Paraspeckles: where long noncoding RNA meets phase separation

Archa H Fox*¹, ⁴, Shinichi Nakagawa², Tetsuro Hirose³ and Charles S. Bond⁴
1. School of Human Sciences, University of Western Australia, Crawley, Western Australia, 6009 and Harry Perkins Institute of Medical Research, Nedlands, Western Australia, Australia
2. Faculty of Pharmaceutical Sciences, Hokkaido University, Hokkaido, Japan
3. Institute for Genetic Medicine, Hokkaido University, Hokkaido, Japan
4. School of Molecular Sciences, University of Western Australia, Crawley, Western Australia, 6009, Australia
*Correspondence: archa.fox@uwa.edu.au (Archa Fox)

Keywords
Paraspeckles, long noncoding RNA, membraneless organelles, RNA-protein interactions, gene regulation

Abstract
Long noncoding RNA molecules are some of the newest and least understood players in gene regulation. Hence we need good model systems with well defined RNA and protein components. One such system is paraspeckles - protein-rich nuclear organelles built around a specific long noncoding RNA scaffold. New discoveries show how paraspeckles are formed through multiple RNA-protein and protein-protein interactions, some of which involve extensive polymerisation, and others with multivalent interactions driving phase separation. Once formed, paraspeckles influence gene regulation through sequestration of component proteins and RNAs, with subsequent depletion in other compartments. Here we focus on the dual aspects of paraspeckle structure and function, revealing an emerging role for these dynamic bodies in a multitude of cellular settings.
Main text

New insights into paraspeckles have helped us understand both lncRNA structure/function and membraneless organelle formation

Dynamic cellular organelles that form without a membrane boundary have long fascinated scientists. What triggers such structures to form? How do a whole variety of specific molecules get targeted there? How are such structures maintained in steady state, yet also rapidly broken down when triggered to do so? The nucleoplasm is rich in such membraneless organelles, often termed ‘nuclear bodies’. These bodies regulate gene expression, albeit with mechanisms that have been hard to pin down. Another area that has seen a surge in interest recently is the field of long noncoding RNA (lncRNA) biology – this has arisen with the knowledge that the majority of our genome is transcribed, but not translated, yet we have little idea what all this RNA is doing. Working out the functions of these lncRNAs, how proteins interact with them and how they can be manipulated has been the focus of much high profile research. These two important fields - membraneless organelles and lncRNA biology - coincide spectacularly in the field of nuclear organization, particularly in one important model system - the paraspeckle - a nuclear body built on a specific lncRNA.

Discoveries from the last few years have revealed that paraspeckle formation is dependent on the lncRNA ‘seed’ to recruit specific component proteins by RNA-protein interactions, thus forcing a local high concentration at a specific nuclear location. Once in close proximity and bound to their RNA scaffold, these proteins oligomerise and recruit additional proteins that use low complexity domains to mediate a liquid liquid phase separation (LLPS) process and formation of the paraspeckle. Despite this liquid-like state, there is evidence of exquisite substructure in paraspeckles: super-resolution fluorescence microscopy reveals distinct ‘core’ and ‘shell’ regions containing specific parts of the lncRNA and distinct proteins. In addition, we have new understanding about how these lncRNA-seeded bodies regulate gene expression in a dynamic way in different physiological settings,
particularly in cancer and cell stress scenarios. In this review we summarise recent literature on paraspeckles, focusing on these novel developments.

**Setting the scene – what are paraspeckles?**

Paraspeckles are mammal-specific RNA-protein nuclear bodies that regulate gene expression. First described in 2002 as nuclear foci enriched in the marker protein PSPC1 (paraspeckle component protein 1) [1], paraspeckles are now known to contain over 40 different proteins (generally ubiquitously expressed RNA binding proteins) and one architectural long noncoding RNA, NEAT1 (nuclear paraspeckle assembly transcript 1, see Table 1). Some component proteins mediate critical RNA-protein or protein-protein interactions essential for paraspeckle formation (marked in Table 1). In addition to these structural components there are several different molecules targeted to paraspeckles, such as various types of RNAs, as well as proteins that are not required to build paraspeckles, but nevertheless may be regulated by them.

Paraspeckles of varying size and abundance are found in the nucleus of most cultured cells, including primary and transformed cell lines, but are absent from embryonic stem cells (ESC) and induced pluripotent stem cells (iPSC) [2]. Paraspeckles have been detected in many different tissues, however, prominent paraspeckles are found only in a sub-population of cells in murine tissues, such as luteal cells in the ovary and cells at the tip of gut epithelium [3].

