ARCHITECTURAL RNAs ARE ESSENTIAL SCAFFOLDS OF PHASE-SEPARATED MEMBRANELESS BODIES

Recently, membraneless organelles and biomolecular condensates, which form through phase separation, have attracted increasing interest (Banani et al. 2017; Alberti et al. 2019). These subcellular compartments play key roles in a number of processes, including the sequestration of biomolecules and the formation of organization hubs to coordinate biochemical reactions and macromolecular assembly (Shin and Brangwynne 2017). A subset of phase-separated cellular bodies are constructed by specific long noncoding RNAs (lncRNAs) (Yamazaki and Hirose 2015; Chujo et al. 2016; Chujo and Hirose 2017), which play diverse regulatory roles in various biological processes, including gene expression (Guttman and Rinn 2012; Geisler and Coller 2013; Quinn and Chang 2016; Schmitt and Chang 2017). Multiple lncRNAs can act as essential architectural scaffolds of nuclear bodies in a variety of eukaryotes, including humans, mice, Drosophila, and yeast; hence, we termed them architectural RNAs (arcRNAs) (Yamazaki and Hirose 2015; Chujo et al. 2016; Chujo and Hirose 2017).

A number of arcRNAs are involved in the formation of various nuclear bodies. For example, NEAT1_2, the long isoform of the nuclear paraspeckle assembly transcript 1 (NEAT1) lncRNA, is an essential scaffold of paraspeckle nuclear bodies (Chen and Carmichael 2009; Clemson et al. 2009; Sasaki et al. 2009; Sunwoo et al. 2009); highly repetitive satellite III (HSATIII) lncRNAs, which are induced by conditions such as thermal stress, promote the formation of nuclear stress bodies (nSBs) in primates (Jolly et al. 1999; Aly et al. 2019; Ninomiya et al. 2019); HSATII lncRNAs are core components of cancer-associated satellite transcript bodies (Hall et al. 2017); intergenic spacer lncRNAs induce the formation of static but reversible amyloid-like solid structures termed amyloid bodies (Audas et al. 2016); histone pre-mRNAs are essential for histone locus body formation (Shevtsov and Dundr 2011); tumor-associated NBL2 transcript (TNBL) aggregates are induced by TNBL RNAs (Dumbović et al. 2018); the perinucleolar compartment is induced by the PNCRT RNA (Yap et al. 2018); Hsro RNAs are essential cores of ω-speckles in Drosophila (Mallik and Lakhotia 2009); and the Mei2 dot is a nuclear body induced by meiRNA in fission yeast (Yamashita et al. 1998). The Sam68 nuclear body and DBC1 bodies are also RNA-dependent bodies, but their arcRNAs have not yet been identified (Mannen et al. 2016). In addition, disease-associated repeat RNAs can also be regarded as a type of arcRNA (Zhang and Ashizawa 2017).

The existence of arcRNAs in a large number of organisms suggests that they are suitable for nucleating subnuclear bodies through phase separation. RNA-binding proteins (RBPs) are enriched with prion-like domains (PLDs), low-complexity domains (LCDs), and intrinsically disordered regions (IDRs), which form weak multivalent molecular interactions that promote biological phase separation (Hennig et al. 2015; Yamazaki and Hirose 2015). Thus, arcRNAs can induce phase separation by increasing the local concentration of these RBPs. More than 1000 human RBPs have been reported to date; these proteins possess a wide variety of biological functions, such as transcriptional regulation and epigenetic chromatin regulation (Baltz et al. 2012; Castello et al. 2012, 2016). RNAs can assemble a specific set of RBPs to combine and integrate their functions, even when the RBPs...
would not directly interact (Chujo et al. 2016; Engreitz et al. 2016). This property allows arcRNA-induced bodies to perform multiple tasks in the cell, including acting as molecular sponges to sequestrate proteins and/or RNAs, forming reaction crucible to enhance biochemical reactions, and organizational hubs of chromatin (Chujo et al. 2016; Chujo and Hirose 2017; Hirose et al. 2019). In addition, noncoding RNAs can act as functional molecules themselves, with roles that depend on their specific sequences and do not require protein coding capacity. Architectural noncoding RNAs are able to alter their specific sequences to adapt to circumstantial changes. Furthermore, multiple copies of IncRNAs are synthesized from their gene loci; this process is important for phase separation as it promotes multivalent interactions between high local concentrations of associated IDR proteins. Therefore, IncRNAs act as markers of specific genomic loci. This property of IncRNAs is beneficial to the amplification of signals from specific genomic loci, such as super-enhancers, and phase separation would enhance this effect (Engreitz et al. 2016). Transcription of arcRNAs is essential for the formation of phase-separated nuclear bodies, and inhibition of their transcription leads to rapid disintegration of the bodies. As most arcRNAs are stress-inducible or specifically expressed in certain developmental and pathological conditions, phase-separated nuclear bodies are transiently formed upon the transcription of arcRNAs. Therefore, it is conceivable that arcRNAs spatiotemporally control cellular signals and act as integral hubs for gene regulation.

In the sections below, we focus on NEAT1_2 as a model arcRNA. We describe how NEAT1_2 arcRNAs construct paraspeckle nuclear bodies and determine their biophysical properties, with a special focus on the functional RNA domains of NEAT1 and their partner proteins.

