



Title	Functional Domains of NEAT1 Architectural lncRNA Induce Paraspeckle Assembly through Phase Separation
Author(s)	Yamazaki, Tomohiro; Souquere, Sylvie; Chujo, Takeshi; Kobelke, Simon; Chong, Yee Seng; Fox, Archa H.; Bond, Charles S.; Nakagawa, Shinichi; Pierron, Gerard; Hirose, Tetsuro
Citation	Molecular cell, 70(6), 1038-1053 <a href="https://doi.org/10.1016/j.molcel.2018.05.019">https://doi.org/10.1016/j.molcel.2018.05.019</a>
Issue Date	2018-06-21
Doc URL	<a href="http://hdl.handle.net/2115/74758">http://hdl.handle.net/2115/74758</a>
Rights	© 2018. This manuscript version is made available under the CC-BY-NC-ND 4.0 license <a href="http://creativecommons.org/licenses/by-nc-nd/4.0/">http://creativecommons.org/licenses/by-nc-nd/4.0/</a>
Rights(URL)	<a href="http://creativecommons.org/licenses/by-nc-nd/4.0/">http://creativecommons.org/licenses/by-nc-nd/4.0/</a>
Type	article (author version)
File Information	Yamazaki_et_al.(HUSCAP).pdf



[Instructions for use](#)

## **Functional domains of NEAT1 architectural lncRNA induce paraspeckle assembly through phase separation**

Tomohiro Yamazaki<sup>1</sup>, Sylvie Souquere<sup>2</sup>, Takeshi Chujo<sup>1</sup>, Simon Kobelke<sup>3</sup>, Yee Seng Chong<sup>4</sup>, Archa H. Fox<sup>3</sup>, Charles S. Bond<sup>4</sup>, Shinichi Nakagawa<sup>5</sup>, Gerard Pierron<sup>2</sup>, and Tetsuro Hirose<sup>1, 6, \*</sup>

<sup>1</sup>Institute for Genetic Medicine, Hokkaido University, Sapporo 060-0815, Japan

<sup>2</sup>Centre National de la Recherche Scientifique, UMR-8122, Institut Gustave Roussy, Villejuif 94805, France

<sup>3</sup>School of Human Sciences, University of Western Australia, Crawley, Western Australia, 6009, Australia

<sup>4</sup>School of Molecular Sciences, University of Western Australia, Crawley, Western Australia, 6009, Australia

<sup>5</sup>Faculty of Pharmaceutical Sciences, Hokkaido University, Sapporo 060-0812, Japan

<sup>6</sup>Lead contact

\*Correspondence: [hirose@igm.hokudai.ac.jp](mailto:hirose@igm.hokudai.ac.jp)

Keywords: Noncoding RNA, architectural RNA, nuclear bodies, paraspeckle, RNA-binding proteins, prion-like domain, phase separation, 1,6-hexanediol, CRISPR/Cas9, haploid cell

## Summary

A class of long noncoding RNAs (lncRNAs) has architectural functions in nuclear body construction; however, specific RNA domains dictating their architectural functions remain uninvestigated. Here, we identified the domains of the architectural NEAT1 lncRNA that construct paraspeckles. Systematic deletion of NEAT1 portions using CRISPR/Cas9 in haploid cells revealed modular domains of NEAT1 important for RNA stability, isoform switching, and paraspeckle assembly. The middle domain, containing functionally redundant subdomains, was responsible for paraspeckle assembly. Artificial tethering of the NONO protein to a NEAT1\_2 mutant lacking the functional subdomains rescued paraspeckle assembly, and this required the NOPS dimerization domain of NONO. Paraspeckles exhibit phase-separated properties including susceptibility to 1,6-hexanediol treatment. RNA fragments of the NEAT1\_2 subdomains preferentially bound NONO/SFPQ, leading to phase-separated aggregates *in vitro*. Thus, we demonstrate that the enrichment of NONO dimers on the redundant NEAT1\_2 subdomains initiates construction of phase-separated paraspeckles, providing mechanistic insights into lncRNA-based nuclear body formation.

## Introduction

Long noncoding RNAs (lncRNAs) have recently been recognized as fundamental regulators of gene expression, but their mechanisms of action remain largely unknown. Unlike protein-coding mRNAs, lncRNAs *per se* form complex machineries with multiple RNA-binding proteins (RBPs) (Engreitz et al., 2016; Quinn and Chang, 2016). However, specific sequences or structural elements embedded in lncRNA sequences required for building the functional machinery have been poorly investigated. A limited number of lncRNAs are known to have modular domain compositions; for example, XIST requires the A-repeat domain comprising nine repetitive stem-loop structures for gene silencing on the inactivated X chromosome in mammalian cells (Wutz et al., 2002). SPEN interacts with the A-repeat domain and is required for gene silencing (Chu et al., 2015).

Among tens of thousands of human lncRNAs, several lncRNAs function as the structural scaffolds of membraneless subnuclear structures or nuclear bodies (NBs) (Chujo et al., 2016). NBs are located in the inter-chromatin spaces in the highly organized nucleus and consist of specific factors that function in various nuclear processes (Banani et al., 2017). Some NBs exhibit liquid-like properties, and RBP components of NBs containing the prion-like domain (PLD), low complexity domains, or intrinsically disordered regions induce liquid-liquid phase separation (LLPS) to form liquid droplets. This raises the intriguing possibility that LLPS is the driving force for NB formation (Banani et al., 2017; Stanek and Fox, 2017). RNA accelerates LLPS *in vitro* by sequestration of PLD-containing RBPs and/or alteration of their conformation (Shin and Brangwynne, 2017). Thus, RNA sequences that specifically interact with PLD-containing

RBPs can induce LLPS.

Several NBs contain specific lncRNA species that play architectural roles in their formation (Chujo et al., 2016). For example, NEAT1 serves as an essential architectural component of paraspeckle NBs (Clemson et al., 2009; Sasaki et al., 2009; Sunwoo et al., 2009). Paraspeckles were initially defined as the foci found in close proximity to nuclear speckles and are enriched with characteristic DBHS (Drosophila Behavior Human Splicing) RBPs including NONO, SFPQ, and PSPC1 (Fox et al., 2002; Visa et al., 1993). NEAT1 is transcribed from a specific locus in human chromosome 11 by RNA polymerase II (RNAPII) to produce two distinct isoforms, NEAT1\_1 (3.7 kb) or NEAT1\_2 (22.7 kb), as a consequence of alternative 3' end processing (Naganuma et al., 2012). NEAT1\_2 but not NEAT1\_1 is indispensable for paraspeckle formation (Naganuma et al., 2012). NEAT1\_1 is canonically polyadenylated, whereas NEAT1\_2 lacks a poly(A) tail and instead contains a triple helix (TH) structure at its 3' terminus (Sunwoo et al., 2009; Wilusz et al., 2012). Electron microscopic (EM) and super-resolution microscopic (SRM) analyses revealed that the paraspeckle is a highly ordered, spheroidal structure in which the 5' terminus of NEAT1 and 3' terminus of NEAT1\_2 are located in the outer shell, whereas the middle region of NEAT1\_2 is located in the inner core (Souquere et al., 2010; West et al., 2016). Paraspeckle formation is tightly coupled with ongoing RNAPII transcription of NEAT1\_2 and with the binding of multiple paraspeckle proteins (PSPs) to NEAT1 (Mao et al., 2011). More than 60 PSPs, most of which are RBPs, have been identified (Yamazaki and Hirose, 2015). Functional categorization of PSPs by extensive RNAi analysis revealed seven category 1 proteins

essential for formation of a paraspeckle with ~50 NEAT1\_2 molecules (Chujo et al., 2017; Naganuma et al., 2012). Among the seven category 1 proteins, six possess characteristic PLDs (Yamazaki and Hirose, 2015). The PLDs of FUS and RBM14 form hydrogels *in vitro* and are required for paraspeckle formation *in vivo* (Hennig et al., 2015). In addition, SWI/SNF chromatin remodeling complexes play an essential role in paraspeckle assembly without requiring their ATP-dependent remodeling activity (Kawaguchi et al., 2015). Transcription of NEAT1 is upregulated by various conditions, resulting in an increase in the size and number of paraspeckles, which can sequester specific RBPs and/or RNAs away from the nucleoplasm to control gene expression (Chen and Carmichael, 2009; Hirose et al., 2014). Physiologically, NEAT1 is required for the development of specific tissues such as the corpus luteum in mice, and is involved in the progression of various cancers (Adriaens et al., 2016; Nakagawa et al., 2014).

Despite growing knowledge of the roles of PSPs during paraspeckle formation, it is not known what parts of NEAT1\_2 are responsible for its architectural roles. By carrying out CRISPR/Cas9-mediated systematic deletions of portions of NEAT1\_2 in the human haploid cell line, we revealed the modular domain structure of NEAT1\_2 with three functionally distinct domains important for (A) stabilization of NEAT1\_2, (B) NEAT1 isoform switching, and (C) paraspeckle assembly. The middle domain, which is necessary and sufficient for paraspeckle assembly, contains three functional subdomains. Artificial tethering of each of three category 1 proteins to a functionally defective NEAT1\_2 mutant rescued paraspeckle formation, and the rescue activity required the NOPS dimerization domain of NONO. Paraspeckles were disintegrated by

treatment with 1,6-hexanediol (1,6-HD) in a similar manner to other phase-separated NBs. The RNA fragments of the NEAT1\_2 subdomain preferentially bound NONO and SFPQ, and induced higher-order assembly *in vitro*. In summary, we argue that the identified NEAT1\_2 subdomains recruit the NONO dimer to initiate oligomerization with other PSPs, promoting formation of the phase-separated paraspeckle structure. Our findings provide fundamental insights into the formation of NBs built on architectural lncRNAs (arcRNAs).

## Results

### **The 3' terminal TH structure stabilizes NEAT1\_2 lncRNA *in vivo***

To investigate the functional domains of NEAT1 lncRNA, we set up an experimental system using CRISPR/Cas9 genome editing in the human haploid cell line HAP1. We deleted various portions of NEAT1 by expressing CRISPR/Cas9 proteins and one or two single guide RNAs (sgRNAs) (Figure 1A). We first targeted the TH structure of NEAT1\_2 to validate our system. We established two HAP1 clones carrying deletions of a part of the conserved stem loop and U-rich motif 2 in the TH ( $\Delta$ TH#1 and #2 in Figure 1B). Single molecule FISH (smFISH) was employed to observe the paraspeckles in these clones. Our smFISH specifically detected NEAT1 because no smFISH signals were detected in NEAT1 knockout (KO) cells ( $\Delta$ NEAT1) (Figure S1A–S1D). In the  $\Delta$ TH clones, paraspeckles were almost undetectable by smFISH and conventional RNA-FISH (Figure 1C and 1D; Figure S1E–S1G). RT-qPCR and RNase protection assay (RPA) revealed considerably lower levels of NEAT1\_2, but not NEAT1\_1, in the  $\Delta$ TH clones than in the wild type (WT) (Figure 1E and 1F; Figure S1H). Furthermore, NEAT1\_2 in  $\Delta$ TH cells was much more unstable than in WT cells (Figure 1G). Thus, TH stabilizes NEAT1\_2 *in vivo*.

### **NEAT1 lncRNA domains required for expression, isoform switching, and paraspeckle formation**

We next investigated the RNA domains of NEAT1 required for its expression and for paraspeckle formation. To explore the entire region of NEAT1, we established HAP1 clones lacking portions of NEAT1 (0.6–1.9 kb) as shown in Figure 2A, and examined two

independent clones for each mutant to validate their phenotypes by smFISH and RT-qPCR of NEAT1\_1 and NEAT1\_2 (Figure S2A and S2B). Overall, the expression levels of NEAT1\_2 were positively correlated with the presence of paraspeckles (Figure S2A–S2C). Among these deletion clones, paraspeckles did not form in the deletion lines  $\Delta 0$ –1k,  $\Delta 2.1$ –2.8k, and  $\Delta 4$ –5.1k (k for kb) (Figure 2B and 2C; Figure S2A–S2C).

In the  $\Delta 0$ –1k clones, levels of NEAT1\_1 and NEAT1\_2 were markedly lower than in the WT, as determined by RT-qPCR and RPA (Figure 2D and 2F). This reduction was partly caused by destabilization of NEAT1\_1 and NEAT1\_2 (Figure 2G). In  $\Delta 0$ –1k cells, the marked disappearance of paraspeckles is likely to be a consequence of the strong reduction in NEAT1\_2, suggesting that this 5' terminal region contains specific elements for NEAT1 expression, which is further supported by experiments to narrow down the functional domain within the 0–1 kb region (Figure S2D–S2H).

In the  $\Delta 2.1$ –2.8k and  $\Delta 4$ –5.1k clones, the expression levels of NEAT1\_2 were significantly lower than in the WT, whereas NEAT1\_1 levels were 2–2.5-fold higher (Figures 2E and S2C–S2H). As these regions were located close to the polyadenylation signal (PAS) of NEAT1\_1, we supposed that these domains are involved in the alternative 3' end processing, facilitating NEAT1\_2 synthesis by inhibiting the usage of NEAT1\_1 PAS. To test this possibility, we established HAP1 cell lines carrying deletions of PAS alone ( $\Delta$ PAS), both  $\Delta$ PAS and  $\Delta 2.1$ –2.8k ( $\Delta 2.1$ –2.8k/ $\Delta$ PAS), and both  $\Delta$ PAS and  $\Delta 4$ –5.1k ( $\Delta 4$ –5.1k/ $\Delta$ PAS) (Figure 2A). Deletion of PAS alone caused a strong reduction in NEAT1\_1 and slight upregulation of NEAT1\_2, not affecting the appearance of paraspeckles (Figure 2H–2J). Unlike  $\Delta 2.1$ –2.8k or  $\Delta 4$ –5.1k clones,  $\Delta 2.1$ –2.8k/ $\Delta$ PAS and

$\Delta$ 4–5.1k/ $\Delta$ PAS clones showed comparable expression of NEAT1\_2 to the WT, and paraspeckles were clearly observed (Figure 2H–2J). Therefore, NEAT1 2.1–2.8 kb and 4–5.1 kb domains suppress NEAT1\_1 PAS-dependent polyadenylation.

### **The middle domain of NEAT1\_2 is required for ordered paraspeckle assembly**

Several lncRNAs like XIST reportedly possess repeated sequence stretches that have important functional roles (Chu et al., 2015; Wutz et al., 2002). Comparison of the NEAT1\_2 sequence with itself using the BLAST algorithm revealed the presence of several repetitive sequences with varying length and similarity among three mammalian species (Figure S3A). Some of the repetitive sequences in human NEAT1\_2 overlapped with annotated LINE and SINE elements (Figure S3B). To investigate the importance of these repetitive sequences, we deleted the middle portion (8–16.6 kb) of NEAT1\_2 ( $\Delta$ middle in Figure 3A), which contains multiple long repetitive sequences and is localized in the core of the paraspeckle. In the  $\Delta$ middle cells, we observed tiny dispersed foci (Figures 3B and S3C). A similar phenotype was observed in another distinct  $\Delta$ middle cell line ( $\Delta$ middle#2), which was established with another set of sgRNAs, under both MG132-untreated and -treated conditions (Figure S3C). We determined paraspeckle sizes by measuring the area and sum signal intensity of each focus in cells with or without MG132 treatment (Figures 3C and S3D). MG132 induces NEAT1 transcription to enlarge paraspeckles (Hirose et al., 2014b). The size and smFISH signal intensity of paraspeckles in  $\Delta$ middle were significantly lower than in the WT (Figures 3C and S3D), although the NEAT1\_2 expression levels were comparable to those in the WT (Figure

S3E). Therefore, the middle region is required for paraspeckle assembly rather than NEAT1\_2 expression. Similarly, dispersed foci were observed in FUS KO cells, which are defective in paraspeckle assembly without affecting NEAT1\_2 levels (Chujo et al., 2017; West et al., 2016) (Figure 3C; Figure S3C–S3E). Another large deletion of a different region of NEAT1\_2 ( $\Delta$ 13–20.2k) did not significantly affect the integrity of paraspeckles (Figure 3B and 3C; Figure S3C and S3D), excluding the possibility that  $\Delta$ middle NEAT1 was too short to form paraspeckles. Thus, the middle domain is specifically required for paraspeckle integrity.

To clarify the ultra-structures of the paraspeckles in  $\Delta$ middle cells, we used EM. Previous EM analyses showed that paraspeckles are highly ordered (Souquere et al., 2010). As reported, the probes against the NEAT1 5' region detected the surface of paraspeckles in WT and  $\Delta$ 13–20.2k cells. By contrast, in  $\Delta$ middle cells, the same 5' probes randomly detected both the surface and interior of the paraspeckles (Figure 3D, upper panels), showing that the paraspeckle structure became disordered in  $\Delta$ middle cells. We measured the size of paraspeckles characterized by the electron dense NONO-positive oval foci in the WT,  $\Delta$ middle, and  $\Delta$ 13–20.2k (Figure 3D, lower panels). Diameters of the short (Sx) and long (Lx) axes of paraspeckles in  $\Delta$ middle cells and not  $\Delta$ 13–20.2k cells were significantly smaller than in the WT, thereby shrinking the surface area (Figure 3E; Figure S3F and S3G).

Next, we examined *de novo* formation of paraspeckles in WT,  $\Delta$ middle, and  $\Delta$ 13–20.2k cells using 5,6-dichloro-1- $\beta$ -D-ribofuranosylbenzimidazole (DRB), a reversible inhibitor of RNAPII. During DRB treatment, NEAT1 levels decreased in the

three cell lines and increased upon release from DRB (Figure S3H). Accordingly, DRB treatment disintegrated paraspeckles, and upon washing out DRB, paraspeckles formed *de novo*. SRM analyses revealed that paraspeckles formed *de novo* in the WT and  $\Delta 13\text{--}20.2\text{k}$  had an ordered core-shell structure, whereas the structure was disordered in  $\Delta\text{middle}$  (Figure 3F). Therefore, the assembly step of paraspeckles was defective in  $\Delta\text{middle}$ . Consistent with the importance of the middle domain, the essential PSPs including NONO, SFPQ, and FUS were mainly localized in the core of WT paraspeckles where the NEAT1\_2 middle domain is located (Figures 3F and S3I).

**The middle domain of NEAT1\_2 is sufficient for formation of the ordered paraspeckle structure with a phase separated feature**

Based on the functional domains of NEAT1 identified above, we next explored the minimal regions of NEAT1\_2 required for paraspeckle formation. We established HAP1 clones that contained only the expression domains (5' terminal region and the 3' TH) and the middle domain (8–16.6 kb) required for paraspeckle assembly (Figure 4A). SRM showed that this mutant, termed mini-NEAT1, could form an ordered paraspeckle structure (Figure 4B). EM observation of the mini-NEAT1 paraspeckles also supported their ordered structure with the interior localization of NONO, patchy localization of BRG1, surface localization of NEAT1-5', and interior localization of the NEAT1\_2 middle region, all of which are consistent with the patterns in the WT (Kawaguchi et al., 2015; Souquere et al., 2010) (Figure 4F and 4G). Despite the ordered structure, the mini-NEAT1 paraspeckles were smaller with shorter Lx and Sx than those in WT cells (Figure 4H).

Mini-NEAT1 was expressed at a comparable level to WT NEAT1 (Figure 4D). Unlike WT paraspeckles in which the 3' end of NEAT1\_2 is located in the shell, the 3' end of mini-NEAT1 was located in the core (Figure 4A–4C and 4E). These data suggest that the shortened length of mini-NEAT1 (9.7 kb) and its configuration reflect the decreased size of paraspeckles. An essential PSP, SFPQ was less efficiently recruited to the mini-NEAT1 paraspeckles than to WT paraspeckles, but other essential PSPs were recruited to these foci (Figure S4A).

Paraspeckles, which have a core-shell spheroidal structure, have been reported to be dynamic NBs, into which and from which a proportion of PSPs (~60-75%) dynamically migrate in and out (Mao et al., 2011). Additionally, PLDs of PSPs are required for paraspeckle formation *in vivo*, and recombinant PLD of PSPs can induce phase separation *in vitro* (Hennig et al., 2015). These results raised the possibility that paraspeckles are phase-separated structures. To investigate this, we employed 1,6-HD, which can disrupt several phase-separated subcellular structures *in vivo* by disrupting their multivalent hydrophobic interactions (Kroschwald et al., 2015; Lin et al., 2016; Shin and Brangwynne, 2017). After treatment with 6–10% 1,6-HD, the intact paraspeckles, in which NEAT1 and NONO should be co-localized, disappeared in both HAP1 WT and mini-NEAT1 cells (Figure 4I and 4J), suggesting that the paraspeckles are sensitive to 1,6-HD and thus are likely phase-separated in these cells. In the same experimental conditions, Cajal bodies, which were labeled with COIL and are known to be phase-separated (Lin et al., 2016), were readily disintegrated (Figure S4F). On the other hand, nuclear speckles labeled with SRSF2 were not significantly affected (Figure S4F),

showing that our 1,6-HD condition was mild enough not to disrupt all subcellular structures. 2,5-hexanediol (2,5-HD) is an aliphatic alcohol like 1,6-HD, but does not affect phase-separated cellular structures (Lin et al., 2016). As expected, even in the presence of 10% 2,5-HD, the paraspeckles were clearly visible (Figure S4G). Next, we examined the paraspeckle substructure by SRM in the presence of 1,6-HD or 2,5-HD. Disrupted structures were observed in the 1,6-HD-treated condition, while ordered paraspeckle structures were clearly observed in the 2,5-HD-treated condition (Figure 4K). Co-immunoprecipitation (coIP) revealed that interactions of NONO with RBM14 and FUS were markedly weakened by 1,6-HD washing buffer (Figure 4L), suggesting that these interactions are part of the effector sites for 1,6-HD. Together with previous reports, these data suggest that the paraspeckles in WT and mini-NEAT1 cells are phase-separated *in vivo*.

### **NEAT1\_2 middle domain contains multiple redundant subdomains for ordered paraspeckle assembly**

To identify which part of the NEAT1\_2 middle domain is required for paraspeckle assembly, we established additional deletion mutants (Figure 5A upper). NEAT1 expression levels in these clones were comparable (Figure S5A and S5B), and all new deletion mutants showed significant defects in paraspeckle formation, although these defects were milder than defects in  $\Delta$ middle cells (Figure 5B and 5C; Figure S5C and S5D). Using SRM, we quantified the ratio of ordered and disordered paraspeckles. Consistent with the smFISH results in Figure 5B and 5C, all the deletion mutants showed

defects in formation of ordered paraspeckle structures (Figure 5D and 5E). In particular,  $\Delta 8$ –15.4k mutants showed a strong defect, and the remaining mutants ( $\Delta 8$ –13k,  $\Delta 8$ –11.1k, and  $\Delta 11.1$ –16.6k) exhibited milder but significant defects (Figure 5D and 5E). These data suggest that the whole middle domain contributes to paraspeckle assembly.

