. Buchman. 2014. Compromised paraspeckle formation as a pathogenic factor in FUSopathies. *Hum. Mol. Genet.* 23:2298–2312. http://dx.doi.org/10.1093/hmg/ddt622

Shevtsov, S.P., and M. Dundr. 2011. Nucleation of nuclear bodies by RNA. *Nat. Cell Biol.* 13:167–173. http://dx.doi.org/10.1038/ncb2157

Simon, M.D., C.I. Wang, P.V. Kharchenko, J.A. West, B.A. Chapman, A.A. Alekseyenko, M.L. Borowsky, M.I. Kuroda, and R.E. Kingston. 2011. The genomic binding sites of a noncoding RNA. *Proc. Natl. Acad. Sci. USA.* 108:20497–20502. http://dx.doi.org/10.1073/pnas.1113536108

Simon, M.D., S.F. Pinter, R. Fang, K. Sarma, M. Rutenberg-Schoenberg, S.K. Bowman, B.A. Kesner, V.K. Maier, R.E. Kingston, and J.T. Lee. 2013. High-resolution Xist binding maps reveal two-step spreading during X-chromosome inactivation. *Nature.* 504:465–469. http://dx.doi.org/10.1038/nature12719

Souquere, S., G. Beauclair, F. Harper, A. Fox, and G. Pierron. 2010. Highly ordered spatial organization of the structural long noncoding NEAT1 RNAs within paraspeckle nuclear bodies. *Mol. Biol. Cell.* 21:4020–4027. http://dx.doi.org/10.1091/mbc.E10-08-0690

Spector, D.L. 2006. SnapShot: Cellular bodies. *Cell.* 127:1071. http://dx.doi.org/10.1016/j.cell.2006.11.026

Standaert, L., C. Adriaens, E. Radaelli, A. Van Keymeulen, C. Blanpain, T. Hirose, S. Nakagawa, and J.C. Marine. 2014. The long noncoding RNA Neat1 is required for mammary gland development and lactation. *RNA.* 20:1844–1849. http://dx.doi.org/10.1261/rna.047332.114

Sunwoo, H., M.E. Dinger, J.E. Wilusz, P.P. Amaral, J.S. Mattick, and D.L. Spector. 2009. MEN ε/βnuclear-retained non-coding RNAs are up-regulated upon muscle differentiation and are essential components of paraspeckles. *Genome Res.* 19:347–359. http://dx.doi.org/10.1101/gr.087775.108

Tollervey, J.R., T. Curk, B. Rogelj, M. Briese, M. Cereda, M. Kayikci, J. König, T. Hortobágyi, A.L. Nishimura, V. Zupunski, et al. 2011. Characterizing the RNA targets and position-dependent splicing regulation by TDP-43. *Nat. Neurosci.* 14:452–458. http://dx.doi.org/10.1038/nn.2778

West, J.A., C.P. Davis, H. Sunwoo, M.D. Simon, R.I. Sadreyev, P.I. Wang, M.Y. Tolstorukov, and R.E. Kingston. 2014. The long noncoding RNAs NEAT1 and MALAT1 bind active chromatin sites. *Mol. Cell.* 55:791–802. http://dx.doi.org/10.1016/j.molcel.2014.07.012

Downloaded from on December 11, 2016