THE CHARACTERISTICS OF PHASE-SEPARATED PARASPECKLE NUCLEAR BODIES CONSTRUCTED BY THE NEAT1_2 arcRNA

Paraspeckles were originally identified as distinct nuclear bodies called interchromatin granule-associated zones (IGAZs) (Visa et al. 1993), which were located adjacent to interchromatin granule clusters (also known as nuclear speckles) (Fig. 1A). In 2002, paraspeckles were identified...
as nuclear bodies enriched in PSPC1, a member of the Drosophila behavior human splicing (DBHS) family of proteins that includes SFPQ and NONO (Fig. 1B, left; Fox et al. 2002). Subsequently, paraspeckles were shown to be identical to IGAZs and were found to be RNase-sensitive structures, suggesting a requirement of RNAs for their maintenance (Fox et al. 2005). In 2009, four groups independently reported that the NEAT1 IncRNA, also called Mene/β, is an essential architectural scaffold of paraspeckles, as shown by a lack of discrete paraspeckles in NEAT1 knockdown cells (Fig. 1B, right; Chen and Carmichael 2009; Clemson et al. 2009; Sasaki et al. 2009; Sunwoo et al. 2009). NEAT1 has two isoforms: the longer NEAT1_2/Menβ isoform (22.7 kb in human) is essential for paraspeckle formation, whereas the short NEAT1_1/Menε isoform (3.7 kb in human) is dispensable (Fig. 1C; Naganuma et al. 2012; Yamazaki and Hirose 2015). NEAT1 has two isoforms: the longer NEAT1_2/Menβ isoform (22.7 kb in human) is essential for paraspeckle formation, whereas the short NEAT1_1/Menε isoform (3.7 kb in human) is dispensable (Fig. 1C; Naganuma et al. 2012; Yamazaki and Hirose 2015). Most PSPs are RBPs that contain PLDs, LCDs, or IDRs that show significant overlaps with a number of genes associated with amyotrophic lateral sclerosis (Yamazaki and Hirose 2015). Electron microscopy studies revealed that paraspeckles are roundish or oblong structures with constant short axes (~360 nm in human HeLa cells) (Fig. 1A) and become elongated when NEAT1_2 is up-regulated (Souquere et al. 2010; Hirose et al. 2014b). Notably, the NEAT1_2 IncRNA is spatially organized within paraspeckles; the 5′ and 3′ regions of NEAT1_2 are located in the shells of paraspeckles, whereas the middle region is located in the core region, suggesting that NEAT1_2 forms a looped structure within paraspeckles (Fig. 1A, D; Souquere et al. 2010; West et al. 2016). In addition, PSPs are also spatially organized within paraspeckles, with distinct localizations at the core, shell, and patch regions (Fig. 1E; Kawaguchi et al. 2015; West et al. 2016).

Multiple lines of evidence support the concept that paraspeckles are phase-separated nuclear bodies. First, fluorescence recovery after photobleaching (FRAP) experiments revealed that PSPs are dynamically exchanged between paraspeckles and the surrounding nucleoplasm (Mao et al. 2011; Audas et al. 2016; Wang et al. 2018). Second, paraspeckles are sensitive to 1,6-hexanediol, which mainly disrupts hydrophobic interactions and contributes to the formation of phase-separated subcellular bodies (Yamazaki et al. 2018b). Third, the abilities of the PSPs FUS and RBM14, which are essential for paraspeckle formation, to form hydrogels in vitro are essential for paraspeckle assembly (Hennig et al. 2015; Fox et al. 2018). Fourth, paraspeckles undergo a fusion and fission process termed “kiss-and-run fusion” (Mao et al. 2011; Yang et al. 2019). A recent review article reported that paraspeckles are induced by polymer–polymer phase separation rather than the typical liquid–liquid phase separation (Peng and Weber 2019). Taken together, these findings suggest that paraspeckles are massive phase-separated subnuclear structures with a distinct shape and ordered interior configuration.

### BIORGENESIS OF THE NEAT1_2 arcRNA AND PARASPECKLES

Synthesis of the NEAT1 isoforms is regulated by alternative 3′-end processing governed by HNRNPK and TDP-43 (Naganuma et al. 2012; Modic et al. 2019). The 3′ end of NEAT1_2 contains an unusual triple helix structure (3′TH) that is essential for stabilization of the arcRNA (Fig. 1C; Brown et al. 2012; Wilusz et al. 2012; Yamazaki et al. 2018b). Seven PSPs and the SWI/SNF complex are essential for paraspeckle formation (Fig. 2; Naganuma et al. 2012; Kawaguchi et al. 2015; West et al. 2016). Specifically, SFPQ, NONO, and RBM14 are essential for NEAT1_2 stability (Naganuma et al. 2012), and HNRNPK controls isoform switching to the NEAT1_2 isoform by inhibiting polyadenylation of the NEAT1_1 arcRNA.
isoform (Fig. 2; Naganuma et al. 2012). NONO, FUS, DAZAP1, HNRNPH3, and the SWI/SNF complex are essential for paraspeckle assembly (Naganuma et al. 2012; Kawaguchi et al. 2015; Yamazaki et al. 2018b), and RNA polymerase II transcription is also essential for paraspeckle formation (Fig. 2). Upon transcriptional inhibition, paraspeckles are rapidly disassembled alongside NEAT1 degradation and relocation of PSPs to nucleolar cap structures (Fox et al. 2005; Chujo et al. 2017). Paraspeckles are formed in close proximity to the NEAT1 gene loci, likely because NEAT1 lncRNAs are highly concentrated at these regions during transcription, and thus the environment favors phase separation by facilitating molecular interactions among RNAs and proteins. This evidence suggests cotranscriptional formation of paraspeckles via phase separation (Mao et al. 2011; Yamazaki et al. 2018b). During this process, approximately 50 NEAT1 lncRNAs are incorporated into a single spherical paraspeckle (Fig. 2; Chujo et al. 2017). The paraspeckle is a highly dynamic structure; FRAP studies revealed that 60%–70% of PSPs within paraspeckles are recovered within several minutes (Mao et al. 2011; Audas et al. 2016; Wang et al. 2018). In addition, a recent study showed that elongated, oblong paraspeckles show slower mobility than spherical paraspeckles (Wang et al. 2018). These data indicate the existence of slow or immobile fractions of paraspeckles and that paraspeckle dynamics are related to their shape.

FUNCTIONAL RNA DOMAINS OF THE NEAT1_2 lncRNA

Although the biological importance of various lncRNAs has been shown, the specific sequences and structural elements required for their biogenesis and functions remain poorly understood (Hirose et al. 2014a). Nonetheless, some lncRNAs have been studied intensively. A previous study examining the XIST lncRNA revealed the existence of functional modular repeat domains involved in X-chromosome inactivation (Wutz et al. 2002). Subsequent studies identified the subdomains of the XIST lncRNA and their partner proteins required for specific functions (Chu et al. 2015; Monfort and Wutz 2017; Colognori et al. 2019). We recently identified multiple modular domains of NEAT1 involved in its biogenesis and the formation of paraspeckles (Yamazaki et al. 2018b). Similar to that of many other lncRNAs, the cross-species homology of mammalian NEAT1_2 sequences is markedly lower than those of protein-coding mRNAs. In general, mammalian NEAT1_2 sequences only display homology at the 5′ region (~1.6 kb), short stretches in the internal regions, and the 3′TH (Fig. 3A, B). Nevertheless, mouse Neat1 can form paraspeckles in human cells, suggesting that the functions of NEAT1_2, at least in paraspeckle formation, are conserved between mouse and human (Fig. 3C; Naganuma et al. 2012).