As an alternative approach to dissect the middle domain of NEAT1\_2, deletion mutants were established based on mini-NEAT1 (Figure 5A lower). NEAT1\_2 was detected at comparable levels in all the new mutants (Figure S5E), and SRM showed that the mutants m9.2–16.6k, m9.8–16.6k, m11.1–16.6k, and m12–16.6k (m for mini-NEAT1) could predominantly form ordered structures (Figure 5F and 5G). By contrast, m13–16.6k and m8–15.4k could only form tiny disordered structures (Figure 5F and 5G), suggesting that the 12–13 kb and 15.4–16.6 kb regions are functional subdomains for paraspeckle assembly. It should be noted that the stepwise deletions from 9.8 to 12 kb gradually decreased the ratio of ordered paraspeckles (Figure 5G), suggesting the presence of redundant subdomains in this region. We thus examined whether the 12–13 kb region was solely required for the formation of ordered paraspeckles in the m9.8–16.6k mutant in which ordered paraspeckles were formed. To this end, we established HAP1 clones lacking the 12–13 kb region in m9.8–16.6k ( $\Delta 12$ –13k/m9.8–16.6k in Figure 5A lower). SRM showed the presence of ordered paraspeckles in these cells (Figure 5F and 5G), showing that 9.8–12 kb and 12–13 kb regions function redundantly in the formation of ordered paraspeckles. These data collectively demonstrate that the middle domain contains multiple redundant subdomains for paraspeckle assembly.

## **An essential NEAT1\_2 subdomain can be functionally replaced by artificially tethered essential PSPs**

We next investigated the role of the identified functional subdomains of NEAT1\_2 (9.8–12 kb, 12–13 kb, and 15.4–16.6 kb) in paraspeckle assembly. We hypothesized that these subdomains recruit the PSPs essential for paraspeckle assembly. We examined whether tethering of PSPs to a functionally defective NEAT1\_2 can rescue paraspeckle assembly. We introduced 6×MS2-binding sites (6×MS2BS) at 14 kb of NEAT1\_2 in m13–16.6k cells (m13–16.6k/6×MS2BS) (Figure 6A). We selected four essential PSPs (NONO, SFPQ, FUS, and RBM14) for this tethering experiment. We expressed MS2 coat protein (MCP) fused with PSPs (e.g., MCP-NONO) in this cell line (Figure S6A and S6B). By expressing MCP-NONO, MCP-SFPQ, or MCP-FUS, paraspeckle formation was rescued in the m13–16.6k/6×MS2BS cells (Figures 6B left, 6E, and S6C). Neither negative control MCP-GFP-NLS nor MCP-RBM14 rescued paraspeckle formation. As controls, we expressed the MCP-fusion proteins in the parental m13–16.6k cells without 6×MS2BS, but none of them rescued paraspeckle formation (Figures 6B right, 6E, and S6C). SRM revealed that the paraspeckles rescued by the tethering possessed the properly ordered core-shell structure (Figure 6C). Treatment with 1,6-HD but not 2,5-HD disrupted the tethering-induced paraspeckles (Figure S6D). These data suggest that tethering of NONO, SFPQ, or FUS compensates for the lack of the functional NEAT1\_2 subdomains and induces phase separation.

We next investigated which domains of NONO were essential for the rescue activity. Of the three proteins that rescued mini-paraspeckle formation by tethering to

m13–16.6k cells, NONO was chosen because SFPQ was hardly detected in mini-paraspeckles (Figure S4A) and FUS exhibited nonspecific RNA-binding features as shown later (Figure 7B). We constructed MCP-tagged NONO mutants lacking the coiled-coil domain (NONO $\Delta$ CC) essential for paraspeckle localization and intermolecular polymerization of NONO (Passon et al., 2012). We also made NONO lacking the NOPS domain (NONO $\Delta$ NOPS), which is essential for dimer formation with DBHS family members (Passon et al., 2012), or PLD (NONO $\Delta$ PLD) (Hennig et al., 2015) (Figure S6A and S6B). Tethering of MCP-NONO mutants revealed that NONO $\Delta$ NOPS lost the rescue activity, whereas NONO $\Delta$ CC and NONO $\Delta$ PLD did not (Figure 6D and 6E). Since the NOPS domain is essential for dimer formation of two DBHS proteins (Passon et al., 2012), NONO $\Delta$ NOPS is likely defective in interaction with endogenous NONO or other DBHS proteins such as SFPQ and PSPC1. To determine whether NOPS is required to interact with other DBHS proteins experimentally, we performed coIP using MCP-NONO WT,  $\Delta$ NOPS, or  $\Delta$ CC to examine the interactions with DBHS proteins (Figure 6F). As expected, MCP-NONO WT interacted with NONO, SFPQ, and PSPC1, whereas MCP-NONO $\Delta$ NOPS did not. MCP-NONO $\Delta$ CC still interacted with these proteins, although the interaction was weaker than with the WT. From these data, we argue that the actual function of the subdomains is to recruit NONO to form the primary dimers on NEAT1\_2 that become the scaffold to initiate oligomerization with other PSPs to form the structure of massive paraspeckles.

NONO plays an essential role in paraspeckle formation by maintaining NEAT1\_2 levels (Naganuma et al., 2012). The above data strongly suggest that NONO

is also involved in the assembly of paraspeckles. In our previous experiments, NEAT1\_2 and paraspeckle foci were undetectable after depletion of NONO (Naganuma et al., 2012); therefore, it was impossible to judge the requirement for NONO in the assembly step. However, we recently found that paraspeckle-like foci with substantial amounts of NEAT1\_2 were detectable in MG132-treated NONO KO cells ( $\Delta$ NONO), although they were much smaller than in WT cells (Figure S6E–S6H). SRM revealed that the paraspeckle-like foci in  $\Delta$ NONO were structurally disordered (Figure S6I), suggesting that NONO has a distinct, essential function for assembly of the ordered paraspeckle structure.

**RNA fragments of the NEAT1\_2 subdomain specifically bind NONO and SFPQ, and induce higher-order assembly *in vitro***

We next asked whether the identified NEAT1\_2 subdomains (9.8–12 kb and 12–13 kb), deletion of which could be rescued by tethering NONO, SFPQ, or FUS, can interact with these proteins. We thus carried out *in vitro* RNA pulldown from nuclear extracts (NE) using biotinylated sense or antisense RNA fragments of the NEAT1\_2 subdomains. By Coomassie Brilliant Blue (CBB) staining of the co-precipitated proteins, specific proteins were precipitated with the sense RNAs but not with the antisense RNAs (Figure 7A). Immunoblotting analysis revealed that three essential PSPs (NONO, SFPQ, and RBM14) were specifically precipitated with the sense RNA. Interestingly, FUS nonspecifically bound both to the sense and antisense RNAs (Figure 7B). Next we investigated the binding preferences of PSPs for sense RNA fragments from the middle

domain of NEAT1\_2. CBB staining and immunoblotting revealed strong binding of NONO and SFPQ to 12-13 kb and 13-14 kb RNAs, which contain the functional subdomain and its adjacent domain, respectively (Figure 7C and 7D). Both proteins also bound substantially to other sense RNAs (Figure 7D). Moreover, CLIP-seq data indicated that NONO, SFPQ, and FUS bound numerous sites on NEAT1\_2, some of which overlapped with the NEAT1\_2 subdomains (Figure S7A). Together, these data strongly suggest that the binding preference of these PSPs with the sense RNA reflects the requirement of the NEAT1\_2 subdomains for paraspeckle formation *in vivo*.

During the RNA pulldown experiments (Figure 7A), we often observed magnetic bead aggregates visible to the eye in the presence of NEAT1 12–13k sense RNAs (Figure S7B). This suggests the ability of the NEAT1\_2 RNA fragments, which strongly bind NONO and SFPQ, to facilitate higher-order assembly *in vitro*. We quantified the bead aggregation by microscopy, which enabled us to detect single magnetic beads conjugated with RNAs (Figure S7C and S7D). Upon mixing 12–13k sense RNAs, but not 12–13k antisense RNAs, with NE, large bead aggregates formed (Figure 7E). Importantly, without NE, 12–13k sense RNAs themselves did not induce bead aggregates (Figure 7E), suggesting that proteins interacting with the sense RNAs are required for the aggregate formation. When we used washing buffer containing 10% 1,6-HD, the aggregates were disrupted and the size of the aggregates induced by the sense and antisense RNAs became similar (Figure 7E). In addition, we washed the beads with washing buffer containing 10% 2,5-HD and clearly observed much larger aggregates in the presence of sense RNAs than antisense RNAs (Figure 7E), suggesting that the

proteins dissociated by 1,6-HD washing are involved in the aggregation. We thus examined the binding proteins in these conditions. CBB staining clearly showed a loss of protein binding to the 12–13k sense RNAs after washing with 1,6-HD, but not with 2,5-HD (Figure 7F). NONO, SFPQ, RBM14, and FUS were dissociated from the beads, although FUS dissociation was less prominent than dissociation of the other proteins (Figure 7F). Specific depletion of NONO/SFPQ dimers from NE ( $\Delta$  in Figures 7G and S7E) resulted in loss of interactions between these proteins and 12-13k sense RNA (Figure 7G) and significant reduction in bead aggregate size (Figure 7H). Despite the presence of RBM14 and FUS in  $\Delta$ NE, they could not bind to 12-13k sense RNA (Figure 7G), suggesting that NONO/SFPQ are required for binding of RBM14 and FUS to 12-13k sense RNA. These data strongly suggest that the RNA fragments of NEAT1\_2 subdomains can induce higher-order assembly *in vitro* by recruiting proteins including essential PSPs.

## Discussion

We established various NEAT1 deletion mutant lines and revealed the modular domain structure of the NEAT1 lncRNA. Our strategy combining CRISPR/Cas9 technology and HAP1 haploid cells is a powerful one to efficiently establish unambiguous targeted deletions. We found that the 5' terminal domain (0–1 kb) and the 3' TH are both required for stabilization of NEAT1\_2. The decay rate of NEAT1\_2 in  $\Delta 0-1k$  cells was more moderate than that in  $\Delta TH$ , suggesting that the 0–1k domain is required not only for RNA stability but also for efficient transcription. The locations of these functional domains are consistent with the configuration of NEAT1\_2 molecules in paraspeckles; both NEAT1\_2 termini are located at the surface of the body where they can be exposed to attacks by exoribonucleases, and the middle part of NEAT1\_2 that initiates paraspeckle formation is located in the interior core.

The middle domain (8–16.6 kb) is sufficient for paraspeckle assembly and contains at least three functional subdomains (C1: 9.8–12k, C2: 12–13k, and C3: 15.4–16.6k) (Figure 7I). Based on the analysis using deletion mutants derived from mini-NEAT1, deletion of C1 and C2 abolished the paraspeckle assembly. However, sole deletion of C1 or C2 did not abolish paraspeckle assembly, indicating that C1 and C2 act redundantly. On the other hand, C3 might be a distinct functional subdomain since the sole deletion of C3 abolished the function of mini-NEAT1. Similar redundancy in the function of the middle domain is also exerted by the regions outside of the middle domain, as suggested from comparison of the phenotype of deletion mutants constructed from mini-NEAT1 and those from the WT (Figure 5A). Remarkably, deletion of the C3

subdomain (15.4–16.6k) from the WT did not affect paraspeckle integrity, whereas its deletion from mini-NEAT1 did (Figure S2A and S2B). Both  $\Delta 8$ –13k and m13–16.6k lacked the essential region including the C1 and C2 subdomains, and m13–16.6k showed a more severe defective phenotype (<20%) than that in  $\Delta 8$ –13k (>40%) (Figure 5E and 5G). Thus, there are multiple functional redundancies between the middle domain and the outside region, as well as within the middle domain.

The functional subdomains in the middle domain are required for the assembly step in which some of the essential PSPs likely interact to dictate the function. Indeed, tethering of NONO, SFPQ, or FUS but not RBM14 to m13–16.6k rescued paraspeckle assembly, strongly suggesting that the C2 subdomain, which binds essential PSPs including NONO and SFPQ *in vitro* and *in vivo*, shows an ability to facilitate higher-order assembly *in vitro* and functionally recruits these PSPs to initiate paraspeckle assembly. Tethering of mutant NONO showed that the rescue activity requires the NOPS domain, which is required for dimerization with the DBHS proteins (Passon et al., 2012). Interestingly, tethering of mutant NONO lacking the CC maintains the rescue activity even though the CC is required for the polymerization of NONO that likely underlies paraspeckle assembly. Considering that endogenous WT NONO coexists in the cells, the main function of the NEAT1\_2 subdomain is likely to recruit and form NONO dimers with other DBHS proteins such as SFPQ. The subsequent polymerization, which is one of the distinct characteristics of DBHS proteins, can be conducted via the CC of the endogenous NONO in the primary dimer interacting with the C1 and C2 subdomains (Figure 7I). RNA fragments derived from the C2 subdomain and its adjacent region

preferentially bound NONO and SFPQ (Figure 7D), suggesting that the subdomain is enriched for sequence stretches with affinity for PSPs. The other parts of the middle domain including the C1 and C3 subdomains exhibited lower but substantial affinity for NONO and SFPQ (Figure 7D). PAR-CLIP data indicate that PSP-binding sites broadly cover the whole NEAT1\_2 region, and that the C2 subdomain contains multiple prominent binding peaks for NONO and SFPQ (Figure S7A). The difference between the preferential binding sites in the C2 subdomain and other scattered binding sites remains to be elucidated. An intriguing possibility is that the identified subdomains are primary binding sites of PSPs that act as seeds to spread PSPs throughout the broad NEAT1\_2 region via their polymerization described above (Figure 7I). In support of this idea, a previous study using transmission EM showed that SFPQ can coat DNA to form higher-order complexes by polymerization through the CC domain (Lee et al., 2015). As another example, cooperative spreading of HNRNPA1 on pre-mRNA carrying high affinity binding sites was previously reported (Zhu et al., 2001). In addition, our bead aggregation assay supports the idea that the oligomerization of NEAT1 RNPs occurs through the ability of NEAT1\_2 subdomains to recruit essential PSPs including NONO and SFPQ. Moreover, LLPS is likely to be the driving force for paraspeckle formation for three reasons: 1) paraspeckles are sensitive to 1,6-HD treatment in a similar manner to LLPS-formed Cajal bodies, 2) FUS and RBM14 PLDs cause *in vitro* hydrogel formation and are required for *in vivo* paraspeckle formation (Hennig et al., 2015), and 3) PSPs rapidly move into and out of the paraspeckle (Fox et al., 2002; Mao et al., 2011), and thus the paraspeckle is highly dynamic, which is one of the characteristics of phase-separated bodies. We found

that the tethering of FUS can rescue paraspeckle formation in the m13–16.6k mutant. Since the PLD of FUS is required for paraspeckle assembly and phase separation, the tethered FUS might recruit multiple PSPs and/or form FUS oligomers through its PLD, leading to paraspeckle assembly via phase separation. Consistently, we found that treatment with concentrations of 1,6-HD that disrupted paraspeckle structure also disturbed interactions between NONO and the PLD in RNA binding proteins including FUS (Figure 4L).

Several reports indicate the importance of multivalent interactions in phase separation (Banani et al., 2017; Shin and Brangwynne, 2017). As discussed above, multiple redundant NEAT1\_2 middle domains including the identified functional subdomains likely serve as multiple binding sites for essential PSPs including DBHS proteins (e.g., NONO and SFPQ) and proteins with PLDs (e.g., FUS and RBM14), which have an ability for multivalent interactions. For example, DBHS proteins can bind RNAs via RRM, dimerize via the NOPS domain, and reversibly polymerize via the CC domain (Knott et al., 2016). In addition, the PLD promotes multivalent and transient interactions (Kroschwald et al., 2015). Recently, RNA self-assembly was shown to contribute to formation of phase-separated stress granules (Van Treeck et al., 2018), raising the intriguing possibility that self-assembly of NEAT1 at its transcription site may contribute to paraspeckle integrity. Therefore, it appears that various multivalent interactions between RNAs and proteins, and proteins and proteins, through NOPS, CC, and PLD, and presumably via self-assembly of NEAT1, contribute to the formation of this massive, dynamic, highly ordered NB. It is tempting to speculate that the NEAT1 domains

redundantly required for paraspeckle assembly reduce a critical NEAT1 concentration threshold for phase separation to form paraspeckles.

NEAT1\_2 shows a characteristic spatial organization within the paraspeckle. Our observation of aberrant localization of the 3' terminus of mini-NEAT1 to the core of the paraspeckle suggests the presence of NEAT1 subdomain required for NEAT1\_2 spatial organization (Figure 4E). Further NEAT1\_2 deletion analysis suggested that the 16.6-22.6 kb region, which is distinct from the paraspeckle assembly domain, is required for shell localization of NEAT1\_2 3' terminus (Figure S4B-S4E). This observation will facilitate further research into the biological functions of paraspeckles related to NEAT1\_2 spatial organization. Although we revealed several NEAT1 functional domains, NEAT1 likely possesses additional domains required for other biological functions. Since the number of reports on the normal and disease-related functions of NEAT1 is increasing, our mutant HAP1 cell lines will be a valuable resource to investigate further the links between NEAT1 domains and these functions.

A number of arcRNAs have been reported in various eukaryotic species (Chujo et al., 2016). Using two different approaches, we recently found additional arcRNA candidates as well as NBs built on putative arcRNAs (Chujo et al., 2017; Mannen et al., 2016). Data on the characterized arcRNAs raise the intriguing possibility of a common molecular mechanism underlying arcRNA function (Chujo et al., 2016). Here, we showed that multiple redundant NEAT1\_2 subdomains provide the scaffold for paraspeckle assembly by recruiting NONO dimers that initiate oligomerization of DBHS proteins leading to massive RNP complex formation via phase separation (Figure 7I). It is also

possible that other arcRNAs recruit distinct sets of RBPs that induce phase separation through their repetitive sequences. Indeed, the C3 subdomain overlaps with a repetitive sequence that is strongly bound by NONO and SFPQ *in vivo* (Figures S3B and S7A). In addition, the C1 subdomain, which interacts with SFPQ and NONO *in vitro* and *in vivo*, contains several highly similar repetitive sequences (Figure S3B; Figure 7A and 7B; Figure S7A). Thus, it may be interesting to investigate the combination of RBPs and RNA domains that can efficiently induce phase separation, which would enable us to design arcRNAs using the modular RNA domains. Such fundamental understanding of RNA domains and functional machineries will lead us to deeper understandings of the mechanism, biological roles, and evolution of RNA-instructed architecture of subcellular structures.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - HAP1 cell line
- METHOD DETAILS
  - Plasmids
  - Genome editing using CRISPR/Cas9
  - Plasmid transfection
  - RNA-FISH and immunofluorescence
  - Single molecule FISH (smFISH)
  - Reverse transcription-quantitative PCR (RT-qPCR)
  - RNase protection assay (RPA)
  - 5-bromouridine pulse-chase analysis
  - Immunoblotting
  - Electron microscopy
  - PAR-CLIP and mapping of CLIP-seq data
  - Immunoprecipitation
  - RNA pulldown
  - Bead aggregation assay

- Affinity depletion from NE
- QUANTIFICATION AND STATISTICAL ANALYSIS
- DATA AND SOFTWARE AVAILABILITY

## **STAR METHODS**

### **CONTACT FOR REAGENT AND RESOURCE SHARING**

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Tetsuro Hirose (hirose@igm.hokudai.ac.jp).

### **EXPERIMENTAL MODEL AND SUBJECT DETAILS**

#### **HAP1 cell line**

HAP1 cells were maintained in IMDM (Gibco) supplemented with 10% FBS (Gibco) and Penicillin-Streptomycin (Nacalai Tesque). HAP1 mutant cell lines established in this study are listed in Supplemental Table S1. HeLa cells and U-2 OS cells were maintained in DMEM containing high glucose and pyruvate (Gibco) supplemented with 10% FBS (Serana) and Penicillin-Streptomycin (Gibco).

### **METHOD DETAILS**

#### **Plasmids**

To express Cas9 and two sgRNAs from one vector, the PX330 vector bearing an additional U6-driven sgRNA expression cassette (PX330-B/B) was constructed by

inserting a second sgRNA expression cassette containing a Bsal sgRNA cloning site that was amplified from the PX330 vector bearing a Bsal sgRNA cloning site instead of the original BbsI site by primers (Forward: TCTAGAGAGGGCCTATTTCCCATGAT, Reverse: TCTAGAGCCATTTGTCTGCAGAATTG) into the XbaI site of the PX330 plasmid (Addgene). pcDNA5/FRT/TO/i/FLAG vector was generated by inserting the Ftz intron into the AflIII site of pcDNA5/FRT/TO/FLAG used previously (Yamazaki et al., 2012). To construct MCP fusion protein expression vector (pcDNA5/FRT/TO/i/FLAG/MCP), the MCP gene was inserted into the KpnI site of pcDNA5/FRT/TO/i/FLAG. To generate FLAG-MCP-GFP-NLS, FLAG-MCP-NONO, FLAG-MCP-SFPQ, FLAG-MCP-FUS, and FLAG-MCP-RBM14 expression vectors, the MCP-GFP-NLS was inserted into the KpnI site of pcDNA5/FRT/TO/i/FLAG, the NONO ORF was inserted using the BamHI and XhoI sites of pcDNA5/FRT/TO/i/FLAG/MCP, and the SFPQ, FUS, or RBM14 ORF was inserted into the EcoRV site in pcDNA5/FRT/TO/i/FLAG/MCP. To construct plasmids expressing FLAG-MCP-NONO mutants including NONO $\Delta$ NOPS ( $\Delta$ 219–272aa), NONO $\Delta$ CC (1–312aa fused with SV40 NLS), and NONO $\Delta$ PLD ( $\Delta$ 1–52aa), the deletions were introduced by a conventional cloning or site-directed mutagenesis approach. To construct the knock-in vector for 6 $\times$ MS2BS, knock-in sgRNA targeting sites (GCATCGTACGCGTACGTGTT) were inserted into both sides of the 6 $\times$ MS2BS by PCR amplification and the PCR fragments were sub-cloned into the pCR-Blunt II-TOPO vector (Thermo Fisher Scientific).