Although paraspeckles were first defined as sites of enrichment of PSPC1, subsequent siRNA knockdown experiments showed PSPC1 was in fact dispensable for paraspeckle formation [4]. However, two other closely-related proteins, NONO and SFPQ, are absolutely required for paraspeckle formation, along with several other proteins: HNRNPK, DAZAP1, FUS, RBM14, HNRNPH3 – knockdown of any of these proteins in HeLa cells results in paraspeckle ablation ([5], table 1). Given this seeming minor role for the original paraspeckle marker protein, here we propose a definitive definition of a paraspeckle: a nuclear body in which one of the essential paraspeckle proteins co-localises with the longer isoform of NEAT1.

**The longer isoform of NEAT1 is the structural backbone of paraspeckles**
The long non-coding RNA (lncRNA), NEAT1, is an essential structural component of paraspeckles, forming a scaffold for protein binding [2, 4, 6, 7]. The requirement of NEAT1 for paraspeckle formation was one of the first concrete examples of a functional lncRNA at a time when controversy over their functionality was still raging. As such, paraspeckles have now become one of the most well studied lncRNA-protein complexes in the cell.

Two transcripts are generated by RNA Polymerase II from the NEAT1 gene: NEAT1_1 (3.7kb in humans, 3kb in mice) completely overlaps the 5’ end of NEAT1_2 (23kb in humans, 20kb in mice, Figure 1). Both transcripts contain a single exon and are nucleus-retained. NEAT1_1 is polyadenylated in a manner akin to a typical mRNA, whereas NEAT1_2 has an unusual triple helix structure at the 3’ end that has been used as a characteristic feature for lncRNAs of this class [8]. It is the transcription of NEAT1_2 that is critically important for paraspeckles, indeed NEAT1_2 forms the paraspeckle ‘architectural’ backbone [3, 5, 9, 10]. Paraspeckles fail to form in NEAT1 knockout mice, and only overexpression of NEAT1_2 can rescue paraspeckle formation in murine embryonic fibroblasts from these mice [5]. Given its central role in paraspeckle structure, it is important to know how NEAT1_2 is made (Figure 1). *In vitro* and cellular approaches demonstrated that transcription of NEAT1_2 results from competition between the cleavage factors that would otherwise cleave and polyadenylate NEAT1_1 being prevented from doing so by the essential paraspeckle protein HNRNPK [5]. Mutation of the NEAT1_1 polyA site with genome editing can also mimic this effect [11]. In the absence of transcriptional termination and cleavage to make NEAT1_1, RNA Polymerase II continues to transcribe along the DNA, generating the longer NEAT1_2 transcript. Beyond this initial NEAT1_2 based seeding event, active transcription of NEAT1_2 is required for paraspeckle formation and maintenance, suggesting additional *trans* acting factors that sense NEAT1_2 production [9]. Additional questions remain as to the role of NEAT1_1 in paraspeckle biology, with new data suggesting it may have a paraspeckle independent role [11]. Given these distinctions in NEAT1 isoforms, we urge researchers to consider these differences when designing their experiments and not assume that NEAT1_1 overexpression, or knockdown in the absence of NEAT1_2, necessarily influences paraspeckles.
Paraspeckles are chains of spheroids, each with a core and shell structure

Paraspeckles are readily labelled with specific marker proteins and RNAs using fluorescent microscopy, and under the EM are also distinguishable as electron dense structures in the inter-chromatin space. Despite initially being observed as punctate foci, EM and super-resolution microscopy has revealed paraspeckles are actually elongated prolate structures with a uniform width of approximately 360 nm in human cells and slightly less in murine cells, but with length varying up to 1-2 microns [10]. The number and length of paraspeckles is directly proportional to NEAT1_2 transcription, with more and longer paraspeckles apparent as NEAT1_2 levels increase [12].