Figure 3. Conservation of the sequences and functions of NEAT1_2 between human and mouse. (A) A dot blot showing the homologous regions of the human and mouse NEAT1_2 sequences. The BLAST search parameter was “somewhat similar.” (B) A screenshot of the UCSC genome browser record for the human NEAT1 gene locus with modifications. Sequence conservation (PhastCons) and positions of repetitive elements identified by a RepeatMasker algorithm are shown. (C) Transient expression of mouse Neat1_2 (mNeat1_2) induces paraspeckles in human HeLa cells (magenta signals). Scale bar, 5 µm.
To dissect the roles of specific sequences within the human NEAT1_2 lncRNA, we performed extensive deletion analyses using CRISPR–Cas9 and the human haploid HAP1 cell line (Yamazaki et al. 2018b). These analyses revealed that multiple functional RNA domains of NEAT1_2 dictate its own biogenesis and the coupled formation of paraspeckles. The 5' end (first 1 kb) and 3'TH of NEAT1_2 (termed the A domains) are necessary for stability of the lncRNA (Yamazaki et al. 2018b). The regions upstream (2.1–2.8 kb) and downstream (4–5 kb) of the polyadenylation site of NEAT1_1 (termed the B domains) promote the expression of NEAT1_2 by repressing NEAT1_1 polyadenylation, indicating a role in isoform switching from NEAT1_1 to NEAT1_2 (Yamazaki et al. 2018b). The middle domain of NEAT1_2 (termed the C domain) is necessary and sufficient for paraspeckle assembly. Notably, our CRISPR dissection analysis also revealed that the C domain contains several subdomains (C1: 9.8–12 kb; C2: 12–13 kb; C3: 15.4–16.6 kb) that have redundant functions in paraspeckle assembly (Yamazaki et al. 2018b). Such redundancy is also found among the C domain and the unidentified outside domain, suggesting that the presence of multiple redundant domains confers a robust competency of NEAT1_2 to build paraspeckles (Fig. 4).

In our recent study, artificial tethering of NONO, SFPQ, or FUS, but not RBM14, rescued paraspeckle assembly in cells expressing a NEAT1 mutant lacking the C1 and C2 subdomains (Yamazaki et al. 2018b). Moreover, the rescue activity of NONO required dimerization with DBHS family proteins, including SFPQ, PSPC1, and NONO itself, which is also required for the oligomerization of these proteins along the RNA molecule. CLIP-seq data showed that the C domain of NEAT1_2 contains multiple binding sites for NONO/SFPQ, which are likely to be essential for paraspeckle formation (Naganuma et al. 2012; Passon et al. 2012; Lee et al. 2015; Huang et al. 2018; Yamazaki et al. 2018b). The in vitro–transcribed RNA derived from the C2 domain preferentially binds to NONO/SFPQ and induces the formation of higher-order structures that are sensitive to 1,6-hexanediol and depend on NONO/SFPQ (Yamazaki et al. 2018b). In addition, the NONO/SFPQ oligomer seems to recruit additional PSPs such as FUS and RBM14, which in turn likely induce phase separation (Hennig et al. 2015; Yamazaki et al. 2018b). We recently found that the NEAT1_2 arcRNA is poorly extracted by conventional RNA purification methods using acid guanidinium thiocyanate-phenol-chloroform reagents such as TRIzol (Thermo Fisher Scientific). This unusual feature, which we termed “semixtractability,” depends on NONO, FUS, or the NEAT1_2 C domain, suggesting that the tenuous interactions between these proteins and the C domain, which can probably tolerate guanidinium thiocyanate and phenol-containing harsh conditions, contribute to arcRNA-induced cellular body formation (Chuo et al. 2017; T Yamazaki, unpubl. observ.). SFPQ forms higher-order structures by cooperative spreading and coating the nucleic acids, and a similar spreading mechanism is also proposed to play a role in splicing regulation by HNRNPA1 (Zhu et al. 2001; Lee et al. 2015). Hence, this spreading mechanism might play an important role in the initial process of paraspeckle formation.

In addition to the A, B, and C domains, NEAT1_2 contains UG repeat stretches that are evolutionally conserved among multiple species. These repeat sequences are thought to be the major binding sites of TDP-43 (Tollervey et al. 2011). When we deleted the UG repeats from NEAT1_2 using CRISPR–Cas9, recruitment of TDP-43 to the paraspeckles was reduced dramatically (Modic et al. 2019). Furthermore, knock-in of a long UG repeat stretches into the UG-deleted NEAT1_2 mutant recovered the recruitment of TDP-43, suggesting that the UG repeat stretches are necessary and sufficient for TDP-43 recruitment to paraspeckles (Fig. 4; Modic et al. 2019).

As described above, NEAT1_2 is looped and spatially organized within paraspeckles (Fig. 4). Our super-resolution microscopy analyses revealed that deletion of the 3' region of NEAT1_2 (Δ16.6–22.6-kb mutant) alters its configuration (Yamazaki et al. 2018b). In wild-type cells, the 3' end of NEAT1_2 is located in the shell of the paraspeckle, but the 3' end of the Δ16.6–22.6-kb mutant localizes to the paraspeckle core. This finding suggests that the configuration of NEAT1_2 is determined by one or more of its domains. By extension, we speculate that other NEAT1_2 domains will also contribute to the organization of the lncRNA within paraspeckles.

NEAT1_1, which can be regarded as one of the functional domains of the NEAT1 gene, forms “micro-speckles,” suggesting a role outside of paraspeckles (Li et al. 2017a). A recent study showed that NEAT1_1-specific knockout mice do not show an aberrant phenotype, although a function of NEAT1_1 under specific conditions cannot be excluded (Adriaens et al. 2019; Isobe et al. 2019). Further investigations may reveal specialized roles of NEAT1_1 under certain cellular conditions.