## **Genome editing using CRISPR/Cas9**

PX330 or PX330-B/B was used for genome editing of the HAP1 cells. sgRNAs targeting NEAT1 were designed using the CRISPR Design website ([crispr.mit.edu](http://crispr.mit.edu)) (Table S1). Using PX330-B/B, two sgRNAs were cloned into BbsI and BsaI/Eco31I sites according to the methods described previously (Ran et al., 2013). To delete the portions of NEAT1, the PX330-B/B plasmids (2  $\mu$ g) containing two sgRNA sequences were co-transfected with pcDNA6/TR plasmids (0.2  $\mu$ g) containing the blasticidin resistance gene (Thermo Fisher Scientific) into HAP1 cells ( $1.5 \times 10^6$  cells) by Nucleofector Kit V (Lonza) with a Nucleofector device (Lonza) using program "X-005" according to the manufacturer's instructions. For knock-in of the 6xMS2 site into the NEAT1 locus, three kinds of plasmids including 1) PX330-B/B (2  $\mu$ g) expressing two sgRNAs (a knock-in sgRNA to cut both sides of 6xMS2BS on the knock-in vector to facilitate genomic integration: GCATCGTACGCGTACGTGTT, and a sgRNA targeting the NEAT1 locus: TACCGCATATCTGTGTACAT), 2) a knock-in vector for 6xMS2BS (2  $\mu$ g), and 3) pcDNA6/TR plasmids (0.2  $\mu$ g) were transfected, similarly using a Nucleofector device (Lonza). To establish NONO knockout HAP1 cells, the sgRNA (see Supplemental Table S1) was selected from GeCKO v2 libraries (Sanjana et al., 2014) and cloned into the BbsI site of PX330 (Addgene). The plasmids (2  $\mu$ g) and pcDNA6/TR plasmids (0.2  $\mu$ g) were transfected similarly using a Nucleofector device (Lonza). To enrich the plasmid-transfected cells, the HAP1 cells were treated with 20  $\mu$ g/ml blasticidin (InvivoGen) for 3 days, starting 1 day after transfection. Subsequently, the cells were diluted into 96-well plates for selection of single clones. The selected clones were lysed by proteinase K treatment (200  $\mu$ g/ml proteinase K [Roche], 20 mM Tris-HCl pH 8.0, 5 mM EDTA, 400

mM NaCl, and 0.3% SDS at 55°C for 1 h, followed by proteinase K inactivation; 95°C for 15 min). Then, the lysates were subjected to PCR analysis to amplify the genomic regions flanking the guide RNA target sites for detecting deletions or insertions using KOD FX Neo enzyme (TOYOBO). To detect small deletions in the NONO knockout cell lines, a T7 endonuclease I (NEB) cleavage assay was performed on the amplified PCR products. The indel-positive clones were further confirmed by sequencing. Finally, the absence of protein expression was confirmed by immunoblotting.

### **Plasmid transfection**

For MS2 tethering and immunoprecipitation (IP) experiments, HAP1 cells ( $1.8 \times 10^5$  cells/well) were cultured on coverslips (ZEISS) in a 6-well plate overnight and plasmids (1.5  $\mu$ g) were transfected using 5  $\mu$ l of TransIT LT-1 (Mirus). The cells were cultured for 48 h, including 6 h of MG132 treatment. Then, the cells were subjected to each experiment.

### **RNA-FISH and immunofluorescence**

RNA-FISH and immunofluorescence were performed as previously described (Kawaguchi et al., 2015; West et al., 2016). The RNA probes were synthesized using T7 or SP6 RNA polymerase and a digoxigenin (DIG) or a fluorescein (FITC) RNA labeling kit (Roche). Linearized plasmids (1  $\mu$ g) containing a NEAT1 fragment (see Supplemental Table S2 for details) were used as templates. For confocal microscopic analyses, the cells were grown on coverslips (Matsunami; micro cover glass; 18 mm round; thickness,

0.16–0.19 mm) and fixed with 4% paraformaldehyde/PBS at room temperature for 10 min. Then, the cells were washed with 1× PBS, permeabilized with 0.5% Triton X-100/PBS for 5 min, and washed three times with 1× PBS. Subsequently, the cells were dipped in 100% ethanol for 5 min and then dried. Dehydrated coverslips were incubated with pre-hybridization solution (50% formamide, 1× Denhardt's salt [Sigma-Aldrich], 2× SSC, 100 mM EDTA, 100 µg/ml yeast tRNA, and 0.01% Tween 20) at 55°C for 1 h and then incubated with hybridization solution (50% formamide, 1× Denhardt's salt [Sigma-Aldrich], 2× SSC, 100 mM EDTA, 100 µg/ml yeast tRNA, 0.01% Tween 20, and 5% Dextran sulfate [Sigma-Aldrich]) containing DIG- or FITC-labeled RNA probes (final concentration: 100 ng/coverslip) at 55°C overnight. After hybridization, the coverslips were washed twice with prewarmed wash buffer (50% formamide, 2× SSC, and 0.1% Tween 20) at 55°C for 15 min. Then, excess RNA probes were digested with 10 µg/ml RNase A (Nacalai Tesque) in NTET buffer (10 mM Tris-HCl [pH 8.0], 1 mM EDTA, 500 mM NaCl, and 0.1% Tween 20) at 37°C for 30 min. The coverslips were washed with buffer (2× SSC and 0.01% Tween 20) at 55°C for 15 min and washed twice with a second buffer (0.1× SSC and 0.01% Tween 20) at 55°C for 15 min. The coverslips were subsequently washed with TBST (1× TBS containing 0.1% Tween 20) and incubated with 1× blocking solution (Blocking reagent [Roche] and TBST) for blocking at room temperature for 1 h. Then, the coverslips were incubated with primary antibodies in 1× blocking solution at room temperature for 1 h, washed three times with TBST for 5 min, incubated with secondary antibodies in 1× blocking solution at room temperature for 30 min, and washed three times with TBST for 5 min. The coverslips were mounted with

VECTASHIELD Hard Set Mounting Medium with DAPI (Vector). Confocal images were acquired using a confocal laser scanning microscope FV1000D (Olympus).

For super-resolution microscopic analyses using Structured Illumination Microscopy (SIM), the cells were grown on coverslips (ZEISS; Cat#0109030091; 18 × 18 mm; thickness no. 1.5H,  $170 \pm 5 \mu\text{m}$ ) and fixed with 4% paraformaldehyde/PBS at 4°C overnight. The coverslips were subsequently washed with 1× PBS and then Milli-Q water. Then, the coverslips were treated with 0.2 N HCl at room temperature for 20 min and washed with Milli-Q water. The coverslips were incubated in Proteinase K buffer (10 mM Tris-HCl [pH 8.0] and 10 mM EDTA [pH 8.0]), then an equal volume of 2× Proteinase K solution (3.3  $\mu\text{g/ml}$  Proteinase K [Roche], 10 mM Tris-HCl [pH 8.0], and 10 mM EDTA [pH 8.0]) was added and the coverslips were incubated at 37°C for 7 min. To stop the Proteinase K reaction, the coverslips were treated with 0.2% glycine/PBS at room temperature for 10 min. Then, the coverslips were washed with 1× PBS and fixed again with 4% paraformaldehyde/PBS at room temperature for 10 min. The coverslips were washed twice with 1× PBS, once with Milli-Q water, and once with 100% ethanol for 5 min, and air-dried. Rehydrated coverslips were pre-hybridized with pre-hybridization solution at 55°C for 1 h and then incubated with hybridization solution containing DIG- or FITC-labeled RNA probes (final concentration: 500 ng/coverslip) at 55°C overnight. After hybridization, the coverslips were proceeded to the washing, blocking, and immunofluorescence steps of the conventional method described above. After the incubation with secondary antibodies, the coverslips were washed with TBST and 1× PBS, and then fixed again with 4% paraformaldehyde/PBS at room temperature for 10

min. Then, the coverslips were washed once with PBS and twice with TBST. The coverslips were mounted with 97% TDE (Sigma-Aldrich) containing DABCO 33-LV (Sigma-Aldrich) as an anti-fade reagent. SIM images were captured using ELYRA PS.1 (ZEISS) with 100× objective lens as previously described (West et al., 2016).

For the treatment with 1,6-Hexanediol or 2,5-Hexanediol, 1,6-Hexanediol or 2,5-Hexanediol dissolved in the HAP1 culture medium was added to cells at room temperature for 5 min. After these treatments, the cells were fixed with 4% paraformaldehyde/PBS followed by the RNA-FISH and immunofluorescence procedure described above. The antibodies used in the RNA-FISH and immunofluorescence are listed in the key resources table. NEAT1 FISH probes against 5' (+1 to +1000), middle (+12786 to +13811), 15.4–16.6k (+15401 to +16612), and 3' regions (+21743 to +22580) were transcribed as antisense RNAs.

### **Single molecule FISH (smFISH)**

Stellaris FISH probes to the NEAT1 5' region (LGC Biosearch Technologies) and NEAT1 middle region (LGC Biosearch Technologies) were used for single molecule FISH (smFISH). The smFISH was performed according to the manufacturer's instructions. The cells were cultured on coverslips (18 mm round) and then fixed with 4% paraformaldehyde/PBS at room temperature for 10 min. Then, the fixed cells were washed twice with 1× PBS and immersed in 70% ethanol for at least 1 h at 4°C for permeabilization. After removing the 70% ethanol, wash buffer (2× SSC and 10% formamide) was added and the coverslips were incubated for 5 min. Then, the 50 µl

hybridization solutions (2× SSC, 100 mg/ml dextran sulfate, and 10% formamide) containing Stellaris NEAT1 probes (final concentration: 125 nM) were dropped onto the coverslips in the humidified chamber, and were incubated for 16 h in the dark. The coverslips were washed with wash buffer at 37°C for 30 min and with 2× SSC at room temperature for 5 min. The coverslips were mounted with VECTASHIELD Hard Set Mounting Medium with DAPI (Vector).

### **Reverse transcription-quantitative PCR (RT-qPCR)**

Total RNAs were purified with TRI reagent (Molecular Research Center) as previously described (Chujo et al., 2017). Briefly, homogenates of cells with TRI reagent were heated at 55°C for 20 min with 1000 rpm agitation. Total RNAs were purified according to the manufacturer's instructions. To quantify NEAT1\_2 levels, the total RNAs (1 µg) were reverse-transcribed by High Capacity cDNA Reverse Transcription Kit (Thermo Fisher Scientific) with random hexamer primers. To quantify NEAT1\_1 specifically, cDNA was reverse-transcribed using dT<sub>20</sub> primers. Aliquots of cDNA were amplified by qPCR using LightCycler 480 SYBR Green I Master reagent (Roche) according to the manufacturer's instructions (see Table S2 for primer sets).

### **RNase protection assay (RPA)**

RNase protection assays (RPAs) were performed using RPAIII kit (Thermo Fisher Scientific). Total RNAs were purified from the HAP1 cell lines using TRI reagents according to the manufacturer's instructions with a modification of heat treatment at 55°C

for 20 min with 1000 rpm agitation. The total RNA (5 µg) was hybridized with <sup>32</sup>P-labeled antisense probes (NEAT1\_1 and NEAT1\_2 boundary region, and U12) synthesized using T7 RNA polymerase (Takara) (Naganuma et al., 2012). RNase A/T1 digestion was performed to eliminate unhybridized single-stranded RNA probes. The protected RNA fragments were separated on 6% PAGE gels containing 7 M urea. Radioactive RNA bands were visualized and quantified with an FLA-7000 analyzer (Fuji).

### **5-bromouridine pulse-chase analysis**

The 5-bromouridine pulse-chase analyses were performed using BRIC kit (MBL) according to the manufacturer's instructions. The HAP1 cells were seeded at 1 million cells per 10 cm dish and cultured overnight. Then, BrU solution was added and cultured for BrU-labeling for 24 h. The cells were washed three times with 1× PBS, and then fresh culture medium was added. Samples were collected at several time points. The total RNAs from the samples were purified using TRI reagent (MRC) as described above. The purified BrU-labeled RNAs were immunoprecipitated with anti-BrdU monoclonal antibody immobilized on protein G magnetic beads. The immunoprecipitated RNAs were subjected to reverse transcription-quantitative PCR (RT-qPCR) analyses using GAPDH as a loading control.

### **Immunoblotting**

Cells were lysed in immunoprecipitation (IP) lysis buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 1% Triton X-100, and complete EDTA-free protease inhibitor cocktail [Roche])

and then disrupted by five pulses of sonication for 5 s. The cell lysates were cleared by centrifugation, and the protein concentration was determined using the Bradford method. After adding SDS sample buffer, the samples were heated at 95°C for 5 min and separated by SDS-PAGE. After electrophoresis, the proteins were transferred to FluoroTrans W membrane (PALL) by electroblotting. The antibodies used are listed in the key resources table.

### **Electron microscopy**

Ultrastructural studies were carried out on ultra-thin sections of Lowicryl K4M-embedded cell pellets as previously described (Souquere and Pierron, 2015). Duplicated samples of WT and mutant HAP1 cells grown in the absence or presence of 5  $\mu$ M MG132 were fixed *in situ* with 1.6% glutaraldehyde (Electron Microscopic Sciences) or 4% formaldehyde (Electron Microscopic Sciences), scraped off from plastic containers, and centrifuged. Cell pellets were equilibrated in 30% methanol and deposited in a Leica EM AFS2/FSP automatic reagent handling apparatus (Leica Microsystems). Lowicryl polymerization under UV was performed for 40 h at -20°C and 40 h at +20°C.

For immunogold electron microscopy (I-EM), ultra-thin sections were incubated at room temperature for 1 h with the primary antibody (mouse anti-NONO [BD Biosciences, 1/20 dilution in PBS] or rabbit anti-BRG1 [Bethyl Laboratories, 1/25 dilution in PBS]) and for 30 min with the secondary anti-mouse IgG or anti-rabbit IgG antibody coupled to 10 nm gold particles (BBInternational). Thin sections were briefly contrasted with uranyl acetate and analyzed with a Tecnai Spirit (FEI). Digital images were taken

with a SIS MegaviewIII charge-coupled device camera (Olympus). Short axes (Sx), long axes (Lx), and surface areas of nuclear body sections were determined with AnalySIS (Olympus Soft Imaging Solutions).

For high resolution *in situ* hybridization (EM-ISH), ultra-thin sections of formaldehyde-fixed cells were occasionally pre-treated with protease (0.2 mg/ml) for 15 min at 37°C to enhance access of the biotinylated DNA probes to the NEAT1 targets. Hybridization conditions and detection of RNA/DNA hybrids with goat anti-biotin antibody conjugated to 10 nm gold particles (BBI International) were as previously described (Souquere et al., 2010; Souquere and Pierron, 2015).

DNA probes were PCR-amplified DNA fragments biotinylated by nick translation (Roche) with biotin-16-dUTP (Roche) instead of TTP in the reaction mix. The NEAT1\_5' DNA probe (+230–1721 and +1751–3244) and D2 probe (+12841–14160 and +14735–15897) were as previously described (Souquere et al., 2010). The NEAT1 5' end probe (+1 to +1000) was nick-translated for 40 min at 15°C and following denaturation hybridized at 60°C for 120 min.

### **PAR-CLIP and mapping of CLIP-seq data**

PAR-CLIP experiments were performed based on the method of Hafner *et al.* (Hafner et al., 2010). Cells cultured in normal growth media were supplemented with 4SU for 14 h. Cells were washed with PBS and then irradiated with 365 nm UV at 0.15 J/cm<sup>2</sup> in a Stratalinker 2400. Cells were harvested, lysed in NP-40 lysis buffer, and passed through a 32-gauge needle. Lysates were cleared by centrifugation and then incubated with

RNase T1 (Fermentas). Lysates were precleared using Protein G Dynabeads (Invitrogen), after which RNA-protein complexes containing NONO were immunoprecipitated with NONO antibody (Souquere et al., 2010) conjugated to Protein G Dynabeads (Thermo Fisher Scientific). Following IP, beads were washed and then resuspended in dephosphorylation buffer and Calf intestinal alkaline phosphatase (New England Biolabs). Beads were then washed and incubated in polynucleotide kinase buffer, and RNA was radiolabeled with  $\gamma$ -<sup>32</sup>P-ATP (Perkin Elmer). Radiolabeled protein-RNA complexes were removed from Dynabeads, then resolved by SDS-PAGE on a NuPAGE Novex 4–12% Bis-Tris Protein Gel (Thermo Fisher Scientific). Using <sup>32</sup>P radiography, RNA-protein complexes were excised from the gel, and electroeluted in D-Tube Dialyzer midi tubes (Merck Millipore). Electroeluant was digested with Proteinase K (Fermentas), and RNA was then extracted from samples using miRNAeasy Micro Kit (Qiagen). RNA extracted from PAR-CLIP was converted into a cDNA library using TruSeq small RNA Kit v2 (Illumina) with a size selection range of 145–327 bp (RNA of 18–200 nt) and Illumina sequencing with 50 bp single reads.

Cutadapt (<http://code.google.com/p/cutadapt>) was used to remove the sequencing adapters, and unique reads were aligned to the human reference genome (hg38) using Bowtie (<http://bowtie-bio.sourceforge.net>) allowing up to one mismatch (-v 1 -m 1). FUS HITS-CLIP and TDP-43 iCLIP data were used by reanalyzing deposited data sets (see key resources table).

## **Immunoprecipitation**

HAP1 cells (m13–16.6k/6×MS2BS) transfected with the indicated plasmids using TransIT LT-1 reagent (Mirus) were cultured for 48 h including the last 6 h with MG132 treatment, washed with cold PBS, and harvested by cell scrapers. The collected cells were lysed in the IP lysis buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 1% Triton X-100, and complete EDTA-free protease inhibitor cocktail [Roche]) by three pulses of sonication for 5 s. The cell lysates were cleared by centrifugation, and the protein concentration was determined using the Bradford method. The lysates were diluted with IP lysis buffer (protein concentration: ~2 mg/ml) and treated with RNase A (1 µg/ml) at 4°C for 1 h. Then, the lysates were mixed with anti-FLAG M2 magnetic beads (Sigma-Aldrich) and rotated at 4°C overnight. The beads were washed five times with IP lysis buffer. The IP samples were recovered by adding SDS sample buffer.

For immunoprecipitation of NONO from HeLa nuclear extract (NE), 25 µl of HeLa NE (CILBIOTECH) spun at 20,000 × *g* for 5 min were mixed with 100 µl of IP buffer (1× PBS, 0.1% Triton X-100, 0.6 mM PMSF, and complete EDTA-free protease inhibitor cocktail [Roche]) and 1 µl RNase A (Nacalai Tesque, 10 mg/ml). Control antibodies (Santa Cruz, normal mouse IgG<sub>1</sub>, 3.75 µg) were conjugated with 15 µl of Dynabeads Protein G Magnetic Beads (Thermo Fisher Scientific) by rotating the tubes at room temperature for 1 hour. The beads were washed with IP buffer to wash out unbound antibodies. Then the antibody-conjugated beads were mixed with the NE solution prepared above and rotated at 4°C for 1.5 hours to preclear the NE solution. For IP, control (Santa Cruz, normal mouse IgG<sub>1</sub>) and anti-NONO mouse monoclonal antibodies (3.75 µg) (BD Biosciences) were conjugated with 15 µl of Dynabeads Protein G Magnetic

Beads (Thermo Fisher Scientific) by rotating the tubes at room temperature for 1 hour. The beads were washed with IP buffer to wash out unbound antibodies. Then the antibody-conjugated beads were mixed with the precleared NE solution and rotated at 4°C for overnight. The beads were washed 5 times with different washing buffers shown below: normal wash buffer (1× PBS, 0.1% Triton X-100, and 0.6 mM PMSF), 1,6-hexanediol wash buffer (normal wash buffer containing 10% of 1,6-hexanediol), 2,5-hexanediol wash buffer (normal wash buffer containing 10% of 1,6-hexanediol), or high salt (HS) wash buffer (normal wash buffer containing 1 M NaCl). For elution, to minimize the elution of IgG from the beads, SDS sample buffer without DTT was added to the washed beads and the beads were kept at room temperature for 20 minutes with agitation using an Eppendorf ThermoMixer (1000 rpm). Then the eluates were recovered and DTT (5 mM) was added. The eluates were boiled and subjected to SDS-PAGE, followed by Coomassie blue staining or immunoblotting.

### **RNA pulldown**

Biotinylated RNAs were synthesized with T7 or SP6 RNA polymerase (Roche) and Biotin RNA labeling Mix (Roche), template DNAs were degraded by DNase I (Thermo Fisher Scientific) treatment, and the biotinylated RNAs were purified by gel filtration column (CENTRI SEP Spin Column; PRINCETON SEPARATIONS). HeLa NE was prepared as previously described (Lee and Green, 1990) or purchased from CILBIOTECH. The NE (50 µl) spun at 20,000 × *g* for 5 min was mixed with 150 µl of RNA pulldown buffer (1× PBS, 0.1% Triton X-100, 0.6 mM PMSF, and complete EDTA-free protease inhibitor

cocktail [Roche]) (final protein concentration was approximately 2 mg/ml). The NE (200  $\mu$ l) was precleared by mixing 20  $\mu$ l Tamavidin2-REV Magnetic Beads (WAKO) washed with RNA pulldown buffer and then rotating at 4°C for 1 h. *In vitro* transcribed RNAs (1  $\mu$ g in 10 ml UltraPure Distilled Water [Thermo Scientific]) were heated at 90°C for 2 min and then cooled on ice for 2 min. An equal volume (10  $\mu$ l) of 2 $\times$  RNA structure buffer (20 mM Tris-HCl pH 7.4, 0.2 M KCl, and 20 mM MgCl<sub>2</sub>) was added and kept at room temperature for 20 min to allow proper secondary structure formation. The folded RNAs were mixed with washed Tamavidin2-REV Magnetic Beads (20  $\mu$ l) and rotated at 4°C for 1 h. The unbound RNAs were washed with cold RNA pulldown buffer, and then the RNA-bound Tamavidin2-REV Magnetic Beads (20  $\mu$ l) were mixed with the precleared NE (200  $\mu$ l). They were rotated at 4°C for 3 h and washed five times with cold wash buffer (1 $\times$  PBS, 0.1% Triton X-100, and 0.6 mM PMSF), and then the bound proteins were eluted at 95°C for 5 min in SDS sample buffer. The eluates were subjected to SDS-PAGE, followed by Coomassie blue staining or immunoblotting.

### **Bead aggregation assay**

The same procedure described in the RNA pulldown method was performed until washing except for the use of washing buffer containing 10% 1,6-HD or 2,5-HD in the indicated samples. To visualize RNAs on beads, Fluorescein-Avidin DN (VectorLab, 1:500 dilution) was added to the mixtures of the RNA bound beads and NE solutions. After the washing, 20  $\mu$ l of the beads suspended in the washing buffer was dropped onto a bottom-glass 8-well chamber and observed by EVOS FL AUTO imaging system (Thermo

Fisher Scientific). The obtained images were analyzed using the function “Analyze Particles” with Fiji (NIH).

### **Affinity depletion from NE**

Depletion of NONO/SFPQ from NE was performed using Ni-NTA beads to capture NONO through its N-terminal histidine rich sequence (Snijders et al., 2015). 100  $\mu$ l of Dynabeads His-Tag Isolation and pulldown beads (Thermo Fisher Scientific) or 100  $\mu$ l of Dynabeads M-280 Sheep anti-mouse IgG (Thermo Fisher Scientific) as a mock depletion were mixed with 250  $\mu$ l of NE (CILBIOTECH) containing 0.1% Triton X-100, 0.6 mM PMSF, and 10 mM Imidazole. After 1 hour rotation at 4°C, the recovered NE was subjected to a second depletion step as described above. After the second depletion, the recovered NEs were subjected to the bead aggregation assay and SDS-PAGE, followed by Coomassie blue staining or immunoblotting.

### **QUANTIFICATION AND STATISTICAL ANALYSIS**

Volocity software (PerkinElmer) was used for quantification of the paraspeckle number, intensity, and size of smFISH signals with an intensity threshold. Fiji software (NIH) was used for line profiles and bead aggregation assay. Prism7 software (GraphPad) was used for the statistical analyses. Kruskal-Wallis test with Dunn’s multiple comparison test was used for Figures 3C, 3E, 5C, 7H, S3D, S3F, and S5D. Mann-Whitney test (two-tailed) was used for Figures 4H, 7E, and S6H. Fisher’s exact test (two-sided) was used for Figure 5E and 5G.

## **DATA AND SOFTWARE AVAILABILITY**

PAR-CLIP data has been submitted to GSE113349. Raw image files were deposited on Mendeley Data: <http://dx.doi.org/10.17632/29k48yvp6b.2>.

## **Supplemental information**

Supplemental information includes seven figures and two tables, and can be found with this article online.

## **Acknowledgments**

The authors thank C. Fujikawa, A. Kubota, E. Fortini, A. Takahashi, J. Iwakiri, and T. Mannen for technical support, and the members of the Hirose laboratory for valuable discussions. This research was supported by grants from the Ministry of Education, Culture, Sports, Science, and Technology of Japan [to TY (26891001, 15K18474, and 17K15058), TC (15J01473 and 17K15084), and TH (26113002, 16H06279, 17H03630, and 17K19335)].

## **Author contributions**

TY and TH conceived and designed this study. TY conducted most of the experiments. SS and GP performed EM analyses. TC performed RPA and immunoblotting. AF, SK, CB, and YSC provided the PAR-CLIP data and related information. SN technically supported SRM. TY, TC, and TH wrote the manuscript.

## Declaration of interests

The authors declare no competing interests.

## Figure legends

**Figure 1.** The TH structure stabilizes NEAT1\_2 lncRNA *in vivo*. (A) The experimental strategy of CRISPR/Cas9-mediated partial deletion of the NEAT1 gene in human haploid HAP1 cells. The plasmid harboring two sgRNAs and Cas9 was transfected into HAP1 cells followed by clonal selection of mutant cells. (B) The sequence and structure of the TH at the 3' terminus of wild-type human NEAT1\_2 (WT) and two deletion mutants ( $\Delta$ TH#1 and  $\Delta$ TH#2). Single dash: Watson-Crick base pair, black square: Hoogsteen interaction, black circle: noncanonical base pair, black triangle: A-minor interaction. (C) Paraspeckles in  $\Delta$ TH cells were detected by FISH of NEAT1 (green) and IF of NONO (magenta), and the nuclei were stained with DAPI. Scale bar: 10  $\mu$ m. (D) Quantitation of paraspeckle-positive cells in C (WT: 100%). (E) Quantitation of NEAT1 isoforms by RT-qPCR in WT and  $\Delta$ TH cells. Data are represented as mean  $\pm$  SD (n = 3). (F) Assessment of NEAT1 isoform levels by RPA in WT and  $\Delta$ TH cells. The quantitated band intensities (WT: 100%) are shown below each panel. U12 is a loading control. (G) Stability of NEAT1\_2 was quantitated by RT-qPCR of the pulse-labeled RNA with 5-bromouridine (5-BrU). Data are represented as mean  $\pm$  SD (n = 3).

**Figure 2.** Identification of NEAT1 domains required for expression, isoform switching,

and paraspeckle formation. (A) The scanning deletion mutants of human NEAT1\_2. The NEAT1 transcripts are shown above with a scale. The blue dashed lines represent the deleted regions. (B) NEAT1 FISH (green) and NONO IF (magenta) with DAPI staining (blue) in NEAT1 deletion mutants. Scale bar: 10  $\mu$ m. (C) Quantitation of paraspeckle-positive cells in the cell lines in B (WT: 100%). (D and E) RT-qPCR of NEAT1 isoforms in WT and  $\Delta$ 0–1k cells (D) or  $\Delta$ 2.1–2.8k and  $\Delta$ 4–5.1k cells (E). Data are represented as mean  $\pm$  SD (n = 3). (F) RPA to detect NEAT1 isoforms in WT and  $\Delta$ 0–1k cells. The quantitated band intensities (WT: 100%) are shown below each panel. U12 is a loading control. (G) Quantitation of the RNA stability of NEAT1\_1 and NEAT1\_2 in WT and the mutant cell line by RT-qPCR of the pulse-labeled RNA with 5-BrU. Data are represented as mean  $\pm$  SD (n = 3). (H) Detection of paraspeckles in the WT and five mutant cell lines by smFISH with a NEAT1 middle probe (upper panels) and NEAT1\_5' probe (lower panels) to detect NEAT1\_2 and both NEAT1\_1 and NEAT1\_2, respectively. Scale bar: 10  $\mu$ m. (I) Quantitation of paraspeckle-positive cells in the mutants used in H (WT: 100%). (J) RT-qPCR of the NEAT1 isoforms in the mutant cells used in H (WT = 1). Data are represented as mean  $\pm$  SD (n = 3).

**Figure 3.** The middle domain of NEAT1\_2 is required for ordered paraspeckle assembly.

(A) The schematics of the WT and the mutants with the larger deletions of two middle regions ( $\Delta$ middle and  $\Delta$ 13–20.2k) are shown as in Figure 2A. The positions of the NEAT1\_5' antisense probes used for EM, and NEAT1\_5' and NEAT1\_3' antisense probes for SRM, are shown below with orange and green lines, respectively. (B)

Detection of paraspeckles in the three cell lines in A. Paraspeckles are detected by RNA-FISH with the NEAT1 5' smFISH probe after treatment with 5  $\mu$ M MG132 for 6 h. Magnified images of the areas indicated by the dashed boxes are shown on the right. Scale bar: 10  $\mu$ m. (C) Quantitative analyses of sizes of the NEAT1 foci in the cells treated with 5  $\mu$ M MG132 for 6 h. (\*\*\*\*:  $P < 0.0001$ , \*\*\*:  $P < 0.001$ , ns: not significant.) (D) EM observation of the paraspeckles in the three cell lines in A. The localization of the 5' terminus of NEAT1 (upper panels) or NONO (lower panels) was detected with a NEAT1-5' probe or NONO antibodies labeled with gold particles, respectively. To enlarge the paraspeckles, the HAP1 cells were treated with 5  $\mu$ M MG132 for 17 h (upper panels). Scale bars: 200 nm. (E) Quantitation of the paraspeckle sizes observed by EM after treatment with 5  $\mu$ M MG132 for 4 h. (\*\*\*\*:  $P < 0.0001$ , \*:  $P < 0.05$ , ns: not significant.) (F) Observation of the *de novo* synthesized paraspeckles in the three cell lines in A by SRM. The cells were treated with DRB (50  $\mu$ M) for 4 h and then released for 1, 2, and 4 h to observe paraspeckles by RNA-FISH using a mixture of NEAT1\_5' and NEAT1\_3' probes (green) and IF of NONO (magenta). Line profiles of the paraspeckles at 4 h after DRB release are shown (right). Scale bar: 500 nm.

**Figure 4.** The middle domain of NEAT1\_2 is sufficient for formation of the ordered paraspeckle structure. (A) The schematics of the WT and mini-NEAT1 mutant (m8–16.6k) are shown with a scale as in Figure 2A. The positions of the NEAT1\_5' and NEAT1\_D2 probes used for EM (orange), and the NEAT1\_5' (green), NEAT1\_mid (magenta), and NEAT1\_15.4-16.6k (magenta) probes used for SRM, are shown below.

(B and C) SRM observation of paraspeckles in WT and mini-NEAT1 cells. Indicated probes as shown in A were used. The cells were treated with 5  $\mu$ M MG132 for 6 h. Scale bar: 500 nm. (D) Quantitation of NEAT1\_2 levels in the WT and mini-NEAT1 by RT-qPCR (WT = 1). Data are represented as mean  $\pm$  SD (n = 3). (E) The model of NEAT1\_2 configuration within paraspeckles in WT and mini-NEAT1 cells. 5' and 3' termini and middle region (mid) of NEAT1\_2 are shown in the two-layered paraspeckle structure. (F and G) EM observation of the paraspeckle structure in mini-NEAT1 cells. Localization of NONO and BRG1 (F), and the 5' terminus and middle (D2) of NEAT1 in the paraspeckles bleached by proteinase K treatment (G). No MG132 treatment was performed in F, whereas treatment with 5  $\mu$ M MG132 for 17 h was performed in G. Scale bar: 100 nm (F), 200 nm (G). (H) Quantitation of paraspeckle sizes. The surface area (left), Sx diameter (middle), and Lx diameter (right) of paraspeckles in MG132-untreated WT and mini-NEAT1 cells are plotted (\*\*\*\*:  $P < 0.0001$ ). (I and J) Paraspeckles in HAP1 cells treated with 0–10% 1,6-HD (I: WT, J: mini-NEAT1) observed by NEAT1 RNA-FISH using NEAT1\_5' probes (green) and by NONO IF (magenta). The cells were treated with 5  $\mu$ M MG132 for 6 h. Scale bar: 10  $\mu$ m. (K) SRM observation of paraspeckle substructures in HAP1 WT and mini-NEAT1 cells in the presence of 1,6-HD or 2,5-HD. Paraspeckle substructures are visualized by NEAT1 FISH with the mixture of NEAT1-5' and NEAT1-3' probes to detect the shell (green) and NONO IF to detect the core (magenta). The cells were treated with 5  $\mu$ M MG132 for 6 h. Scale bar: 500 nm. (L) NONO immunoprecipitation under various washing conditions (Upper panel). A CBB stained SDS-PAGE gel image is shown (Lower panels). Immunoblotting with indicated

antibodies is shown. HS: high salt. Filled circles indicate proteins that show reduced binding to NONO proteins after washing with 1,6-HD.

**Figure 5.** NEAT1\_2 middle domain contains multiple redundant subdomains for paraspeckle assembly. (A) The schematics of the WT and five deletion mutants of the NEAT1\_2 middle domain (upper), as well as those of the WT and eight deletion mutants derived from mini-NEAT1 mutants (lower), are shown with a scale. The positions of the NEAT1\_5' (green line), NEAT1\_3' (green line), and NEAT1\_mid (magenta line) antisense probes for SRM are shown below. (B) Observation of the paraspeckles in the cell in A with the NEAT1\_5' (magenta) smFISH probe. The cells were treated with 5  $\mu$ M MG132 for 6 h. Scale bar: 10  $\mu$ m. (C) Quantitation of the paraspeckle sizes observed in B. (\*\*\*\*:  $P < 0.0001$ , \*\*:  $P < 0.01$ , unindicated pairs were not significantly different.) (D) SRM observation of shell and core substructure of paraspeckles in the WT, the five mutants in A, and FUS KO cells. Paraspeckle substructures are visualized by NEAT1 FISH with the mixture of NEAT1\_5' and NEAT1\_3' probes to detect the shell (green) and NONO IF to detect the core (magenta). The cells were treated with 5  $\mu$ M MG132 for 6 h. (E) Ratio (%) of paraspeckles with ordered structure in the cells observed in D. (\*\*\*\*:  $P < 0.0001$ , compared with the WT.) (F) SRM observation of the paraspeckles in the WT and the eight mutants shown in F. For RNA-FISH, NEAT1\_5' (green), and NEAT1\_mid (magenta) were used as shown on the left. The cells were treated with 5  $\mu$ M MG132 for 6 h. Scale bar: 500 nm. (G) Ratio (%) of paraspeckles with ordered structure in the cells observed in G. (\*\*\*\*:  $P < 0.0001$ , \*:  $P < 0.05$ , ns: not significant.)

**Figure 6.** An essential NEAT1\_2 subdomain for mini-paraspeckle formation can be functionally replaced by artificially tethered essential PSPs. (A) The schematics of the m13–16.6k mutant with or without 6×MS2BS for MCP-PSP tethering experiments. The position of the NEAT1\_5' probe used is shown below. (B) Observation of paraspeckle formation with transfection of MCP-PSPs into m13–16.6k/6×MS2BS (left panels) or m13–16.6k cells (right panels) in the MG132 treatment condition (5  $\mu$ M for 6 h). The rescued paraspeckles, which are visualized by IF of MCP protein with  $\alpha$ Flag antibody (magenta) and NEAT1 FISH (green), are indicated by white arrowheads. Scale bar: 10  $\mu$ m. (C) SRM observation of the rescued paraspeckles with MCP-NONO, MCP-SFPQ, and MCP-FUS corresponding to the results in B. Paraspeckle substructures were visualized by NEAT1 FISH (green) and MCP IF (magenta). Scale bar: 500 nm. (D) Tethering of the MCP-tagged mutant NONO lacking NOPS, CC, or PLD domains as in Figure 6B. Scale bar: 10  $\mu$ m. (E) Quantitation of the cells with rescued paraspeckles in B and D. Data are represented as mean  $\pm$  SD (n = 3). (F) Coimmunoprecipitation of MCP-NONO,  $\Delta$ NOPS, and  $\Delta$ CC mutants with a control, MCP-GFP-NLS. Immunoblotting was performed to detect DBHS proteins in the co-IP samples. GAPDH served as a negative control.

**Figure 7.** NEAT1\_2 middle fragments specifically bind NONO and SFPQ, and induce higher-order assembly *in vitro*. (A and B) *In vitro* RNA pulldown to capture the associated proteins with the RNA fragments derived from the functional NEAT1\_2 subdomains. The

captured proteins were separated by SDS-PAGE and visualized by CBB staining. Asterisks indicate major interacting proteins specific to sense RNAs (A). Four essential PSPs were detected by immunoblotting (B). (C and D) *In vitro* RNA pulldown with various NEAT1\_2 middle RNA fragments. CBB stained SDS-PAGE gel image (C) and immunoblotting analysis of the samples (D) are shown. (E) Bead aggregation assay. Magnetic beads were observed by microscopy (upper panels). Quantification of bead aggregate sizes is shown in the lower panels. It should be noted that 10% 1,6-HD and 2,5-HD moderately affected bead aggregation presumably due to the indirect effects of these reagents in NE. Scale bar: 100  $\mu\text{m}$ . (\*\*\*\*:  $P < 0.0001$ , ns: not significant.) (F) (Upper panel) CBB staining analysis of bound proteins on RNAs used in Figure 7E. Asterisks indicate major interacting proteins specific to sense RNAs. (Lower panels) Immunoblotting analysis of the samples shown in the upper panel. (G) Bead aggregation assay with NONO/SFPQ depleted ( $\Delta$ ) NE. Upper panels show CBB stained SDS-PAGE gels and lower panels show immunoblotting analysis of the samples shown in the upper panel. (Left) NE input samples of mock and  $\Delta$  NE. Asterisks indicate major interacting proteins in pulldown samples with the mock depleted NE. (H) Quantification of bead aggregate sizes in the experiment shown in Figure 7G. (\*\*\*\*:  $P < 0.0001$ ) (I) The updated model for the construction of phase-separated paraspeckles with NEAT1 domains and PSPs (see text for details).

## References

Adriaens, C., Standaert, L., Barra, J., Latil, M., Verfaillie, A., Kalev, P., Boeckx, B.,

Wijnhoven, P.W., Radaelli, E., Vermi, W., *et al.* (2016). p53 induces formation of NEAT1 lncRNA-containing paraspeckles that modulate replication stress response and chemosensitivity. *Nat Med* 22, 861-868.

Banani, S.F., Lee, H.O., Hyman, A.A., and Rosen, M.K. (2017). Biomolecular condensates: organizers of cellular biochemistry. *Nat Rev Mol Cell Biol* 18, 285-298.

Chen, L.L. and Carmichael, G.G. (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.

Chu, C., Zhang, Q.C., da Rocha, S.T., Flynn, R.A., Bharadwaj, M., Calabrese, J.M., Magnuson, T., Heard, E., and Chang, H.Y. (2015). Systematic discovery of Xist RNA binding proteins. *Cell* 161, 404-416.

Chujo, T., Yamazaki, T., and Hirose, T. (2016). Architectural RNAs (arcRNAs): A class of long noncoding RNAs that function as the scaffold of nuclear bodies. *Biochim Biophys Acta* 1859, 139-146.

Chujo, T., Yamazaki, T., Kawaguchi, T., Kurosaka, S., Takumi, T., Nakagawa, S., and Hirose, T. (2017). Unusual semi-extractability as a hallmark of nuclear body-associated architectural noncoding RNAs. *EMBO J* 36, 1447-1462.

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

Engreitz, J.M., Ollikainen, N., and Guttman, M. (2016). Long non-coding RNAs: spatial amplifiers that control nuclear structure and gene expression. *Nat Rev Mol Cell Biol* 17, 756-770.

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

Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano, M., Jr., Jungkamp, A.C., Munschauer, M., *et al.* (2010). Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* 141, 129-141.

Hennig, S., Kong, G., Mannen, T., Sadowska, A., Kobelke, S., Blythe, A., Knott, G.J., Iyer, K.S., Ho, D., Newcombe, E.A., *et al.* (2015). Prion-like domains in RNA binding proteins are essential for building subnuclear paraspeckles. *J Cell Biol* 210, 529-539.

Hirose, T., Virnicchi, G., Tanigawa, A., Naganuma, T., Li, R., Kimura, H., Yokoi, T.,

Nakagawa, S., Bénard, M., Fox, A.H., *et al.* (2014b). NEAT1 long noncoding RNA regulates transcription via protein sequestration within subnuclear bodies. *Mol Biol Cell* 25, 169-183.

Kawaguchi, T., Tanigawa, A., Naganuma, T., Ohkawa, Y., Souquere, S., Pierron, G., and Hirose, T. (2015). SWI/SNF chromatin-remodeling complexes function in noncoding RNA-dependent assembly of nuclear bodies. *Proc Natl Acad Sci U S A* 112, 4304-4309.

Knott, G.J., Bond, C.S., and Fox, A.H. (2016). The DBHS proteins SFPQ, NONO and PSPC1: a multipurpose molecular scaffold. *Nucleic Acids Res* 44, 3989-4004.

Kroschwald, S., Maharana, S., Mateju, D., Malinowska, L., Nuske, E., Poser, I., Richter, D., and Alberti, S. (2015). Promiscuous interactions and protein disaggregases determine the material state of stress-inducible RNP granules. *Elife* 4, e06807.

Lee, K.A., and Green, M.R. (1990). Small-scale preparation of extracts from radiolabeled cells efficient in pre-mRNA splicing. *Methods Enzymol* 181, 20-30.

Lee, M., Sadowska, A., Bekere, I., Ho, D., Gully, B.S., Lu, Y., Iyer, K.S., Trewella, J., Fox, A.H., and Bond, C.S. (2015). The structure of human SFPQ reveals a coiled-coil mediated polymer essential for functional aggregation in gene regulation. *Nucleic Acids Res* 43, 3826-3840.

Lin, Y., Mori, E., Kato, M., Xiang, S., Wu, L., Kwon, I., and McKnight, S.L. (2016). Toxic PR Poly-Dipeptides Encoded by the C9orf72 Repeat Expansion Target LC Domain Polymers. *Cell* 167, 789-802 e712.

Mannen, T., Yamashita, S., Tomita, K., Goshima, N., and Hirose, T. (2016). The Sam68 nuclear body is composed of two RNase-sensitive substructures joined by the adaptor HNRNPL. *J Cell Biol* 214, 45-59.

Mao, Y.S., Sunwoo, H., Zhang, B., and Spector, D.L. (2011). Direct visualization of the co-transcriptional assembly of a nuclear body by noncoding RNAs. *Nat Cell Biol* 13, 95-101.

Naganuma, T., Nakagawa, S., Tanigawa, A., Sasaki, Y.F., Goshima, N., and Hirose, T. (2012). Alternative 3'-end processing of long noncoding RNA initiates construction of nuclear paraspeckles. *EMBO J* 31, 4020-4034.

Nakagawa, S., Shimada, M., Yanaka, K., Mito, M., Arai, T., Takahashi, E., Fujita, Y., Fujimori, T., Standaert, L., Marine, J.C., *et al.* (2014). The lncRNA Neat1 is required for corpus luteum formation and the establishment of pregnancy in a subpopulation of mice. *Development* 141, 4618-4627.

Passon, D.M., Lee, M., Rackham, O., Stanley, W.A., Sadowska, A., Filipovska, A., Fox, A.H., and Bond, C.S. (2012). Structure of the heterodimer of human NONO and paraspeckle protein component 1 and analysis of its role in subnuclear body formation. *Proc Natl Acad Sci U S A* *109*, 4846-4850.

Quinn, J.J., and Chang, H.Y. (2016). Unique features of long non-coding RNA biogenesis and function. *Nat Rev Genet* *17*, 47-62.

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* *8*, 2281-2308.

Sanjana, N.E., Shalem, O., and Zhang, F. (2014). Improved vectors and genome-wide libraries for CRISPR screening. *Nat Methods* *11*, 783-784.

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

Shin, Y., and Brangwynne, C.P. (2017). Liquid phase condensation in cell physiology and disease. *Science* *357*.

Snijders, A.P., Hautbergue, G.M., Bloom, A., Williamson, J.C., Minshull, T.C., Phillips, H.L., Mihaylov, S.R., Gjerde, D.T., Hornby, D.P., Wilson, S.A., *et al.* (2015). Arginine methylation and citrullination of splicing factor proline- and glutamine-rich (SFPQ/PSF) regulates its association with mRNA. *RNA* *21*, 347-359

Souquere, S., Beauclair, G., Harper, F., Fox, A., and Pierron, G. (2010). Highly ordered spatial organization of the structural long noncoding NEAT1 RNAs within paraspeckle nuclear bodies. *Mol Biol Cell* *21*, 4020-4027.