Several recent studies have suggested the presence of additional functional NEAT1 domains. Genome-wide

![Figure 4](Figure 4. The functional modular domain structure of the human NEAT1_2 arcRNA. The functional domains of NEAT1_2 and their positions (kb, kilobases from the 5' end of NEAT1_2) are shown. (PAS) Polyadenylation signal.)
mapping of R-loop-forming regions revealed that the 5'-terminal region of NEAT1 (~1400 nt) forms a long R-loop (DNA–RNA hybrid), although its biological importance remains unknown (Fig. 4; Dumelie and Jaffrey 2017). Another recent study showed that the 5'-terminal regions of NEAT1, including the GA- and CU-rich regions, form DNA:RNA triplex structures with many human genomic regions (Fig. 4; Senturk Cetin et al. 2019). Consistent with these two reports, the 5'-terminal sequence of NEAT1 is evolutionarily conserved (Fig. 3A), suggesting a sequence-based mechanism for the biological functions of NEAT1, including target gene recognition. In addition, the 3'-terminal region of human NEAT1_2 harbors a pseudomicroRNA (miR-612) that is not processed into mature microRNA (miRNA) but attracts microprocessors to process pri-miRNAs to pre-miRNAs in the nucleus (Jiang et al. 2017). Overall, as described in this section, the multiple modular domains of NEAT1 determine its biogenesis and the composition, biophysical properties, and functions of paraspeckles.

CURRENT MODEL OF PARASPECKLE FORMATION

Based on the studies described above, we propose a model of paraspeckle formation (Fig. 5). In this model, NEAT1_2 lncRNAs are transcribed by RNA polymerase II and PSPs associate with nascent NEAT1_2 transcripts at the NEAT1 gene loci (Fig. 5, step (1); Mao et al. 2011). PSPs such as NONO and SFPQ, which are essential for paraspeckle assembly, are loaded onto the multiple functionally redundant high-affinity binding sites of NEAT1_2 and then spread and coat NEAT1_2 domains to form base ribonucleoprotein (RNP) complexes through the NOPS and coiled coil domains of NONO and SFPQ (Fig. 5, steps (2) and (3); Lee et al. 2015; Yamazaki et al. 2018b). Subsequently, the NEAT1_2 RNPs further assemble into paraspeckles (50 NEAT1_2 RNPs per paraspeckle sphere), likely through phase separation. Concurrently, FUS and RBM14 are recruited to the RNPs, likely by NONO and SFPQ, and promote paraspeckle assembly via multivalent interactions between their PLDs (Fig. 5, steps (4) and (5); Hennig et al. 2015; West et al. 2016; Chujo et al. 2017; Yamazaki et al. 2018b). During this process, NEAT1_2 RNPs are bundled and folded (Fig. 5, step (5); West et al. 2016).

It is possible that NEAT1_2 lncRNAs interact with each other via direct intermolecular RNA–RNA interactions to contribute to the phase separation process (Fig. 5, step (2); Van Treeck and Parker 2018). This possibility is supported by the fact that weak NEAT1_2 focal signals remain detectable when PSPs are almost completely obliterated by 1,6-hexanediol treatment (Yamazaki et al. 2018b). Several studies have highlighted the importance of RNA–RNA interactions in the formation of cellular bodies, including stress granules, and disease-associated RNA foci caused by aberrant repeat expansions. It is proposed that the sum of the RNA–RNA, RNA–protein, and protein–protein interactions determines whether arcRNA–protein complexes undergo phase separation that eventually leads to cellular body formation (Van Treeck and Parker 2018).

FUTURE PERSPECTIVES

Despite recent advances in our understanding of how the NEAT1_2 domains and their partner proteins dictate paraspeckle formation and functions, many questions still remain. First, although NONO and SFPQ have been identified as partner proteins of the NEAT1_2 C domain, it is important to identify the partner RBPs and specific RNA motifs (e.g., sequences and/or secondary structures) required for proper functioning of the other NEAT1_2 functional domains, especially those involved in inducing phase separation. Second, as NEAT1_2 is involved in various physiological and pathological conditions, it is important to investigate how the molecular functions driven by specific RNA domains/motifs link to physiological functions.

We identified NONO and SFPQ as partner proteins of the NEAT1_2 C domain based on prior knowledge of the protein composition of paraspeckles and the results of experiments using immunofluorescence analyses of PSPs on NEAT1 mutants, MS2 tethering, CLIP-seq, and in vitro RNA pulldown (Yamazaki et al. 2018b). These approaches should also enable the identification of partner proteins of the other functional NEAT1_2 domains. To narrow down the functional domains of NEAT1_2 to precise RNA motifs, a large CLIP-seq data set, including eCLIP data, would be a great resource to differentiate between direct and indirect binding proteins and determine their binding sites (Van Nostrand et al. 2016). Analyses of the secondary structure of NEAT1_2 would also be important (Lu et al. 2016; Lin et al. 2018). For example, it would be useful to identify the specific RNA motifs in the NEAT1_2 C domain required for paraspeckle assembly via phase separation. It would also be interesting to understand how specific NEAT1_2 motifs are conserved between species. Furthermore, it would be helpful to understand how the specific domains and motifs within NEAT1_2 can initiate cellular body formation and determine their properties and functions.

Various molecular and physiological or pathological functions of NEAT1_2 have been reported. One of the molecular functions of paraspeckles is to regulate gene expression by acting as molecular sponges that incorporate specific RBPs and RNAs such as IRAIu-containing RNAs, CTN RNA, AG-rich RNAs, and miRNAs (Chen and Carmichael 2009; Hirose et al. 2014b; Imamura et al. 2014; West et al. 2016; Jiang et al. 2017). In addition, the global association of NEAT1/paraspeckles with chromatin has been shown by multiple methods, including CHART-seq, ChIA-PET, GRID-seq, and RADICL-seq (West et al. 2014; Cai et al. 2016; Li et al. 2017b; Bonetti et al. 2019). Under physiological conditions, NEAT1_2 is required for the establishment of pregnancy, mammary gland development, mitochondrial functions, and efficient regulation of gene expression for pluripotency and...
cellular differentiation (Nakagawa et al. 2014; Standaert et al. 2014; Wang et al. 2018; Modic et al. 2019). A number of studies have highlighted the importance of NEAT1 as a direct target of the p53 central tumor suppressor protein in cancer (Dong et al. 2018; Nakagawa et al. 2018). NEAT1 is also involved in various other pathological conditions, including neurodegenerative diseases and viral/bacterial infections (Nishimoto et al. 2013; Imamura et al. 2018). Additional research into the relationship between the RNA motifs of NEAT1 and its physiological functions is required to clarify the mechanisms of actions of arcRNAs and the biological significance of nuclear body formation on these molecules.