Souquere, S., and Pierron, G. (2015). Ultrastructural analysis of nuclear bodies using electron microscopy. *Methods Mol Biol* *1262*, 105-118.

Stanek, D., and Fox, A.H. (2017). Nuclear bodies: news insights into structure and function. *Curr Opin Cell Biol* *46*, 94-101.

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

Tollervey, J.R., Curk, T., Rogelj, B., Briese, M., Cereda, M., Kayikci, M., König, J., Hortobágyi, T., Nishimura, A.L., Zupunski, V., *et al.* (2011). Characterizing the RNA targets and position-dependent splicing regulation by TDP-43. *Nat Neurosci* *14*, 452-458.

Van Treeck, B., Protter, D.S.W., Matheny, T., Khong, A., Link, C.D., and Parker, R. (2018). RNA self-assembly contributes to stress granule formation and defining the stress granule transcriptome. *Proc Natl Acad Sci U S A* 115, 2734-2739.

Visa, N., Puvion-Dutilleul, F., Bachellerie, J.P., and Puvion, E. (1993). Intranuclear distribution of U1 and U2 snRNAs visualized by high resolution in situ hybridization: revelation of a novel compartment containing U1 but not U2 snRNA in HeLa cells. *Eur J Cell Biol* 60, 308-321.

West, J.A., Mito, M., Kurosaka, S., Takumi, T., Tanegashima, C., Chujo, T., Yanaka, K., Kingston, R.E., Hirose, T., Bond, C., *et al.* (2016). Structural, super-resolution microscopy analysis of paraspeckle nuclear body organization. *J Cell Biol* 214, 817-830.

Wilusz, J.E., JnBaptiste, C.K., Lu, L.Y., Kuhn, C.D., Joshua-Tor, L., and Sharp, P.A. (2012). A triple helix stabilizes the 3' ends of long noncoding RNAs that lack poly(A) tails. *Genes Dev* 26, 2392-2407.

Wutz, A., Rasmussen, T.P., and Jaenisch, R. (2002). Chromosomal silencing and localization are mediated by different domains of Xist RNA. *Nat Genet* 30, 167-174.

Yamazaki, T., Chen, S., Yu, Y., Yan, B., Haertlein, T.C., Carrasco, M.A., Tapia, J.C., Zhai, B., Das, R., Lalancette-Hebert, M., *et al.* (2012). FUS-SMN protein interactions link the motor neuron diseases ALS and SMA. *Cell Rep* 2, 799-806.

Yamazaki, T., and Hirose, T. (2015). The building process of the functional paraspeckle with long non-coding RNAs. *Front Biosci (Elite Ed)* 7, 1-41.

Zhou, Y., Liu, S., Liu, G., Oztürk, A., and Hicks, G.G. (2013). ALS-associated FUS mutations result in compromised FUS alternative splicing and autoregulation. *PLoS Genet* 9, e1003895.

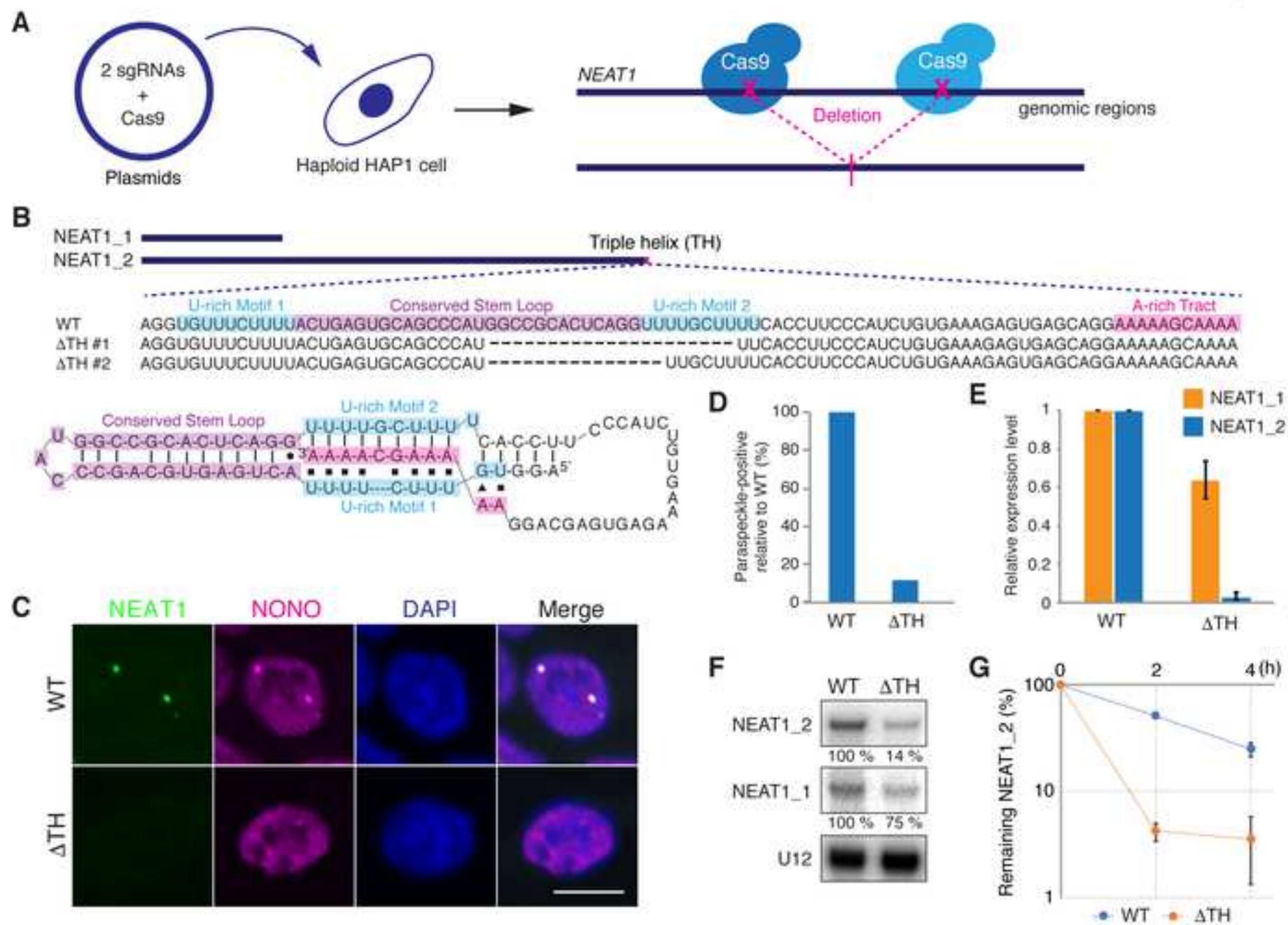
Zhu, J., Mayeda, A., and Krainer, A.R. (2001). Exon identity established through differential antagonism between exonic splicing silencer-bound hnRNP A1 and enhancer-bound SR proteins. *Mol Cell* 8, 1351-1361.

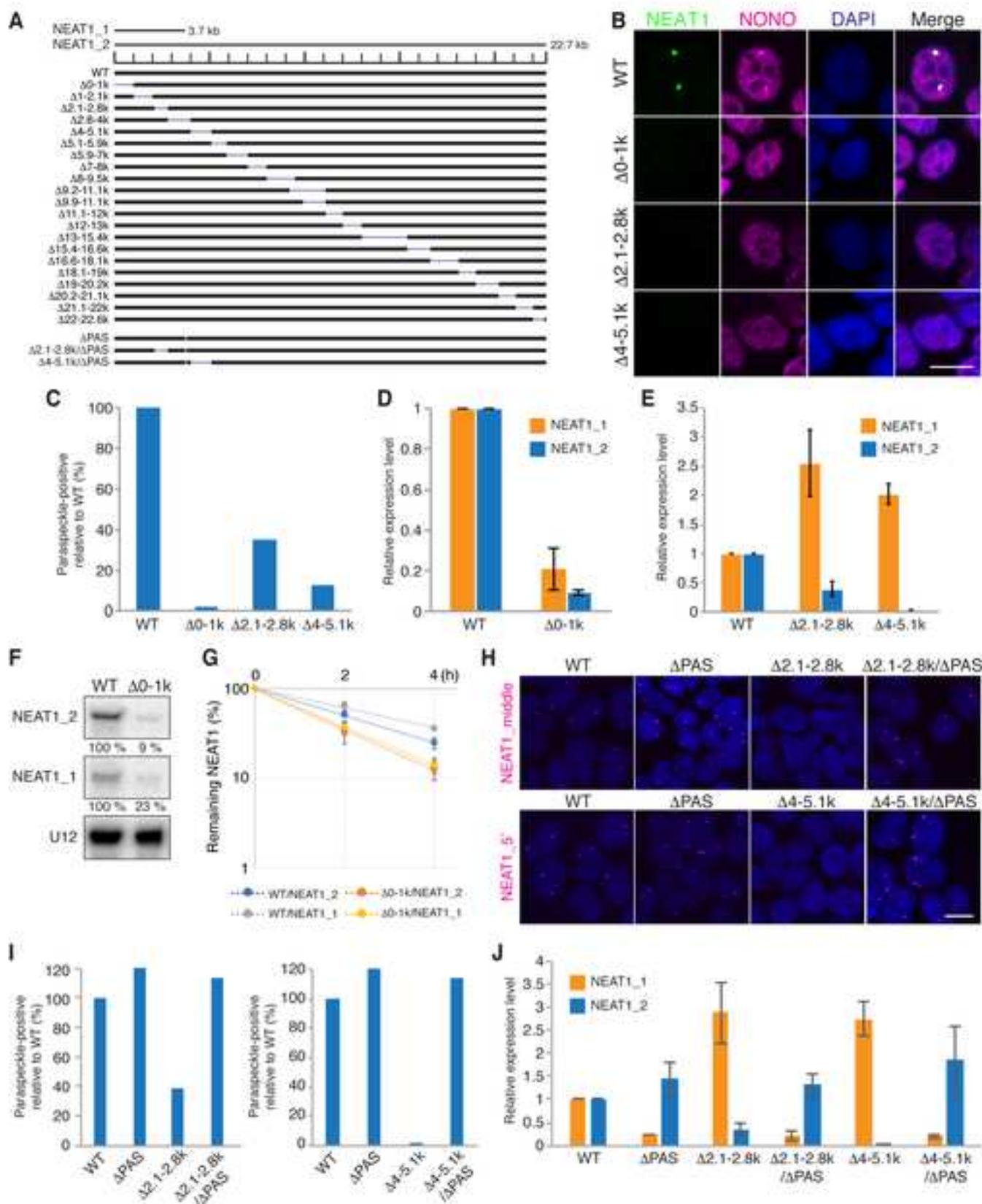
## KEY RESOURCES TABLE

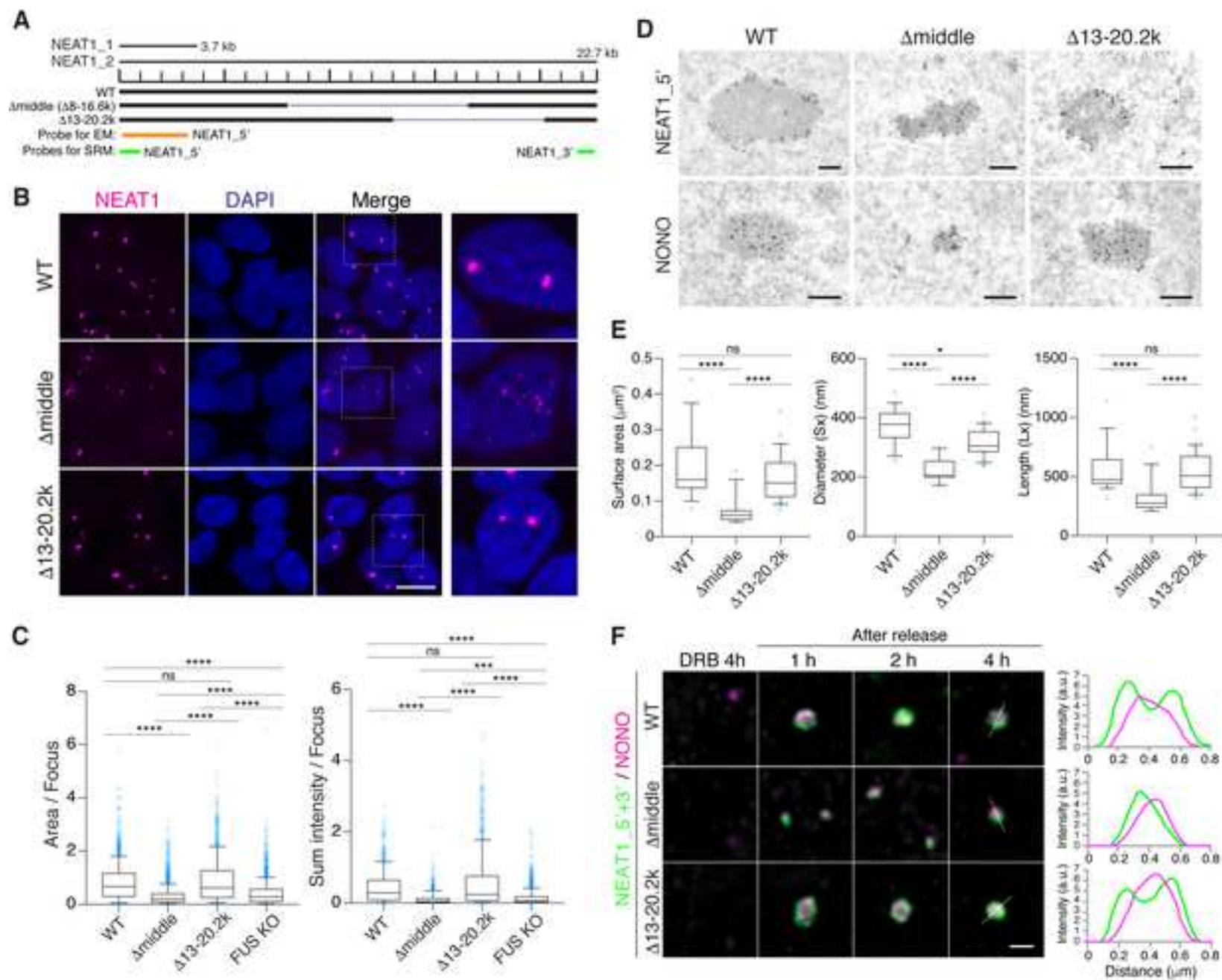
REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse monoclonal anti-NONO (clone 3/p54nrb)	BD Biosciences	Cat#611279
Mouse monoclonal anti-NONO (for PAR-CLIP)	Souquere <i>et al.</i> 2010	N/A
Mouse monoclonal anti-FUS (clone 4H11)	Santa Cruz Biotechnology	Cat#sc-47711
Mouse monoclonal anti-SFPQ (PSF) (clone C23)	MBL	Cat#RN014MW
Rabbit polyclonal anti-RBM14	Bethyl Laboratories	Cat#A300-331A
Rabbit polyclonal anti-PSPC1	Naganuma <i>et al.</i> , 2012	N/A
Rabbit polyclonal anti-BRG1	Bethyl Laboratories	Cat#A300-813A
Mouse monoclonal anti-FLAG (clone M2)	Sigma-Aldrich	Cat#F3165
Rabbit polyclonal anti-GAPDH	Abcam	Cat#ab37168
Mouse monoclonal anti-Digoxigenin (clone 21H8)	Abcam	Cat#ab420
Rabbit polyclonal anti-Fluorescein	Abcam	Cat#ab19491
Anti-mouse IgG, Cy2 conjugate	Abcam	Cat#ab6944
Goat anti-rabbit IgG, Cy2 conjugate	Abcam	Cat#ab6940
Goat anti-mouse IgG, Cy3 conjugate	Merck	Cat#AP124C
Anti-rabbit IgG, Cy3 conjugate	Merck	Cat#AP132C
Anti-mouse IgG, Alexa 488 conjugate	Thermo Fisher Scientific	Cat#A-11029
Anti-rabbit IgG, Alexa 488 conjugate	Thermo Fisher Scientific	Cat#A-11034
Anti-mouse IgG, Alexa 568 conjugate	Thermo Fisher Scientific	Cat#A-11031
Anti-rabbit IgG, Alexa 568 conjugate	Thermo Fisher Scientific	Cat#A-11036
Anti-mouse IgG antibody coupled to 10-nm gold particles	TebuBio	Cat#EM.GMHL10
Anti-rabbit IgG antibody coupled to 10-nm gold particles	TebuBio	Cat#EM.GAR10
Goat anti-biotin antibody conjugated to 10nm gold particles	TebuBio	Cat#EM.GAB10
Bacterial and Virus Strains		
DH5 $\alpha$ Competent Cells	Thermo Fisher Scientific	Cat#18265017
Chemicals, Peptides, and Recombinant Proteins		
Z-Leu-Leu-Leu-al (MG132)	Sigma-Aldrich	Cat#C2211
5,6-Dichlorobenzimidazole 1- $\beta$ -D-ribofuranoside (DRB)	Sigma-Aldrich	Cat#D1916
Blasticidin	InvivoGen	Cat#Ant-bl-1
1,6-Hexanediol	Sigma-Aldrich	Cat#240117
2,5-Hexanediol	Sigma-Aldrich	Cat#H11904
2,2'-Thiodiethanol	Sigma-Aldrich	Cat#166782
DABCO 33-LV	Sigma-Aldrich	Cat#290734
VECTASHIELD Hard Set Mounting Medium with DAPI	Vector laboratories	Cat#H-1500
TRI Reagent	Molecular Research Center	Cat#TR118
cOmplete, EDTA free Protease Inhibitor Cocktail	Roche	5056489001
TransIT LT-1 Reagent	Mirus	Cat#MIR2300
Proteinase K	Roche	Cat#3115828001

Blocking reagent	Roche	Cat#11096176001
Glutaraldehyde (for EM)	Euromedex	Cat#16220
Formaldehyde (for EM)	Euromedex	Cat#19200
Critical Commercial Assays		
High Capacity cDNA Reverse Transcription Kit	Thermo Fisher Scientific	Cat#4368813
Stellaris FISH probe (Human NEAT1_5 with Quaser 570 Dye)	LGC Biosearch Technologies	Cat#SMF-2036-1
Stellaris FISH probe (Human NEAT1_m with Quaser 570 Dye)	LGC Biosearch Technologies	Cat#SMF-2037-1
RPAIII Ribonuclease Protection Assay Kit	Thermo Fisher Scientific	Cat#AM1414
BRIC kit	MBL	Cat#RN1007
T7 Endonuclease I	NEB	Cat#M0302L
Dynabeads protein G	Thermo Fisher Scientific	Cat#10004D
Dynabeads His-Tag Isolation & Pulldown	Thermo Fisher Scientific	Cat#10103D
Anti-FLAG M2 magnetic beads	Sigma-Aldrich	Cat#M8823
Tamavidin2-REV Magnetic Beads	WAKO	Cat#136-18341
Biotin RNA labeling Mix	Sigma-Aldrich	Cat#11685597910
MEGAscript T7 Transcription Kit	Thermo Fisher Scientific	Cat#AM1333
MEGAscript SP6 Transcription Kit	Thermo Fisher Scientific	Cat#AM1330
Nick-translation	Roche	Cat#10976776001
Biotin-16-dUTP	Roche	Cat#11093070910
Deposited Data		
PAR-CLIP data	This paper	GSE113349
Raw image files	This study	Mendeley Data <a href="http://dx.doi.org/10.17632/29k48yvp6b.2">http://dx.doi.org/10.17632/29k48yvp6b.2</a>
TDP-43 iCLIP	Tollervy <i>et al.</i> , 2011	SRA: ERX016992
FUS CLIP-seq	Zhou <i>et al.</i> , 2013	GEO: GSE50178
Experimental Models: Cell Lines		
Human HAP1 cell line	Horizon Discovery	Cat#C631, RRID:CVCL_Y019
HAP1 mutant cell lines	This paper	See Table S1
HAP1 NONO knockout cell line	This paper	See Table S1
HAP1 FUS knockout cell line	Horizon Discovery	Cat#HZGHC001314 c006
U-2 OS cell line	ATCC	Cat#ATCC HTB-96, RRID:CVCL_0042
Oligonucleotides		
Primers for qPCR	Sigma-Aldrich	See Table S2
Recombinant DNA		
Plasmid: pcDNA5/FRT/TO/i/FLAG/MCP/GFP-NLS	This paper	N/A
Plasmid: pcDNA5/FRT/TO/i/FLAG/MCP/NONO	This paper	N/A
Plasmid: pcDNA5/FRT/TO/i/FLAG/MCP/SFPQ	This paper	N/A
Plasmid: pcDNA5/FRT/TO/i/FLAG/MCP/FUS	This paper	N/A