Several recent studies have shown that arcRNAs are still hidden within genomes. Using our recently developed

Figure 5. An updated model of paraspeckle formation showing the processes involved in NEAT1_2 transcription through paraspeckle formation via phase separation. See the main text for more details.
method that relies on the semiextractable feature of arcRNAs, we performed a genome-wide screen of HeLa cells and identified 45 candidate arcRNAs (Chujo et al. 2017). These candidates include IncRNAs, pre-mRNAs, and repeat-derived RNAs and form nuclear foci that are distinct from known nuclear bodies (Chujo et al. 2017). In addition, another recent study showed that short tandem repeat-containing RNAs (STRRNAs) might be a rich source of arcRNAs (Yap et al. 2018). One of the STRRNAs is a core of the perinucleolar compartment. As most arcRNAs are induced by various stimuli and/or specific physiological or pathological conditions, searches performed under various conditions will likely expand the repertoire of known arcRNAs.

Because known arcRNAs have partner proteins, several approaches can be used to identify the components of arcRNA-induced cellular bodies. Most PSPs identified to date were found by colocalization screening using a cDNA library and fluorescent fusion proteins (Naganuma et al. 2012; Fong et al. 2013; Mannen et al. 2016). Our recent comprehensive ChIRP-MS (chromatin isolation by RNA purification followed by mass spectrometry) analysis revealed the compositions of nSBs, among which most of the proteins (141) were newly identified (Ninomiya et al. 2019). Proximity labeling methods such as APEX and BioID have also been used to determine the compositions and dynamic compositional changes of phase-separated cellular bodies (Markmiller et al. 2018; Youn et al. 2018; Padron et al. 2019). These approaches could be used to obtain a comprehensive list of the components and their dynamic exchanges within cellular bodies under various cellular conditions. Future studies should also reveal the mechanisms by which cellular body compositions are determined by arcRNAs and their motifs.

It is well known that arcRNAs can induce phase transitions between several material states (e.g., liquid, hydrogel, and solid) of cellular bodies. Exactly how these material properties are determined remains unanswered. In particular, it is unclear which protein components contribute to each state. The binding sites of the proteins on arcRNAs, including their affinities and densities, are still undefined. The binding sites of the proteins on arcRNAs (141) were newly identified (Ninomiya et al. 2019). Proximity labeling methods such as APEX and BioID have also been used to determine the compositions and dynamic compositional changes of phase-separated cellular bodies (Markmiller et al. 2018; Youn et al. 2018; Padron et al. 2019). These approaches could be used to obtain a comprehensive list of the components and their dynamic exchanges within cellular bodies under various cellular conditions. Future studies should also reveal the mechanisms by which cellular body compositions are determined by arcRNAs and their motifs.

It is well known that arcRNAs can induce phase transitions between several material states (e.g., liquid, hydrogel, and solid) of cellular bodies. Exactly how these material properties are determined remains unanswered. In particular, it is unclear which protein components contribute to each state. The binding sites of the proteins on arcRNAs, including their affinities and densities, are still undefined, and the motifs and properties (e.g., amino acid compositions in PLDs, LCDs, or IDR) of the proteins that determine the material states also require clarification. In vitro reconstitution of arcRNA-induced bodies using purified components or cellular/nuclear extracts should elucidate how they are constructed and how their biophysical properties are determined (Mahanara et al. 2018). In addition to paraspeckles, other arcRNA-induced bodies also show various distinct morphologies. For example, nSBs display island structures that are composed of dense cores and shells (Kawaguchi et al. 2015), and amyloid bodies have filamentous structures (Audas et al. 2016). Therefore, it would be interesting to delineate how these morphologies are determined by arcRNAs and their partner proteins.

Identification of both the partner proteins and the precise binding sites as RNP modules should enable artificial control and reprogramming of arcRNAs. Motif-specific functional inhibition (or activation) might be achievable using antisense oligos, small molecules, or CRISPR-based strategies. In future analyses of arcRNAs, theoretical models and simulations, in addition to new methodologies and quantitative analyses, would offer novel frameworks to explain the mechanisms involved in complex biological phase separation.

ACKNOWLEDGMENTS

This work was supported by MEXT KAKENHI grants (to T.Y. [17K15058, 19K06479, 19H05250], S.N. [17H03604], and T.H. [JP26113002, JP16H06279, JP17H03630, and JP17K19335]), and by The Mochida Memorial Foundation for Medical and Pharmaceutical Research (to T.Y.), The Naito Foundation (to T.Y. and S.N.), The Akiyama Life Science Foundation (to T.Y.), Torey Research Fund (to S.N.), and Tokyo Biochemical Research Foundation (to T.H.).

REFERENCES

Adriaens C, Rambow F, Bervoets G, Silla T, Mito M, Chiba T, Hiroshi A, Hirose T, Nakagawa S, Jensen TH, et al. 2019. The IntraRNA NEAT1.1 is seemingly dispensable for normal tissue homeostasis and cancer cell growth. RNA 25: 1681–1695. doi:10.1261/ma.01456.119

Alberti S, Gladfelter A, Mittag T. 2019. Considerations and challenges in studying liquid-liquid phase separation and biomolecular condensates. Cell 176: 419–434. doi:10.1016/j.cell.2018.12.035

Aly MK, Ninomiya K, Adachi S, Natsume T, Hirose T. 2019. Two distinct nuclear stress bodies containing different sets of RNA-binding proteins are formed with HSATIII architectural noncoding RNAs upon thermal stress exposure. Biochem Biophys Res Commun 516: 419–423. doi:10.1016/j.bbrc.2019.06.061