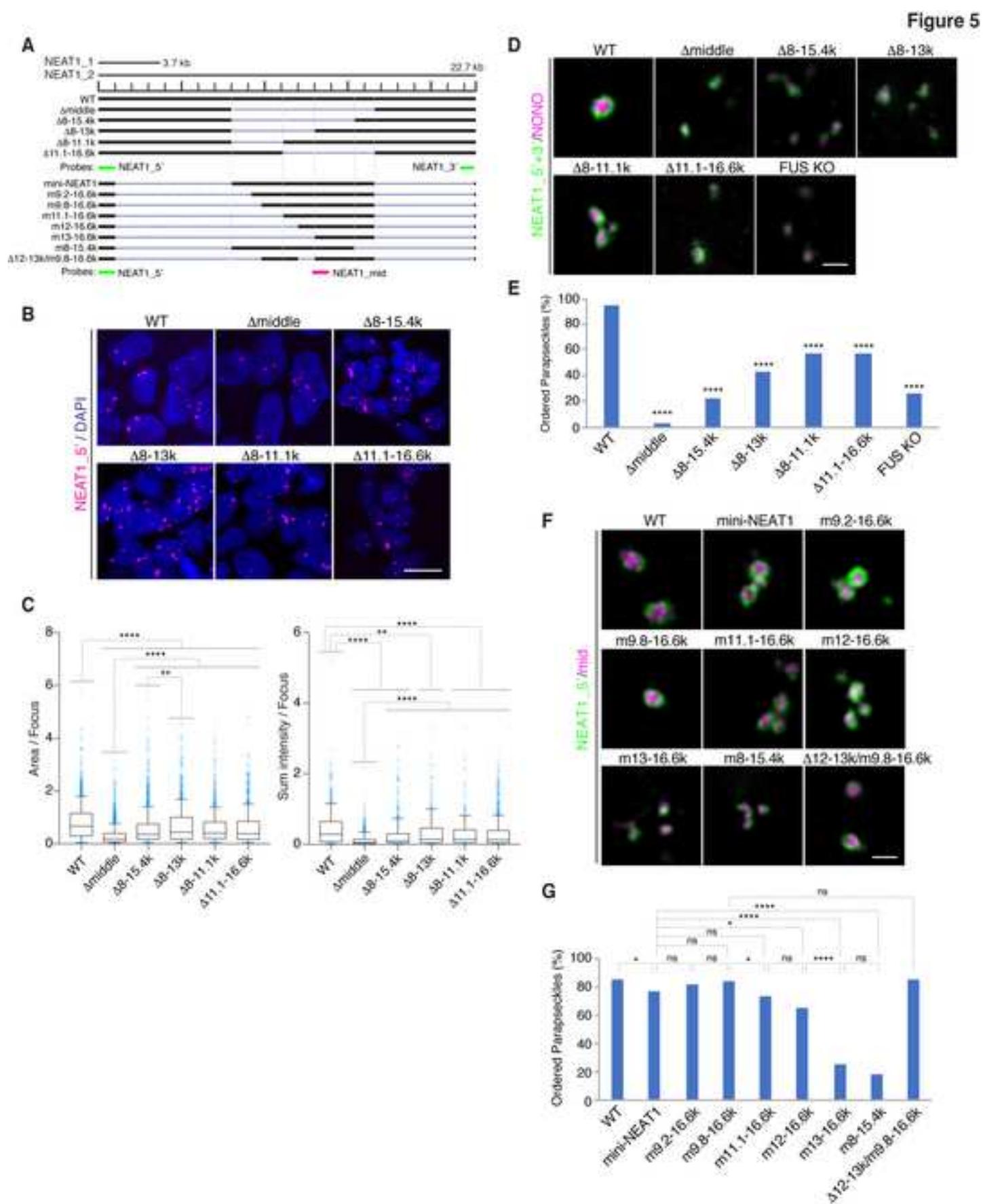
Plasmid: pcDNA5/FRT/TO/i/FLAG/MCP/RBM14	This paper	N/A
Plasmid: pcDNA5/FRT/TO/i/FLAG/MCP/NONO $\Delta$ NOPS	This paper	N/A
Plasmid: pcDNA5/FRT/TO/i/FLAG/MCP/NONO $\Delta$ CC	This paper	N/A
Plasmid: pcDNA5/FRT/TO/i/FLAG/MCP/NONO $\Delta$ PLD	This paper	N/A
Plasmid: PX330-U6-Chimeric_BB-CBh-hSpCas9 (PX330)	Addgene	Plasmid #42230
Plasmid: PX330-B/B	This paper	N/A
pcDNA6/TR	Thermo Fisher Scientific	Cat#V102520
Software and Algorithms		
Velocity 6.3	PerkinElmer	<a href="http://cellularimaging.perkinelmer.com/downloads">http://cellularimaging.perkinelmer.com/downloads</a>
Prism 7	GraphPad Software	<a href="https://www.graphpad.com">https://www.graphpad.com</a>
ImageJ, Fiji, version 1.0	Open source image processing software	<a href="https://fiji.sc/">https://fiji.sc/</a>
Zen	ZEISS	N/A
Metamorph	Molecular Devices	N/A

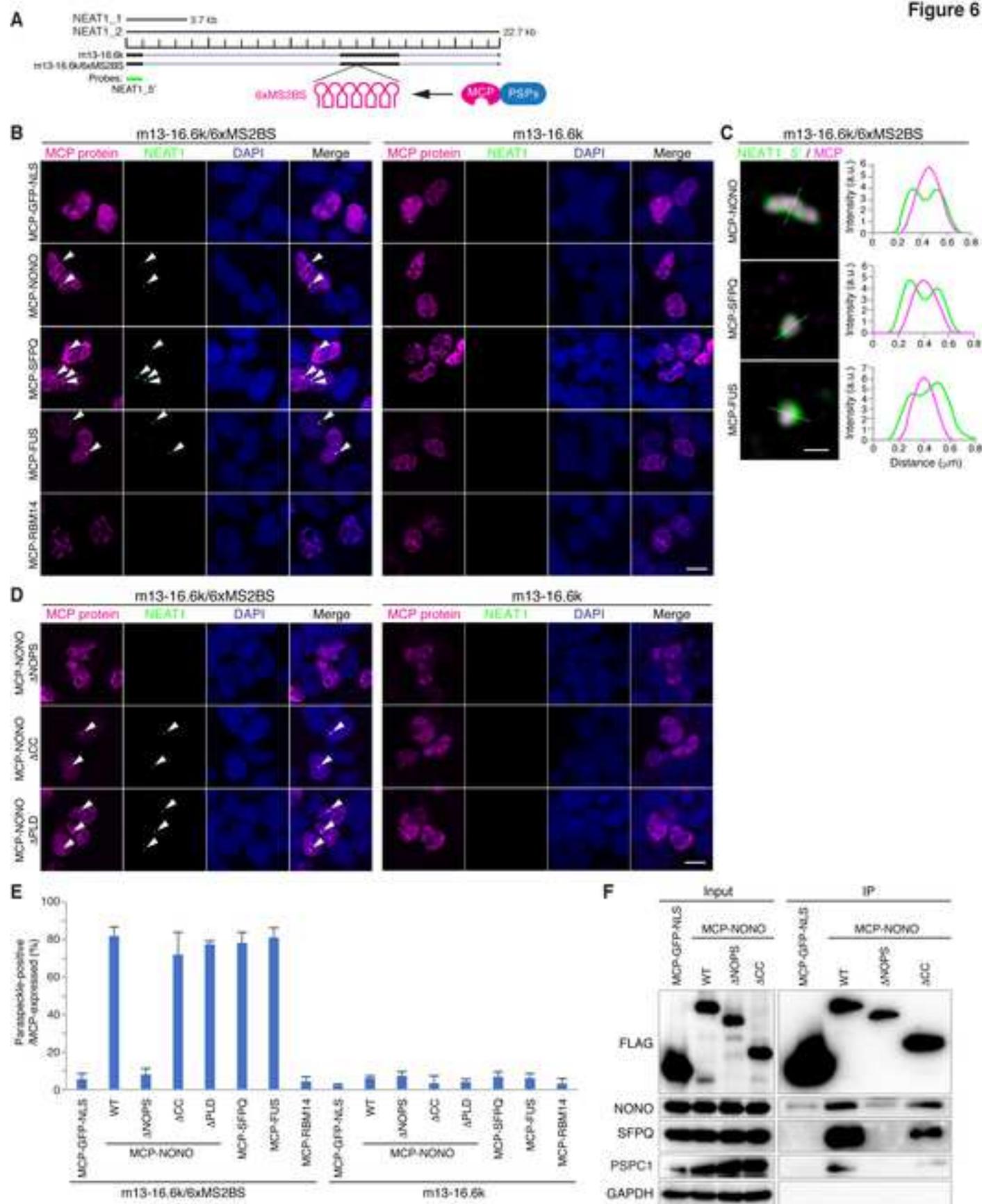


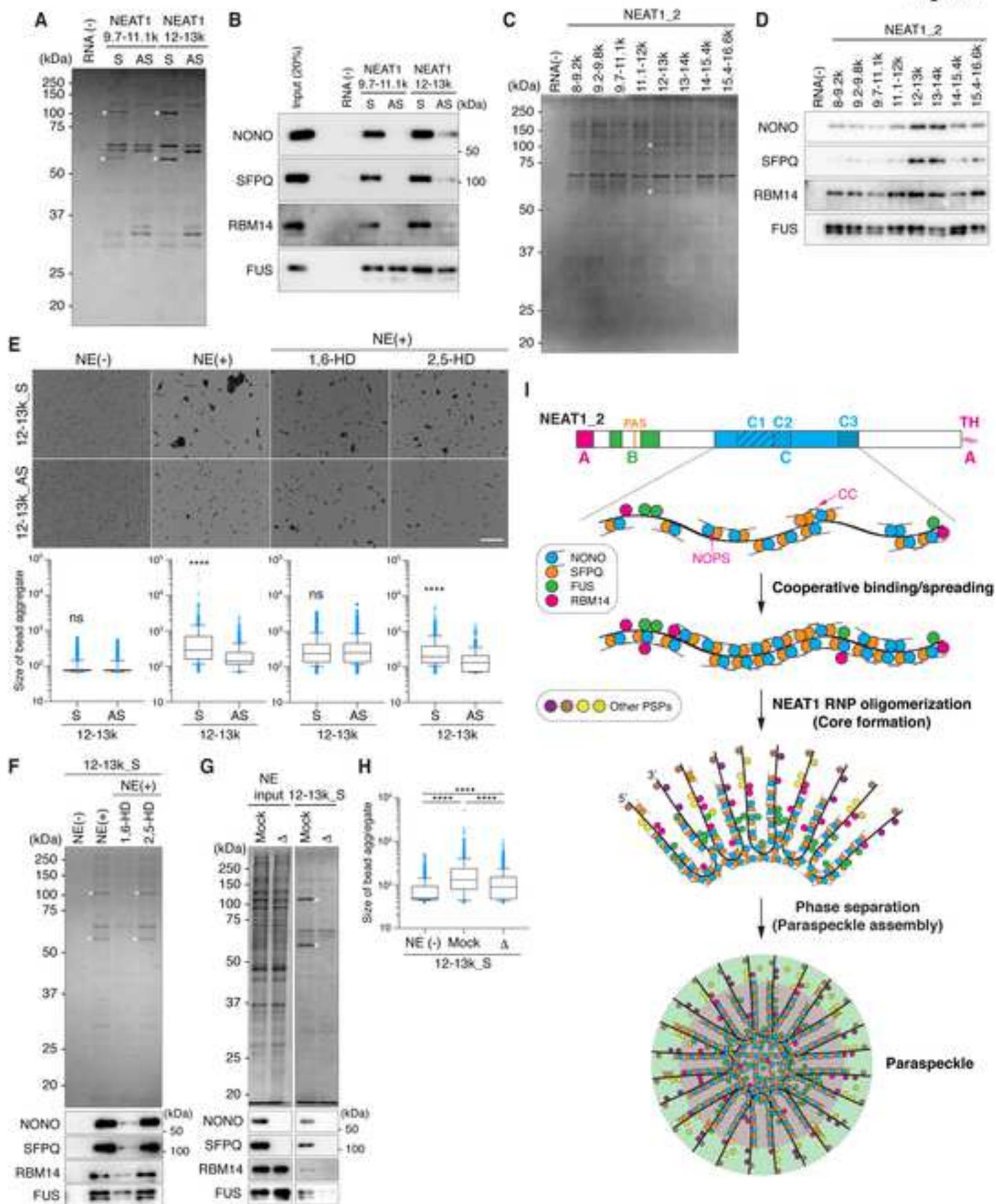












## Supplemental Figures

**Figure S1 related to Figure 1.** Characterization of  $\Delta$ TH and  $\Delta$ NEAT1 mutants. (A) Specific detection of paraspeckles with NEAT1 smFISH probes (NEAT1 5' probe and middle probe). NEAT1 RNA-FISH was carried out with the predesigned smFISH probes that target the NEAT1 5' region (NEAT1\_5' probe shown in the upper panels) and the NEAT1 middle region (NEAT1\_middle probe shown in the lower part) in WT and  $\Delta$ NEAT1 cells. Scale bar: 10  $\mu$ m. (B) Absence of paraspeckles in  $\Delta$ NEAT1 cells. Paraspeckles were visualized with NEAT1 RNA-FISH (NEAT1\_5' probe) and NONO IF. Scale bar: 10  $\mu$ m. (C) Quantitation of paraspeckle-positive cells observed in A (WT: 100%). (D) Relative expression levels of NEAT1 (both NEAT1\_1 and NEAT1\_2) (left), or only NEAT1\_2 (right) were determined by RT-qPCR. Data are represented as mean  $\pm$  SD (n = 3). (E) Requirement of the 3' TH for paraspeckle formation. The second cell clone of  $\Delta$ TH#2 was used for smFISH as in Figure 1C. Scale bar: 10  $\mu$ m. (F) Combination of conventional NEAT1 RNA-FISH and NONO IF was carried out in  $\Delta$ TH#2 cells. Scale bar: 10  $\mu$ m. (G) Quantitation of paraspeckle-positive cells in E. (H) Accumulation levels of NEAT1 isoforms in NEAT1 $\Delta$ TH #2 cells as measured by RT-qPCR. The amounts relative to the WT are plotted. Data are represented as mean  $\pm$  SD (n = 3).

**Figure S2 related to Figure 2.** Compilation of the data on extensive deletion analyses of NEAT1 domains required for NEAT1 expression, isoform switching, and paraspeckle formation. (A) Observation of paraspeckles by smFISH in the WT and the various deletion mutants shown in Figure 2A. Data were obtained from two independent clones

for each deletion mutant. Scale bar: 10  $\mu$ m. (B and C) Quantitation of paraspeckle-positive cells (B, percentage relative to the WT [WT: 100%]) and expression levels of NEAT1 isoforms (C) in the WT and deletion mutants used in A are shown as in Figure 2C–2E. Data are represented as mean  $\pm$  SD (n = 3). It should be noted that NEAT1\_2 was upregulated in  $\Delta$ 5.1–5.9k clones with almost no effect on the NEAT1\_1 level, implicating a distinct role of this region in downregulating NEAT1\_2 expression. (D) The schematics of the WT and the deletion mutants of 5' regions are shown with a scale as in Figure 2A. Although  $\Delta$ 2.1–2.8k lacks part of the G-stretch, the  $\Delta$ 1.9–2.8k mutant lacks the whole G-stretch, indicated by a magenta line. (E and F) Observation of paraspeckles in the cell lines shown in D. (G and H) Quantitation of paraspeckle-positive cells (G, WT: 100%) and expression levels of NEAT1 isoforms (H) in the WT and deletion mutants used in D are shown as in Figure 2C–2E. Data are represented as mean  $\pm$  SD (n = 3). (H) NEAT1\_2 was partially downregulated in  $\Delta$ 0.2–1k; however, it remained unchanged in  $\Delta$ 0–0.8k cells. These data support that the 0–1k region contains domain(s) required for NEAT1 expression. Two additional mutants ( $\Delta$ 4–4.7k or  $\Delta$ 4.4–5.1k) lost the phenotypes observed in  $\Delta$ 4–5.1k. In the  $\Delta$ 1.9–2.8k clones, the paraspeckles were almost completely lost, and the loss of NEAT1\_2 was more prominent than in  $\Delta$ 2.1–2.8k, suggesting the importance of the G-stretch in isoform switching.

**Figure S3 related to Figure 3.** Characterization of the middle domain of NEAT1 lncRNA required for paraspeckle assembly. (A) Dot plots of self-comparison of NEAT1\_2 sequences from human, mouse, and opossum by BLAST (<http://blast.ncbi.nlm.nih.gov/>)

with a setting of “somewhat similar”. (B) Comparison of the repeated sequences in human NEAT1 that were detected in A with the repeat masker sequences (shown by different colors, magenta: acquired using the setting “highly similar”, orange: “more dissimilar”, and blue: “somewhat similar”). (C) Observation of paraspeckles in the cell lines shown in  $\Delta$ middle,  $\Delta$ 13–20.2 mutant, and FUS KO cells that were treated with (+: 5  $\mu$ M for 6 h) or without (-) MG132. Paraspeckles were visualized as in Figure 3B. Two independent  $\Delta$ middle and  $\Delta$ 13–20.2 mutant cells were used. Scale bar: 10  $\mu$ m. (D) Quantitation of the paraspeckle size evaluated by Area/Focus and Sum intensity/Focus. Data for the cells treated with (+: 5  $\mu$ M for 6 h) or without (-) MG132 are shown separately. (\*\*\*\*:  $P < 0.0001$ , \*\*\*:  $P < 0.001$ , \*\*:  $P < 0.01$ , ns: not significant.) (E) Quantitative data of expression levels of NEAT1 isoforms in the cells used in C. Data for the cells treated with (+: 5  $\mu$ M for 6 h) or without (-) MG132 are shown separately. Data are represented as mean  $\pm$  SD ( $n = 3$ ). (F) Quantitative analyses of paraspeckle size measured by EM observation in the indicated conditions. Data for the cells treated with (+: 5  $\mu$ M for 6 h) or without (-) MG132 are shown separately. (\*\*\*\*:  $P < 0.0001$ , \*\*:  $P < 0.01$ , \*:  $P < 0.05$ , ns: not significant.) (G) Time course of paraspeckle enlargement upon MG132 treatment (5  $\mu$ M) in the three cell lines in F. (H) NEAT1 levels in DRB-treated and subsequently released cells. NEAT1<sub>2</sub> levels were quantitated by RT-qPCR in the cells used in Figure 3F. The quantitated NEAT1<sub>2</sub> levels relative to those in the untreated WT cells are plotted. Data are represented as mean  $\pm$  SD ( $n = 3$ ). (I) The core localization of SFPQ and FUS. Localization of NEAT1 (green) and SFPQ or FUS (magenta) was observed by RNA-FISH/IF and SRM. Scale bar: 500 nm.

**Figure S4 related to Figure 4.** Localization of essential PSPs in the paraspeckle of mini-NEAT1 cells and effect of 1,6-HD and 2,5-HD on cellular bodies. (A) Paraspeckle localization of four essential paraspeckle proteins (NONO, SFPQ, FUS, and RBM14) in mini-NEAT1 cells. Line profiles of these PSPs and NEAT1 are shown on the right. Scale bar: 10  $\mu\text{m}$ . (B) The schematics of the WT and  $\Delta 16.6\text{-}22.6\text{k}$  mutant are shown with a scale. The positions of the NEAT1\_5' (green) and NEAT1\_15.4-16.6k (magenta) probes used for SRM are shown below. (C) SRM observation of paraspeckles in  $\Delta 16.6\text{-}22.6\text{k}$  cells. Indicated probes as shown in B were used. The cells were treated with 5  $\mu\text{M}$  MG132 for 6 h. Scale bar: 500 nm. (D) Quantitation of NEAT1 levels in the WT and  $\Delta 16.6\text{-}22.6\text{k}$  cells by RT-qPCR (WT = 1). Data are represented as mean  $\pm$  SD (n = 3). (E) A model of the NEAT1\_2 configuration within paraspeckles in  $\Delta 16.6\text{-}22.6\text{k}$  cells. (F) Effect of 1,6-HD treatment on Cajal bodies and nuclear speckles. COIL and SRSF2 are the markers of Cajal bodies and nuclear speckles, respectively. Scale bar: 10  $\mu\text{m}$ . (G) Effect of 2,5-HD treatment on paraspeckles in HAP1 WT and mini-NEAT1 cells. Scale bar: 10  $\mu\text{m}$ .

**Figure S5 related to Figure 5.** NEAT1 middle domain contains multiple redundant subdomains for paraspeckle assembly. (A and B) Quantitation of the expression levels of the mutant cells in Figure 5A by RT-qPCR. The NEAT1 levels in the mutant cells relative to that in the WT are plotted. The data using the cells treated with or without MG132 (5  $\mu\text{M}$ ) are shown in A and B, respectively. Data are represented as mean  $\pm$  SD

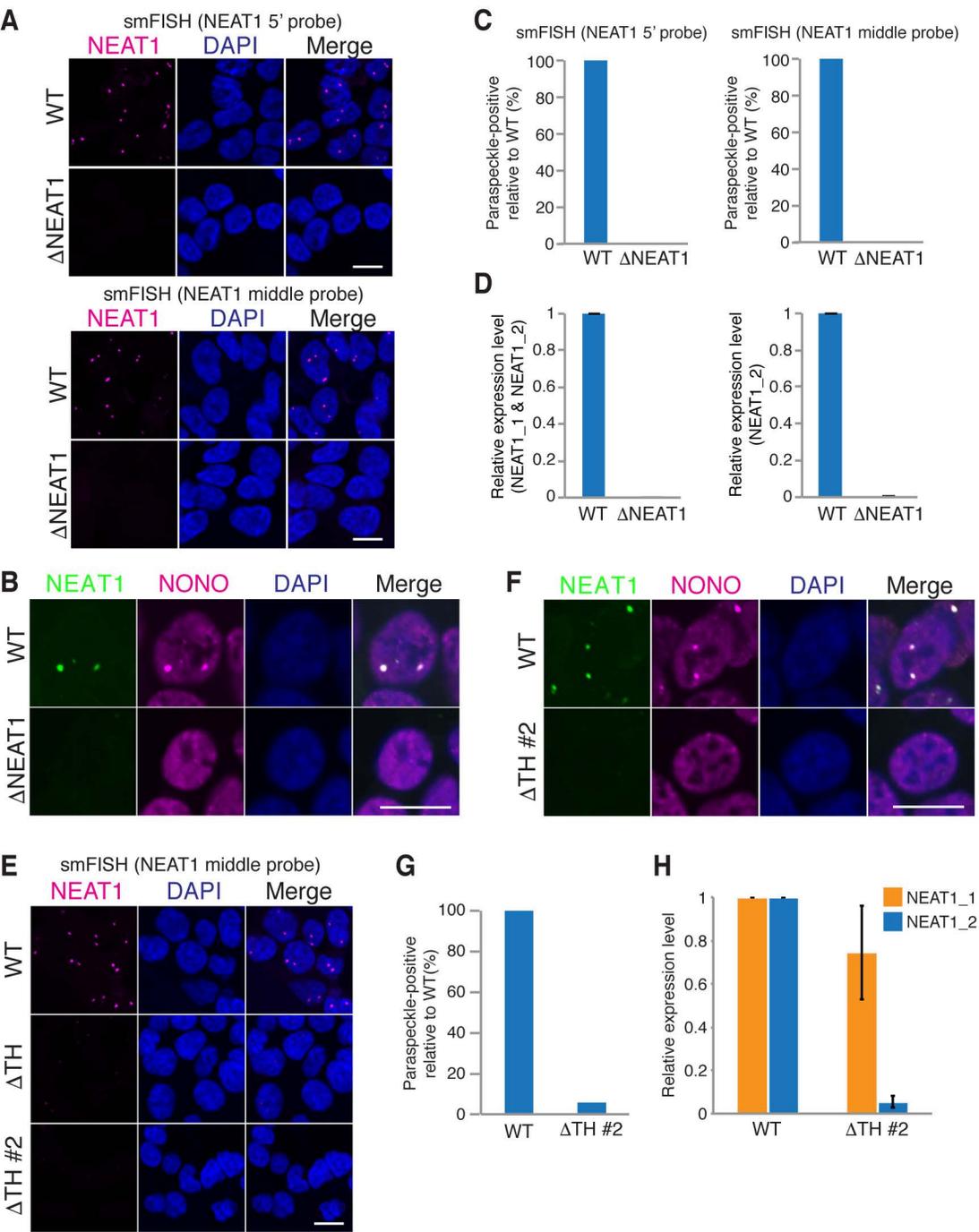
(n = 3). (C) Observation of paraspeckles in the mutant cells (without MG132 treatment) as in Figure 5B. Scale bar: 10  $\mu$ m. (D) Quantitation of the paraspeckle sizes observed in C were carried out as in Figure 5C. (\*\*\*\*:  $P < 0.0001$ , \*\*:  $P < 0.01$ , \*:  $P < 0.05$ , unindicated pairs were not significantly different.) (E) Quantitation of the expression levels of the mutant cells with MG132 treatment (5  $\mu$ M for 6 h) in Figure 5F by RT-qPCR. The NEAT1 levels in the mutant cells relative to those in the WT are plotted. Data are represented as mean  $\pm$  SD (n = 3).

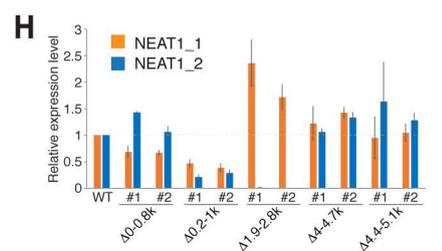
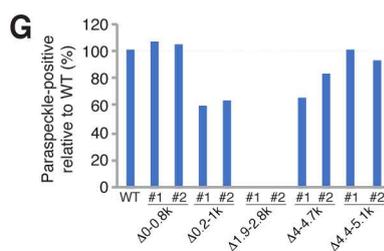
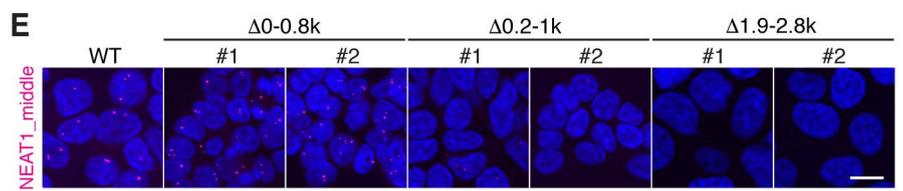
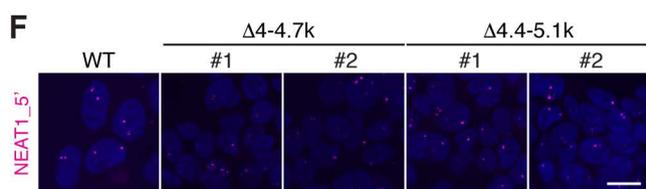
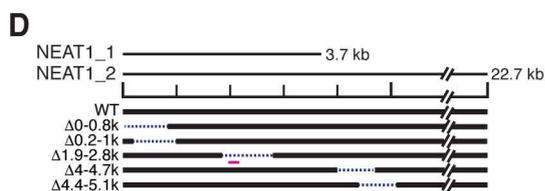
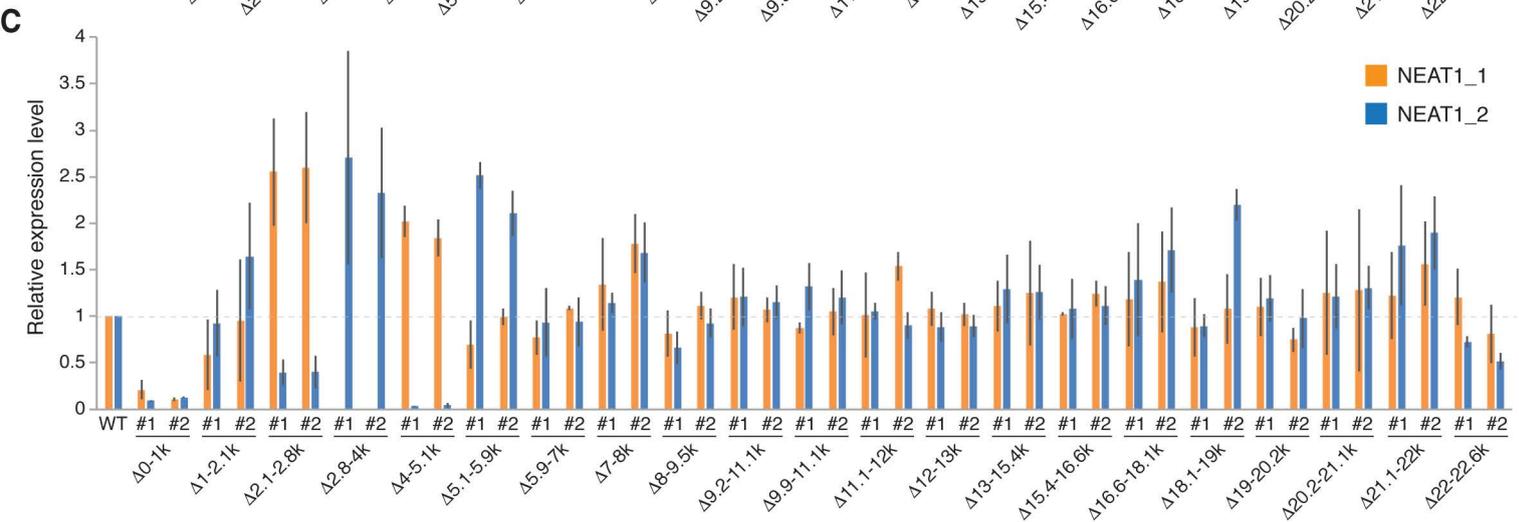
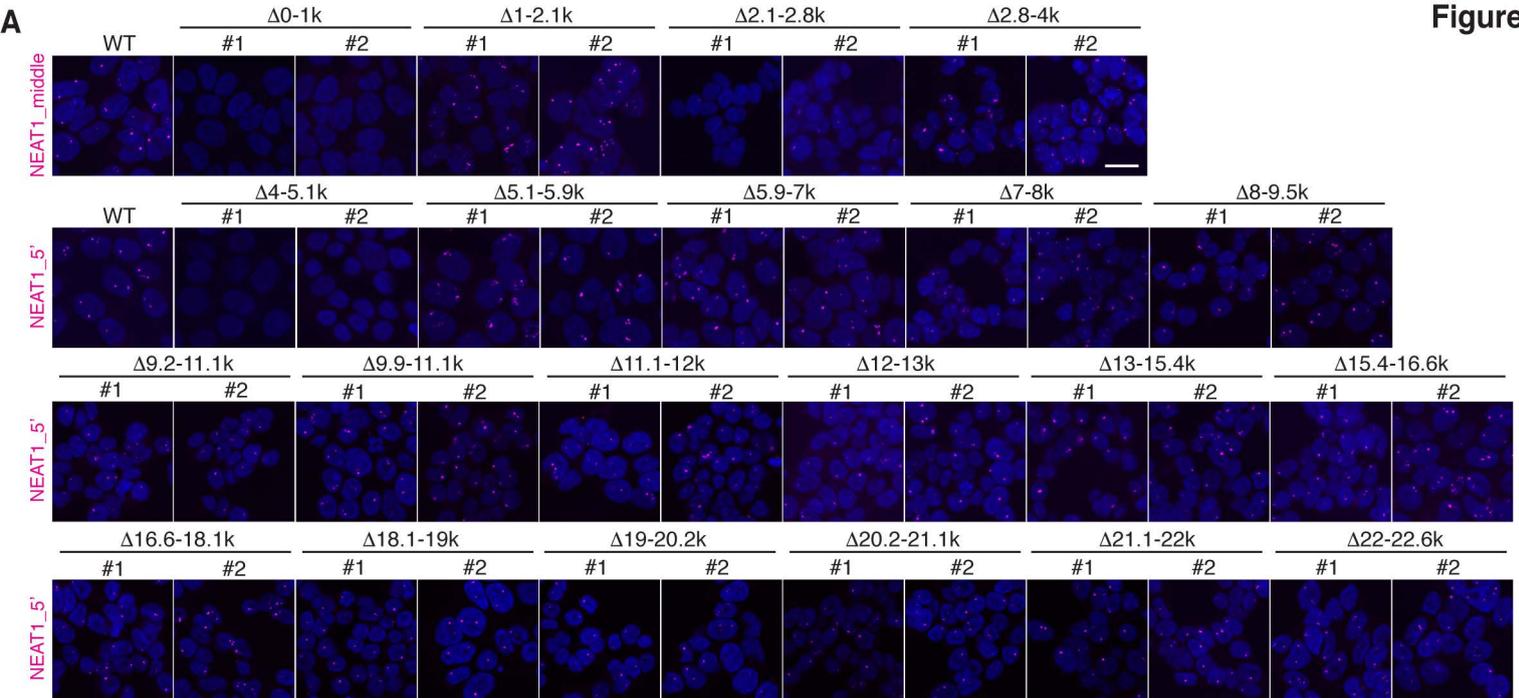
**Figure S6 related to Figure 6.** Expression of MCP-fusion proteins, effects of 1,6-HD and 2,5-HD on tethering-induced paraspeckles, and requirement of NONO in paraspeckle assembly. (A) Schematics of MCP-fused PSPs and their mutants used for tethering experiments in Figure 6 are shown with the domain structures. The numbers of amino acids are shown above each scheme. F represents Flag-tag. (B) Immunoblotting of Flag-tagged MCP-PSPs in the experiments in Figure 6B and 6D. GAPDH was used as a control. Molecular weight markers are shown on the right. (C) Quantitation of NEAT1\_2 expression levels of the cells in Figure 6B and 6D by RT-qPCR. The NEAT1 levels in the cells relative to those in the MCP-GFP-NLS expressing m13-16.6k/6 $\times$ MS2 cells are plotted. Data are represented as mean  $\pm$  SD (n = 3). (D) The paraspeckle structures induced by tethering NONO, SFPQ, and FUS in the presence of 1,6-HD or 2,5-HD. Tethering-induced paraspeckles are indicated by white arrowheads. The cells were treated with 5  $\mu$ M MG132 for 6 h. Scale bar: 10  $\mu$ m. (E) Immunoblotting to detect NONO in WT and  $\Delta$ NONO cells. GAPDH was used as a loading control. (F) Quantitation

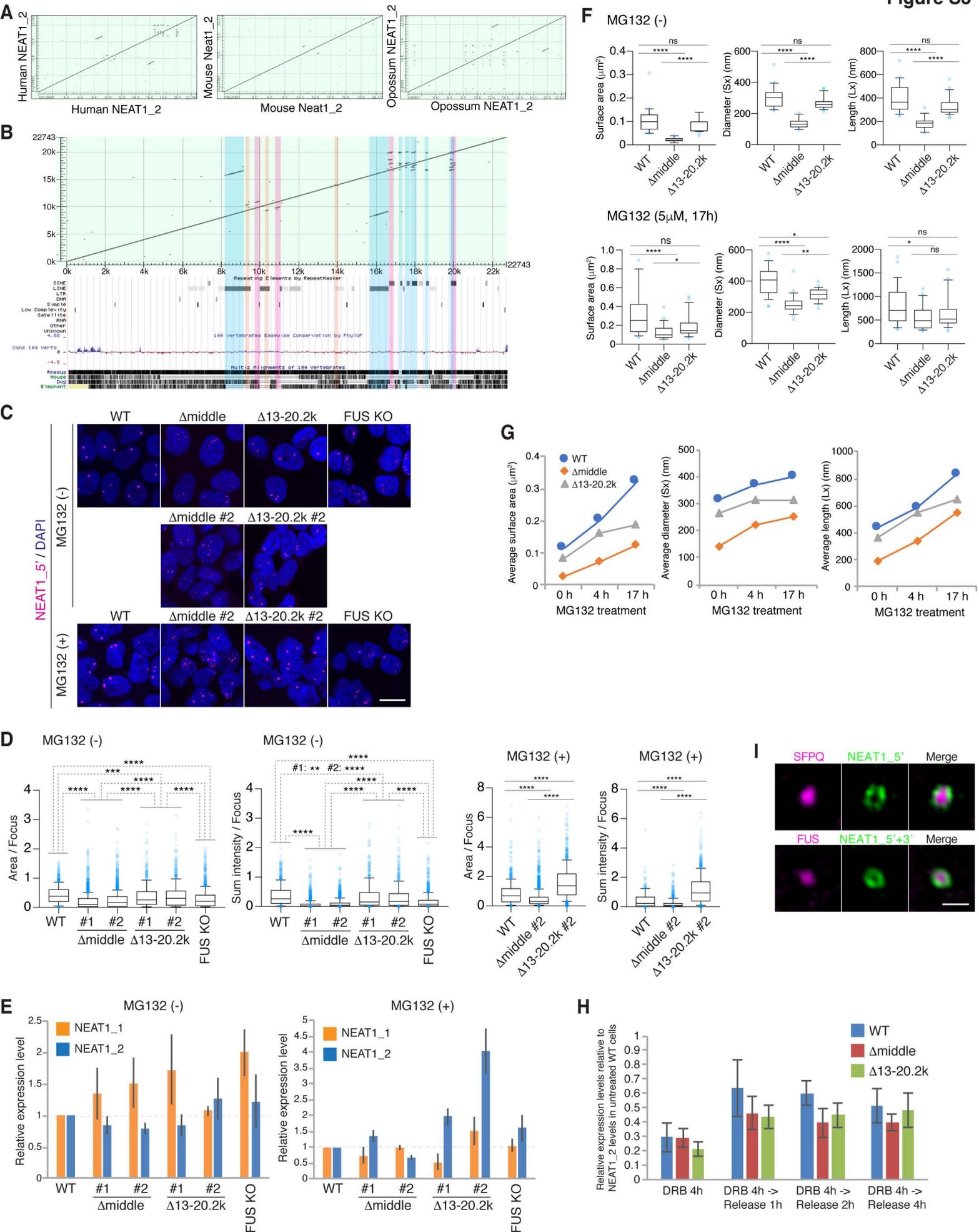
of NEAT1 levels in WT and  $\Delta$ NONO cells treated with (+: 5  $\mu$ M MG132 for 6 h) or without (-) MG132. Data are represented as mean  $\pm$  SD (n = 3). (G) NEAT1 smFISH in the cells shown in F. (H) Quantitation of the paraspeckle sizes observed in F (with 5  $\mu$ M MG132 for 6 h) was carried out. (\*\*\*\*:  $P < 0.0001$ .) Scale bar: 10  $\mu$ m. (I) SRM observation of the NEAT1 foci detected in E. Paraspeckle substructures were detected by NEAT1 FISH with the mixture of NEAT1\_5' and 3' probes (green) and NEAT1 middle probe (magenta). Paraspeckles in  $\Delta$ NONO in the absence of MG132 were undetected. Scale bar: 500 nm.

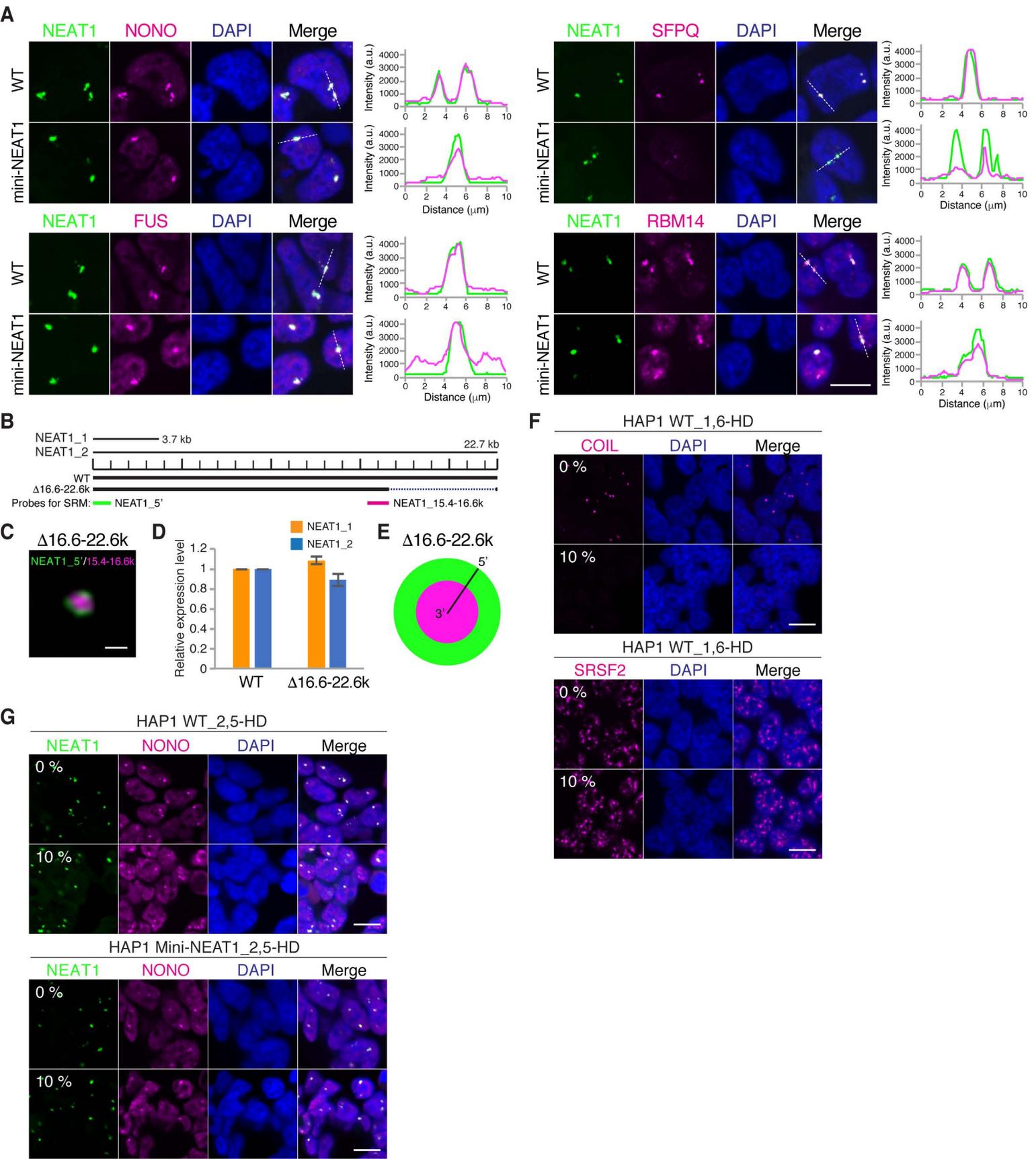
**Figure S7 related to Figure 7.** *In vivo* binding sites of several PSPs on NEAT1 and bead aggregation assay. (A) PAR-CLIP data of NONO-, SFPQ-, and FUS-binding sites mapped on NEAT1\_2. Mapped reads of NONO and SFPQ on NEAT1 (#1: U-2 OS and #2: HeLa cells), and FUS, are shown. TDP-43-binding sites on NEAT1 are shown as a comparison. The domain covered by a blue box indicates the middle domain that is necessary and sufficient for the formation of paraspeckles. The blue dotted boxes indicate the functional subdomains (C1, C2, and C3). (B) Bead aggregates visible to the eye. Insets show magnified bottom parts of 1.5 ml tubes. Arrowheads indicate the brown magnetic bead aggregates. Scale bar: 5 mm. (C) Magnetic beads observed by microscopy showing that single particles can be distinguished. Scale bar: 10  $\mu$ m. (D) Visualization of biotinylated RNAs on the beads in the bead aggregation assay as shown in Figure 7E by using FITC-labeled avidin. (E) Affinity depletion of NONO/SFPQ from NE. Depletion steps were monitored by CBB staining (upper panel) and immunoblotting (lower panels) during 2-rounds of depletion. Molecular weight markers are shown on the

left. Asterisks indicate major interacting proteins in the bead-bound sample.