Audas TE, Audas DE, Jacob MD, Ho JJ, Khacho M, Wang M, Perera JK, Gardiner C, Bennett CA, Head T, et al. 2016. Adaptation to stressors by systemic protein amyloidogenesis. Dev Cell 39: 155–168. doi:10.1016/j.devcel.2016.09.002

Balitz AG, Munschauer M, Schwanhausser B, Vasilie A, Munkawka Y, Schueller M, Youngs N, Penfold-Brown D, Drew K, Milek M, et al. 2012. The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. Mol Cell 46: 674–690. doi:10.1016/j.molcel.2012.05.021

Banani SF, Lee HO, Hyman AA, Rosen MK. 2017. Biomolecular condensates: organizers of cellular biochemistry. Nat Rev Mol Cell Biol 18: 285–298. doi:10.1038/nrm.2017.7

Bonetti A, Agostini F, Suzuki AM, Hashimoto K, Pascarella G, Gimenez J, Roos L, Nash AJ, Ghilotti M, Cameron CJF, et al. 2019. RADICL-seq identifies general and cell type-specific principles of genome-wide RNA-chromatin interactions. bioRxiv doi:10.1101/681924

Brown JA, Valenstein ML, Zingaro TA, Tycowski KT, Steitz JA. 2012. Formation of triple-helical structures by the 3′-end sequences of MALAT1 and MEN1 noncoding RNAs. Proc Natl Acad Sci USA 109: 19202–19207. doi:10.1073/pnas.1217338109

Cao L, Chang H, Fang Y, Li G. 2016. A comprehensive characterization of the function of LincRNAs in transcriptional regulation through long-range chromatin interactions. Sci Rep 6: 36572. doi:10.1038/srep36572

Castello A, Fischer B, Eichelbaum K, Horos R, Beckmann BM, Strein C, Davey NE, Humphreys DT, Press T, Steinmetz LM, et al. 2018. Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. Cell 149: 1393–1406. doi:10.1016/j.cell.2012.04.031

Castello A, Fischer B, Frese CK, Horos R, Alleaume AM, Foehr S, Curk T, Krijgsveel J, Hentze MW. 2016. Comprehensive
ARCHITECTURAL RNAs

identification of RNA-binding domains in human cells. Mol Cell 63: 696–710. doi:10.1016/j.molcel.2016.06.029

Chen L, Li S, Li G. 2009. Nuclear retention of mRNAs containing inverted repeats in human embryonic stem cells: functional role of a nuclear noncoding RNA. Mol Cell 35: 467–478. doi:10.1016/j.molcel.2009.06.027

Chu C, Zhang QC, da Rocha ST, Flynn RA, Bharadwaj M, Calabrese JM, Magnuson T, Heard E, Chang HY. 2015. Systematic discovery of Xist RNA binding proteins. Cell 161: 404–417. doi:10.1016/j.cell.2015.03.025

Chujo T, Hirose T. 2017. Nuclear bodies built on architectural RNAs: unifying principles of their construction and function. Mol Cells 40: 889–896. doi:10.14348/mol.cells.2017.0263

Chujo T, Yamazaki T, Hirose T. 2016. Architectural RNAs (arcRNAs): a class of long non-coding RNAs that function as the scaffold of nuclear bodies. Biochim Biophys Acta 1859: 139–146. doi:10.1016/j.bbamcr.2015.05.007

Chujo T, Yamazaki T, Kawaguchi T, Kurosaka S, Takumi T, Nakagawa S, Hirose T. 2017. Unusual semi-extractability as a hallmark of nuclear body-associated architectural noncoding RNAs. EMBO J 36: 1447–1462. doi:10.15252/embr.201695848

Clopton CM, Hutchinson JN, Sara SA, Ensminger AW, Fox AH, Chen J, Lawrence JB. 2009. An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. Mol Cell 33: 717–726. doi:10.1016/j.molcel.2009.01.026

Colognori D, Sunwoo H, Kriz AJ, Wang CY, Lee JT. 2019. Xist deletional analysis reveals an interdependency between Xist RNA and Polycomb complexes for spreading along the inactive X. Mol Cell 74: 101–117.e110. doi:10.1016/j.molcel.2019.01.015

Dong P, Xiong Y, Yue J, Hanley SJB, Kobayashi N, Todo Y, Watari H. 2018. Long non-coding RNA NEAT1: a novel target for diagnosis and therapy in human tumors. Front Genet 9: 471. doi:10.3389/fgene.2018.00471

Dumbović G, Biyana J, Bantis J, Samuelsson J, Roth A, Diedrichs S, Alonso S, Buschbeck M, Peruche M, Formele SV. 2018. A novel long non-coding RNA from NBL2 pericentromeric macrolocates forms a perinucleolar aggregate structure in colon cancer. Nucleic Acids Res 46: 55045–5524. doi:10.1093/nar/gky263

Dumelie JG, Jaffrey SR. 2017. Defining the location of promoter-associated R-loops at near-nucleotide resolution using bis-DIP-seq. eLife 6: e28306. doi:10.7554/eLife.28306

Engertz JM, Olikainen N, Guttman M. 2016. Long non-coding RNAs: spatial regulators that control nuclear structure and gene expression. Nat Rev Mol Cell Biol 17: 756–770. doi:10.1038/nrm.2016.126

Fong KW, Li X, Zhou B, Chen L, Gou LT, Li H, Fu XD. 2017b. GRID-seq regulates antibacterial noncoding RNAs. Mol Cell 69: 339–346. doi:10.1016/j.molcel.2017.02.072

Fox AH, Lam YW, Leung AK, Lyon CE, Andersen J, Mann M, Trewhella J, Fox AH, Bond CS. 2015. The structure of human PRP18 reveals a coiled-coil mediated polymer essential for protein sequestration within subnuclear bodies. Mol Biol Cell 25: 169–183. doi:10.1091/mbc.e13-09-0558

Frigyesi A, Xie J, Zhang QC, da Rocha ST, Flynn RA, Bharadwaj M, Chen LL, Carmichael GG. 2009. Altered nuclear retention of NEAT1 RNA is essential for the structural localization of NEAT1 long non-coding RNA: the domains, interactors, and biogenesis pathway required to build phase-separated nuclear paraspeckles. Wiley Interdiscip Rev RNA 10: e1545. doi:10.1002/wrna.e1545

Huang J, Casas Garcia GP, Perugini MA, Fox AH, Bond CS, Lee MC. 2018. Crystal structure of a SFQ/PSP1 heterodimer provides insights into preferential heterodimerization of human DHHS family proteins. J Biol Chem 293: 6593–6602. doi:10.1074/jbc.RA117.01451

Imamura K, Ishii T, Nakagawa S, Nakanuma T, Li R, Kimura H, Yokoi T, Nakagawa S, Bénard M, Fox AH, et al. 2014b. NEAT1 regulates nuclear paraspeckles. Nature 510: 189–193. doi:10.1038/nbt.3396

Jolly C, Usson Y, Morimoto RI. 1999. Rapid and reversible rec localization of heat shock factor 1 within seconds to nuclear stress granules. Proc Natl Acad Sci USA 96: 6769–6774. doi:10.1073/pnas.96.12.6769

Kawaguchi T, Tanigawa A, Naganuma T, Ohkawa Y, Souqueire S, Pierron G, Hirose T. 2015. SWI/SNF chromatin-remodeling complexes function in noncoding RNA-dependent assembly of nuclear bodies. Proc Natl Acad Sci USA 112: 4304–4309. doi:10.1073/pnas.141891112

Kang S, Sadowska A, Bekere I, Ho D, Gully BS, Lu Y, Iyer KS, Trewheella J, Fox AH, Bond CS. 2015. The structure of human SFQ reveals a coiled-coil mediated polymer essential for functional aggregation in gene regulation. Nucleic Acids Res 43: 3826–3840. doi:10.1093/nar/gkv283

Kavran J, Sadowska A, Schmidt BF, Bruchez MP, McManus CJ. 2018. Structural analyses of NEAT1 long noncoding RNA interactions that may contribute to paraspeckle architecture. Nucleic Acids Res 46: 3742–3752. doi:10.1093/nar/gky046
Lu Z, Zhang QC, Lee B, Flynn RA, Smith MA, Robinson JT, Davidovich C, Gooding AR, Goodrich JK, Mattick JS, et al. 2016. Rloop-based eQTL map links SNPs to RNA expression in human brain. *Cell* 165: 1267–1279. doi: 10.1016/j.cell.2016.04.028

Maharana S, Wang J, Papadopoulos DK, Richter D, Pozniakowski S, Pospischil M, Ritsch K, Guillen-Boixed J, Franzmann TM, et al. 2018. RNA buffers the phase separation behavior of prion-like RNA binding proteins. *Science* 360: 918–921. doi: 10.1126/science.aat7366

Mallik M, Lakhotia SC. 2009. RNAi for the large non-coding RNA: an emerging and disease-specific diversity in protein interactions with prion-like RNA binding proteins. *Cell Biol* 214: 45–59. doi: 10.1083/jcb.201602024

Mannen T, Yamashita S, Tomita K, Goshima N, Hirose T. 2016. The Sam68 nuclear body is composed of two RNA-sensitive substructures joined by the adaptor HNRNPL. *J Cell Biol* 165: 2721–2735. doi: 10.1083/jcb.201610010

Mao YS, Sunwoo H, Zhang B, Spector DL. 2011. Direct visualization of the co-transcriptional assembly of a nuclear body by noncoding RNAs. *Nat Cell Biol* 13: 95–101. doi: 10.1038/ncb2140

Markmiller S, Soltaehen S, Server KL, Mak R, Jin W, Fang MY, Luo EC, Krach F, Yang D, Sen A, et al. 2018. Context-dependent and ase-specific diversity in protein interactions within stress granules. *Cell* 172: 590–604 e513. doi: 10.1016/j.cell.2017.12.032

Modic M, Grosch M, Rot G, Schirge S, Lepko T, Yamazaki T, Lee FCY, Rusha E, Shaposhnikov D, Pal D, et al. 2019. Cross-regulation between TDP-43 and paraspeckles promotes pluripotency-differentiation transition. *Mol Cell* 74: 951–965. doi: 10.1016/j.molcel.2019.03.041

Monfort A, Wutz A. 2017. Progress in understanding the molecular mechanism of Xist RNA function through genetics. *Philos Trans R Soc Lond B Biol Sci* 372: 20163068. doi: 10.1098/rstb.2016.0368

Naganuma T, Nakagawa S, Tanigawa A, Sasaki YF, Goshima N, Hirose T. 2012. Alternative 3’-end processing of long noncoding RNA initiates construction of nuclear paraspeckles. *EMBO J* 31: 4020–4034. doi: 10.1038/emboj.2012.251

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

Nakagawa S, Yamazaki T, Hirose T. 2018. Molecular dissection of nuclear paraspeckles: towards understanding the emerging world of the RNP milieu. *Open Biol* 8: 180150. doi: 10.1098/rsob.180150

Ninomiya K, Adachi S, Natsume T, Iwakiri J, Terai G, Asai K, Nakagawa S, Yamazaki T, Hirose T. 2009. Molecular dissection of nuclear size and shape. *Cell* 165: 1267–1279. doi: 10.1016/j.cell.2016.04.028

Quinn JJ, Chang HY. 2016. Unique features of long non-coding RNA biogenesis and function. *Nat Rev Genet* 17: 47–62. doi: 10.1038/nrg.2015.198

Sasaki YT, Ideue T, Sano M, Mituyama T, Hirose T. 2009. MEN e/b noncoding RNAs are essential for structural integrity of nuclear paraspeckles. *Proc Natl Acad Sci* 106: 2525–2530. doi: 10.1073/pnas.0807899106

Schmitt AM, Chang HY. 2017. Long noncoding RNAs: at the intersection of cancer and chromatin biology. *Cold Spring Harbor Perspect Med* 7: a026492. doi: 10.1101/cshperspect.a026492

Senturk Cetin N, Kuo CC, Ribarska T, Li R, Costa IG, Grummt I. 2019. Isolation and genome-wide characterization of cellular DNA:RNA triplex structures. *Nucleic Acids Res* 47: 2306–2321. doi: 10.1093/nar/gky1305

Shevtsov SP, Dundr M. 2011. Nucleation of nuclear bodies by noncoding RNAs. *Nat Cell Biol* 13: 167–173. doi: 10.1038/ncb2157