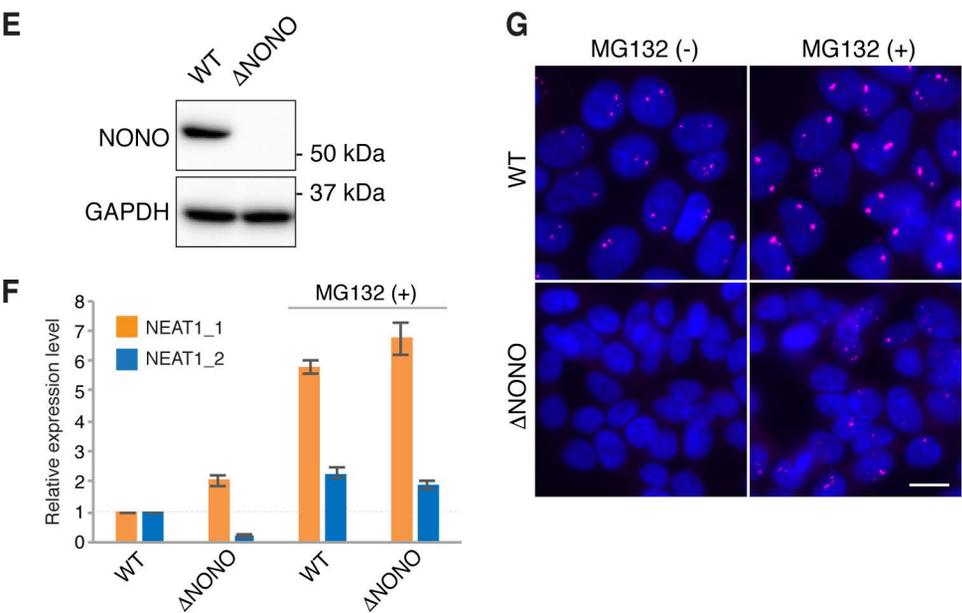
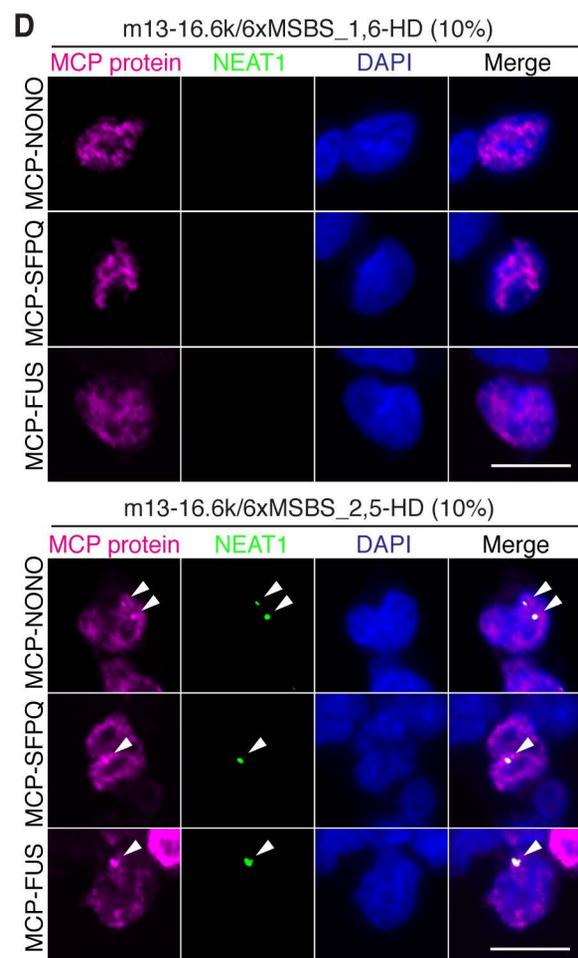
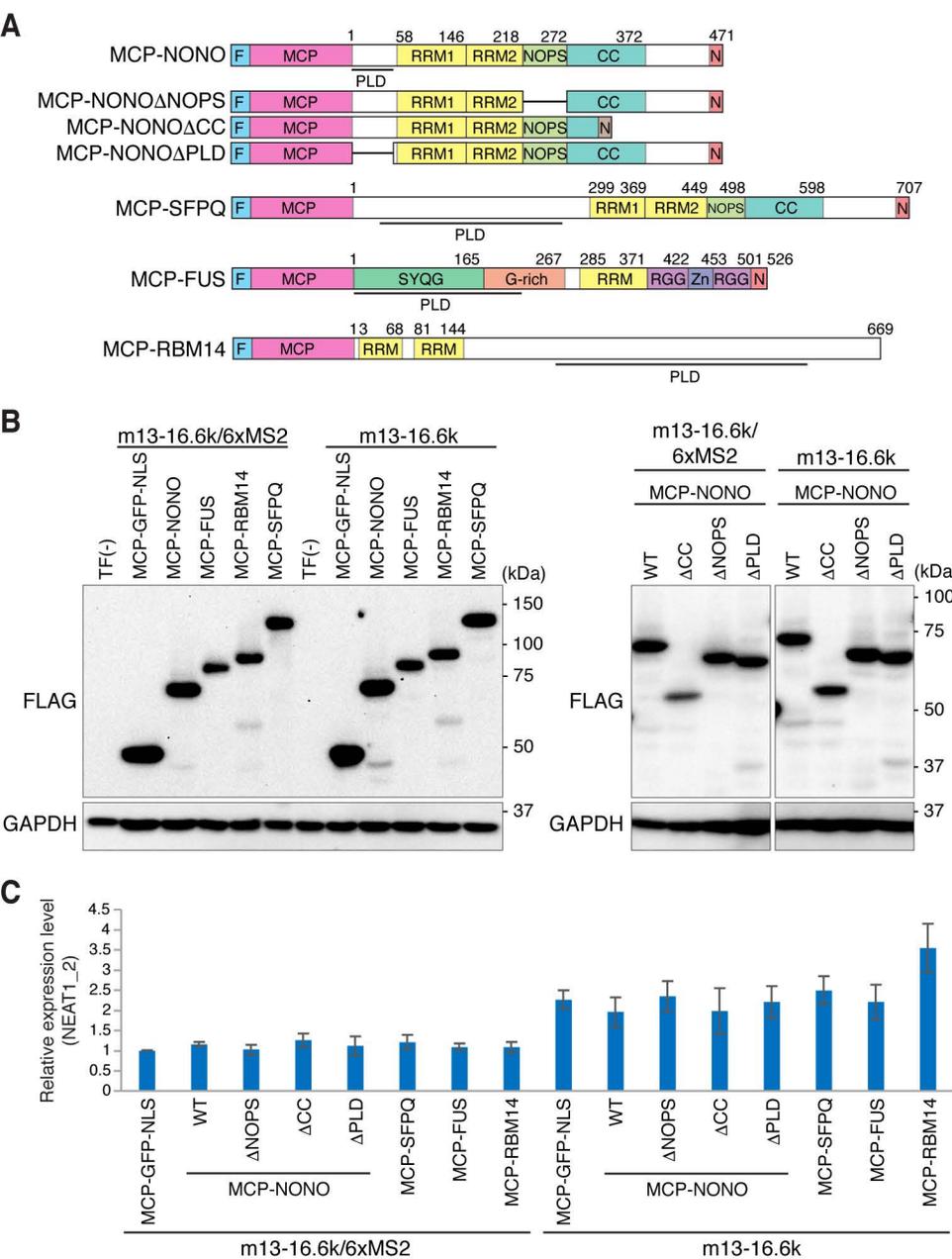












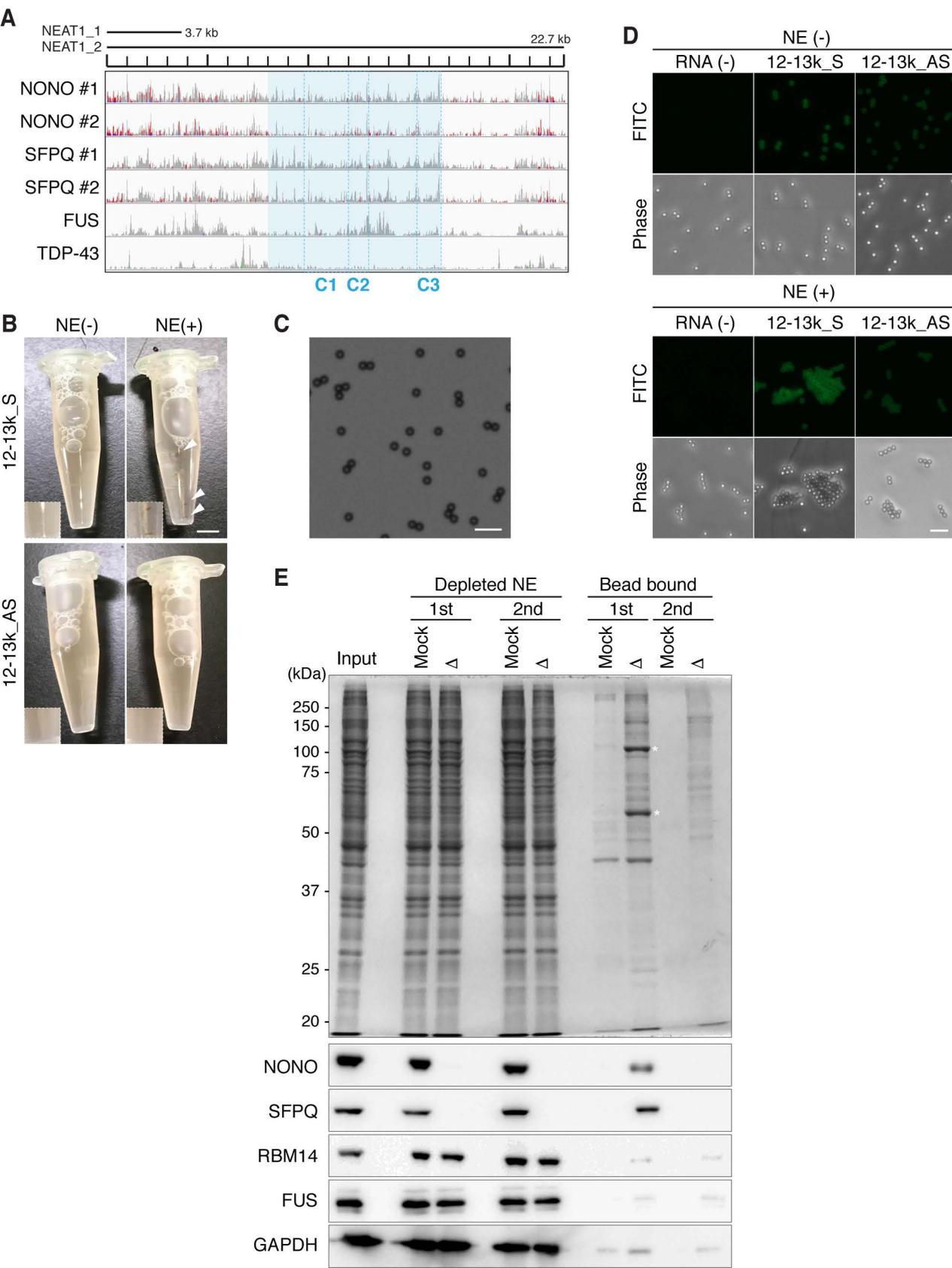


Table S1. HAP1 cell lines used in this study, Related to Key Resources Table

Name	indel regions in hNEAT1 (22743 nt)	sgRNA#1	sgRNA#2	Cell lines modified
NEAT1 ΔTH #1	22680-22700	GCAAAACCTGAGTGGCGCCA	N/A	HAP1 WT
NEAT1 ΔTH #2	22680-22694	GCAAAACCTGAGTGGCGCCA	N/A	HAP1 WT
ΔNEAT1	5-22681	TCCTCCCTGTCGCTAACTC	GCAAAACCTGAGTGGCGCCA	HAP1 WT
NEAT1 Δ0-1k #1	5-1032	TCCTCCCTGTCGCTAACTC	AGAGGCCTCCCCTGAGGC	HAP1 WT
NEAT1 Δ0-1k #2	5-1032	TCCTCCCTGTCGCTAACTC	AGAGGCCTCCCCTGAGGC	HAP1 WT
NEAT1 Δ1-2.1k #1	1033-2115	AGAGGCCTCCCCTGAGGC	CAGCTTGGCACCGTCCCCTG	HAP1 WT
NEAT1 Δ1-2.1k #2	1033-2115	AGAGGCCTCCCCTGAGGC	CAGCTTGGCACCGTCCCCTG	HAP1 WT
NEAT1 Δ2.1-2.8k #1	2116-2805	CAGCTTGGCACCGTCCCCTG	GTGGGCTTTAACGTATTTA	HAP1 WT
NEAT1 Δ2.1-2.8k #2	2115-2806	CAGCTTGGCACCGTCCCCTG	GTGGGCTTTAACGTATTTA	HAP1 WT
NEAT1 Δ2.8-4k #1	2807-4053	GTGGGCTTTAACGTATTTA	TACCACTAGAAGCTCTAAC	HAP1 WT
NEAT1 Δ2.8-4k #2	2808-4053	GTGGGCTTTAACGTATTTA	TACCACTAGAAGCTCTAAC	HAP1 WT
NEAT1 Δ4-5.1k #1	4054-5116	TACCACTAGAAGCTCTAAC	GATTACGGGGTATAAATGT	HAP1 WT
NEAT1 Δ4-5.1k #2	4041-5115	TACCACTAGAAGCTCTAAC	GATTACGGGGTATAAATGT	HAP1 WT
NEAT1 Δ5.1-5.9k #1	5114-5945	GATTACGGGGTATAAATGT	GGCGGGTCTGCTCAACTAGC	HAP1 WT
NEAT1 Δ5.1-5.9k #2	5114-5945	GATTACGGGGTATAAATGT	GGCGGGTCTGCTCAACTAGC	HAP1 WT
NEAT1 Δ5.9-7k #1	5942-7025	GGCGGGTCTGCTCAACTAGC	CATTAAACCTTCTCCCG	HAP1 WT
NEAT1 Δ5.9-7k #2	5942-7025	GGCGGGTCTGCTCAACTAGC	CATTAAACCTTCTCCCG	HAP1 WT
NEAT1 Δ7-8k #1	7024-8005	CATTAAACCTTCTCCCG	GGATATCATATACTCACAAC	HAP1 WT
NEAT1 Δ7-8k #2	7023-8005	CATTAAACCTTCTCCCG	GGATATCATATACTCACAAC	HAP1 WT
NEAT1 Δ8-9.5k #1	8004-9549	GGATATCATATACTCACAAC	CACATGTTAGTGGTCAAAG	HAP1 WT
NEAT1 Δ8-9.5k #2	8004-9549	GGATATCATATACTCACAAC	CACATGTTAGTGGTCAAAG	HAP1 WT
NEAT1 Δ9.2-11.1k #1	9201-11091	GTACCTTATAACGTGGATT	TATTACCTTGGCTAGGGGG	HAP1 WT
NEAT1 Δ9.2-11.1k #2	9201-11091	GTACCTTATAACGTGGATT	TATTACCTTGGCTAGGGGG	HAP1 WT
NEAT1 Δ9.9-11.1k #1	9898-10974	AATGGCTTTGGGGTACAAA	N/A	HAP1 WT
NEAT1 Δ9.9-11.1k #2	9900-10952	AATGGCTTTGGGGTACAAA	N/A	HAP1 WT
NEAT1 Δ11.1-12k #1	11091-(Int insertion)-12009	TATTACCTTGGGGTACAAA	TCATATATGGAGCGTCTAG	HAP1 WT
NEAT1 Δ11.1-12k #2	11091-12011	TATTACCTTGGGGTACAAA	TCATATATGGAGCGTCTAG	HAP1 WT
NEAT1 Δ12-13k #1	12011-12986	TCATATATGGAGCGTCTAG	CACCGTGTATACTCGATT	HAP1 WT
NEAT1 Δ12-13k #2	12011-12986	TCATATATGGAGCGTCTAG	CACCGTGTATACTCGATT	HAP1 WT
NEAT1 Δ13-15.4k #1	12997-15409	CACCGTGTATACTCGATT	GGCACTGGACATAAGCCCCC	HAP1 WT
NEAT1 Δ13-15.4k #2	12997-15409	CACCGTGTATACTCGATT	GGCACTGGACATAAGCCCCC	HAP1 WT
NEAT1 Δ15.4-16.6k #1	15400-16612	GGCACTGGACATAAGCCCCC	TGACTTGATTGGCTGTGCA	HAP1 WT
NEAT1 Δ15.4-16.6k #2	15400-16612	GGCACTGGACATAAGCCCCC	TGACTTGATTGGCTGTGCA	HAP1 WT
NEAT1 Δ16.6-18.1k #1	16613-18121	TGACTTGATTGGCTGTGCA	TTGGACTTAAGGGCATCATC	HAP1 WT
NEAT1 Δ16.6-18.1k #2	16614-18121	TGACTTGATTGGCTGTGCA	TTGGACTTAAGGGCATCATC	HAP1 WT
NEAT1 Δ18.1-19k #1	18120-19104	TTGGACTTAAGGGCATCATC	GAGGGTGGCTGACAGTCTGA	HAP1 WT
NEAT1 Δ18.1-19k #2	18120-19104	TTGGACTTAAGGGCATCATC	GAGGGTGGCTGACAGTCTGA	HAP1 WT
NEAT1 Δ19-20.2k #1	19106-20155	GAGGGTGGCTGACAGTCTGA	TCAGATAACCACTCTAACTA	HAP1 WT
NEAT1 Δ19-20.2k #2	19106-20155	GAGGGTGGCTGACAGTCTGA	TCAGATAACCACTCTAACTA	HAP1 WT
NEAT1 Δ20.2-21.1k #1	20156-21060	TCAGATAACCACTCTAACTA	TACCGGGGAAACGTGCCAC	HAP1 WT
NEAT1 Δ20.2-21.1k #2	20156-21060	TCAGATAACCACTCTAACTA	TACCGGGGAAACGTGCCAC	HAP1 WT
NEAT1 Δ21.1-22k #1	21033-(66nt insertion)-22025	TACCGGGGAAACGTGCCAC	TCCGACTTCATTTCGAGTGA	HAP1 WT
NEAT1 Δ21.1-22k #2	20161-22012	TACCGGGGAAACGTGCCAC	TCCGACTTCATTTCGAGTGA	HAP1 WT
NEAT1 Δ22-22.6k #1	22013-22647	TCCGACTTCATTTCGAGTGA	GGGAGGCGGGAGCCGCCGC	HAP1 WT
NEAT1 Δ22-22.6k #2	22013-22647	TCCGACTTCATTTCGAGTGA	GGGAGGCGGGAGCCGCCGC	HAP1 WT
NEAT1 ΔPAS	3699-3786	ACTTTATTTGTGCTGTAAG	TACATCCAAAGTCGTTATGA	HAP1 WT
NEAT1 Δ2.1-2.8k/ΔPAS	2116-2805	CACTTGGCACCGTCCCCTG	GTGGGCTTTAACGTATTTA	HAP1 NEAT1 ΔPAS
NEAT1 Δ4-5.1k/ΔPAS	4046-5115	TACCACTAGAAGCTCTAAC	GATTACGGGGTATAAATGT	HAP1 NEAT1 ΔPAS
NEAT1 Δ0-0.8k #1	5-816	TCCTCCCTGTCGCTAACTC	CCACCTGAAAAATAAGCGT	HAP1 WT
NEAT1 Δ0-0.8k #2	5-816	TCCTCCCTGTCGCTAACTC	CCACCTGAAAAATAAGCGT	HAP1 WT
NEAT1 Δ0.2-1k #1	211-1032	AAATATCCGAGGGTCCCC	AGAGGCCTCCCCTGAGGC	HAP1 WT
NEAT1 Δ0.2-1k #2	211-1032	AAATATCCGAGGGTCCCC	AGAGGCCTCCCCTGAGGC	HAP1 WT
NEAT1 Δ1.9-2.8k #1	1855-2805	ATTGTGTTGCTAAACGCT	GTGGGCTTTAACGTATTTA	HAP1 WT
NEAT1 Δ1.9-2.8k #2	1853-2817	ATTGTGTTGCTAAACGCT	GTGGGCTTTAACGTATTTA	HAP1 WT
NEAT1 Δ4-4.7k #1	4053-4739	TACCACTAGAAGCTCTAAC	GTACACTACTGACAAACGGG	HAP1 WT
NEAT1 Δ4-4.7k #2	4054-4740	TACCACTAGAAGCTCTAAC	GTACACTACTGACAAACGGG	HAP1 WT
NEAT1 Δ4.4-5.1k #1	4413-5115	GCAAAACCTACACCCCAAG	GATTACGGGGTATAAATGT	HAP1 WT
NEAT1 Δ4.4-5.1k #2	4413-5115	GCAAAACCTACACCCCAAG	GATTACGGGGTATAAATGT	HAP1 WT
NEAT1 Δ8-16.6k #1 (Amiddle #1)	8034-16610	CCTGATGTTAGTGGCTATGT	CTTGATTGGCTGTGCAAGG	HAP1 WT
NEAT1 Δ8-16.6k #2 (Amiddle #2)	8006-16611	GGATATCATATACTCACAAC	TGACTTGATTGGCTGTGCA	HAP1 WT
NEAT1 Δ13-20.2k #1	12988-20156	CACCGTGTATACTCGATT	TCAGATAACCACTCTAACTA	HAP1 WT
NEAT1 Δ13-20.2k #2	12988-20156	CACCGTGTATACTCGATT	TCAGATAACCACTCTAACTA	HAP1 WT
NEAT1 Δ8-15.4k	8006-15399	GGATATCATATACTCACAAC	GGCACTGGACATAAGCCCCC	HAP1 WT
NEAT1 Δ8-13k	7993-13004	GGATATCATATACTCACAAC	CACCGTGTATACTCGATT	HAP1 WT
NEAT1 Δ8-11.1k	8007-11089	GGATATCATATACTCACAAC	TATTACCTTGGCTAGGGGG	HAP1 WT
NEAT1 Δ11.1-16.6k	11019-16629	TATTACCTTGGCTAGGGGG	TGACTTGATTGGCTGTGCA	HAP1 WT
NEAT1 Δ16.6-22.6k	16613-22647	CTTGATTGGCTGTGCAAGG	GGGAGGCGGGAGCCGCCGC	HAP1 WT
NEAT1 m8-16.6k (mini-NEAT1)	1034-8006	AGAGGCCTCCCCTGAGGC	GGATATCATATACTCACAAC	HAP1 WT
	16613-22647	TGACTTGATTGGCTGTGCA	GGGAGGCGGGAGCCGCCGC	HAP1 Δ1-8k
NEAT1 m9.2-16.6k	16613-22647	TGACTTGATTGGCTGTGCA	GGGAGGCGGGAGCCGCCGC	HAP1 WT
	1033-9198	AGAGGCCTCCCCTGAGGC	GTACCTTATAACGTGGATT	HAP1 Δ16.6-22.6k
NEAT1 m9.8-16.6k	16613-22647	TGACTTGATTGGCTGTGCA	GGGAGGCGGGAGCCGCCGC	HAP1 WT
	1033-9830	AGAGGCCTCCCCTGAGGC	GAAGTTAATATTGACTGACC	HAP1 Δ16.6-22.6k
NEAT1 m11.1-16.6k	16613-22647	TGACTTGATTGGCTGTGCA	GGGAGGCGGGAGCCGCCGC	HAP1 WT
	1033-11089	AGAGGCCTCCCCTGAGGC	TATTACCTTGGCTAGGGGG	HAP1 Δ16.6-22.6k
NEAT1 m12-16.6k	16613-22647	TGACTTGATTGGCTGTGCA	GGGAGGCGGGAGCCGCCGC	HAP1 WT
	1033-12010	AGAGGCCTCCCCTGAGGC	TCATATATGGAGCGTCTAG	HAP1 Δ16.6-22.6k
NEAT1 m13-16.6k	16613-22647	TGACTTGATTGGCTGTGCA	GGGAGGCGGGAGCCGCCGC	HAP1 WT
	1033-12986	AGAGGCCTCCCCTGAGGC	CACCGTGTATACTCGATT	HAP1 Δ16.6-22.6k
NEAT1 m8-15.4k	1034-8006	AGAGGCCTCCCCTGAGGC	GGATATCATATACTCACAAC	HAP1 WT
	15401-22648	GGCACTGGACATAAGCCCCC	GGGAGGCGGGAGCCGCCGC	HAP1 Δ1-8k
NEAT1 Δ12-13k/m9.8-16.6k	12006-12996	TCATATATGGAGCGTCTAG	CACCGTGTATACTCGATT	HAP1 m9.8-16.6k
NEAT1 m13-16.6k/6xMS2BS	14020-(6xMS2BS insertion)-14021	TACCGATATCTGTGTACAT	N/A	HAP1 NEAT1 m13-16.6k

Name	Deleted regions	sgRNA#1	sgRNA#2	Cell lines modified
NONO KO	2 nt deletion, frameshift	CCCTCACTACGTGGTTAGC	N/A	HAP1 WT
FUS KO		Purchased from Horizon Discovery (Cat#HZGH001314c006)		HAP1 WT