Shin Y, Brangwynne CP. 2017. Liquid phase condensation in cell physiology and disease. *Science* 357: 1255. doi: 10.1126/science.aaf4382

Souquere S, Beaulclair G, Harper F, Fox A, Pierron G. 2010. Highly ordered spatial organization of the structural long non-coding NEAT1 RNAs within paraspeckle nuclear bodies. *Mol Cell Biol* 21: 4020–4027. doi: 10.1097/MCB.0b013e3181e08-0690

Standlaer L, Adriaen V, Van Kemenade A, Blanpain C, Hirose T, Nakagawa S, Marine JC. 2014. The long noncoding RNA Neat1 is required for mammary gland development and lactation. *RNA* 20: 1844–1849. doi: 10.1261/ma.047332.114

Sunwoo H, Dinger ME, Wilusz JE, Amaral PP, Mattick JS, Specter DL. 2009. MEN e/b nuclear-retained non-coding RNAs are up-regulated upon muscle differentiation and are essential components of paraspeckles. *Genome Res* 19: 347–359. doi: 10.1101/gr.087779.108

Tollervey JR, Curt K, Rogel J, Briese M, Cereda M, Kayikci M, Konig J, Hortobagyi T, Nishimura AL, Zupunski V, et al. 2011. Characterizing the RNA targets and position-dependent splicing regulation by TDP-43. *Nat Neurosci* 14: 452–458. doi: 10.1038/nn.2778

Van Nostrand EL, Pratt GA, Shishkin AA, Gelboin-Burkhart C, Fang MY, Sundararaman B, Blue SM, Nguyen TB, Surka C, Elkins K, et al. 2016. Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). *Nat Methods* 13: 508–514. doi: 10.1038/nmeth.3810

Van Treeck B, Parker R. 2018. Emerging roles for intermolecular RNA:RNA interactions in RNA-BPs assemblies. *Cell* 174: 791–802. doi: 10.1016/j.cell.2018.07.023

Visa N, Puvion-Dutilleul F, Bachelier JP, Puvion E. 1993. Intra-nuclear 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.

Wang Y, Hu SB, Wang MR, Yao RW, Wu D, Yang L, Chen LL. 2018. Genome-wide screening of NEAT1 regulators reveals cross-regulation between paraspeckles and mitochondria. *Nat Cell Biol* 20: 1145–1158. doi: 10.1038/s41556-018-0124-2

West JA, Davis CP, Sunwoo H, Simon MD, Sadreyev RI, Wang PI, Tolstorukov MY, Kingston RE. 2014. The long noncoding RNAs NEAT1 and MALAT1 bind active chromatin sites. *Mol Cell* 55: 791–802. doi: 10.1016/j.molcel.2014.07.012

West JA, Mito M, Kurosa K, Takumi T, Tanegashima C, Chuijo T, Yanaka K, Kingston RE, Hirose T, Bond C, et al. 2016. Structural, super-resolution microscopy analysis of paraspeckle nuclear body organization. *J Cell Biol* 214: 817–830. doi: 10.1083/jcb.201601071

Wilusz JE, JinKistapser K, Lu Y, Kuhn CD, Joshua-Tor L, Sharp PA. 2012. A triple helix replaces the 3’ ends of long noncoding RNAs that lack poly(A) tails. *Genes Dev* 26: 2392–2407. doi: 10.1101/gad.204438.112

Wutz A, Rasmussen TP, Jaenisch R. 2002. Chromosomal silencing and localization are mediated by different domains of Xist RNA. *Nat Genet* 30: 167–174. doi: 10.1038/ng820
Yamashita A, Watanabe Y, Nukina N, Yamamoto M. 1998. RNA-assisted nuclear transport of the meiotic regulator Mei2p in fission yeast. *Cell* 95: 115–123. doi:10.1016/S0092-8674(00)81787-0

Yamazaki T, Hirose T. 2015. The building process of the functional paraspeckle with long non-coding RNAs. *Front Biosci (Elite Ed)* 7: 1–41. doi:10.2741/s420

Yamazaki T, Fujikawa C, Kubota A, Takahashi A, Hirose T. 2018a. CRISPRa-mediated NEAT1 IncRNA upregulation induces formation of intact paraspeckles. *Biochem Biophys Res Commun* 504: 218–224. doi:10.1016/j.bbrc.2018.08.158

Yamazaki T, Souquère S, Chujo T, Kobelke S, Chong YS, Fox AH, Bond CS, Nakagawa S, Pierron G, Hirose T. 2018b. Functional domains of NEAT1 architectural IncRNA induce paraspeckle assembly through phase separation. *Mol Cell* 70: 1038–1053.e1037. doi:10.1016/j.molcel.2018.05.019

Yang L-Z, Wang Y, Li S-Q, Yao R-W, Luan P-F, Wu H, Carmichael GG, Chen L-L. 2019. Dynamic imaging of RNA in living cells by CRISPR-Cas13 systems. *Mol Cell* doi:10.1016/j.molcel.2019.10.024

Yap K, Mukhina S, Zhang G, Tan JSC, Ong HS, Makeyev EV. 2018. A short tandem repeat-enriched RNA assembles a nuclear compartment to control alternative splicing and promote cell survival. *Mol Cell* 72: 525–540.e513. doi:10.1016/j.molcel.2018.08.041

Youn JY, Dunham WH, Hong SJ, Knight JDR, Bashkurov M, Chen GI, Bagci H, Rathod B, MacLeod G, Eng SWM, et al. 2018. High-density proximity mapping reveals the subcellular organization of mRNA-associated granules and bodies. *Mol Cell* 69: 517–532.e511. doi:10.1016/j.molcel.2017.12.020

Zhang N, Ashizawa T. 2017. RNA toxicity and foci formation in microsatellite expansion diseases. *Curr Opin Genet Dev* 44: 17–29. doi:10.1016/j.gde.2017.01.005

Zhu J, Mayeda A, Krainer AR. 2001. Exon identity established through differential antagonism between exonic splicing silencer-bound hnRNP A1 and enhancer-bound SR proteins. *Mol Cell* 8: 1351–1361. doi:10.1016/S1097-2765(01)00409-9