EM and super-resolution microscopy have been used to reveal that paraspeckles are composed of fused chains of spherical subcompartments each of which have a distinct core-shell molecular organisation (Figure 2). Fluorescent In situ hybridisation (FISH) against different regions of NEAT1_2 in the super-resolution microscope revealed that the 5’ and 3’ ends of this 23,000 nucleotide RNA are clearly confined to the shell of the paraspeckle, whereas the central region of the RNA is located in the core [13]. By combining differently coloured probes to either the 3’ or 5’ ends of NEAT1_2, it was shown that the ends do not individually intermix, but are rather separated into bundles [13]. Immunofluorescence showed several proteins localised into distinct zones of the paraspeckle: the essential proteins NONO, SFPQ, FUS are located in the core, with RBM14 and BRG1 in patches, whereas TDP43 is located in the shell with the 5’ and 3’ ends of NEAT1 ([13] Figure 2). Although such ultrastructural insights are largely descriptive, they do provide a critical framework for studies that aim to understand how this substructure is established in paraspeckle biogenesis.

How do paraspeckles form?

Paraspeckle biogenesis starts with transcription of NEAT1_2, with paraspeckles forming close to the NEAT1 gene and often appearing clustered near this locus [6]. Once NEAT1_2 is made, the essential paraspeckle proteins NONO and SFPQ rapidly bind to form and stabilise a minimal NEAT1_2 RNP particle that is an intermediate in paraspeckle formation ([13], Figure 1). It is not yet known precisely which RNA sequence or structure in NEAT1_2 is responsible for NONO/SFPQ
binding, but several CLIP (Cross-linking and immunoprecipitation) datasets indicate extensive binding of the proteins along the length of NEAT1_2 [14]. Again, although the precise RNA binding modalities are not known, NONO and SFPQ both contain classical RRM domains that are important in paraspeckle targeting [15]. It is hypothesised that initial RNA-protein interaction seeding events then trigger polymerisation of NONO and SFPQ along the RNA. Indeed, coiled-coil domains in NONO and SFPQ form extensive polymers that are important for the integrity of paraspeckles [16], with contribution from a conserved charged single charged helix region that may be involved in the molecular spacing of these proteins along the NEAT1_2 RNA [17] and low complexity domains that may be involved in protein-protein interactions. This initial protein binding is thought to be important for NEAT1_2 stability as both isoforms of NEAT1 have a relatively short half life [18, 19].

Additional core proteins - principally FUS - are required to knit together these NONO/SFPQ/NEAT1_2 RNPs into a mature paraspeckle [13, 20], Figure 1). Interestingly, RBM14 and the swi/snf component BRG1 are both found in patches bridging the core and shell regions of the paraspeckle and these two proteins are both required for essential additional protein-protein interactions to form paraspeckles [21, 22]. However, BRG1 does not require its ATPase catalytic site for this structural role. In terms of stoichiometry, recent quantitative analysis revealed the presence of ~50 NEAT1_2 molecules in a single paraspeckle particle [23], but the stoichiometry of different paraspeckle proteins is yet to be determined.

The proteins that build paraspeckles have RNA binding and polymerisation propensity that is also observed in other contexts. For example, NONO and SFPQ can be rapidly recruited to artificial scaffolds such as transfected antisense oligonucleotides to form ‘pseudo-paraspeckles’, devoid of NEAT1 [24]. However, it remains to be tested if the ordered core-shell structures of paraspeckle proteins is also found in the protein assembly induced by these artificial scaffolds.

**Liquid liquid phase separation is critical for paraspeckle formation**

Paraspeckles actually share many features with cytoplasmic stress granules, another type of membraneless organelle. Paraspeckles and stress granules both contain common component proteins [25], become more abundant with stress, seem to
function through molecular sequestration (more on this below) and both have distinct subsets of molecules found either within the core, or shell regions [13, 26]. Our understanding of the type of molecular interactions driving formation of such membraneless organelles has been changing with the appreciation that these structures may be more liquid than solid, indeed they could be undergoing liquid liquid phase separation (LLPS). Exciting work on protein low complexity domains - domains previously thought to be flexible linkers or regions incapable of folding - show that these regions can mediate LLPS [27]. The rapid aggregation of these proteins can result in a continuum of states, from monomers, to liquids, to fibrillar amyloid-like structures that contain cross beta interactions [28]. However, unlike pathological amyloid, such fibrillar structures are capable of being broken down. These fibrillar structures can form hydrogels in vitro [29]. Interestingly, nuclear bodies have long been observed to have liquid like properties involving fusion and splitting [30].

The sub-class of low complexity domain that is responsible for fibril formation is the ‘prion-like domain’, named due to the similarity to the yeast prion protein. Found in almost half of the paraspeckle proteins (Table 1), prion-like domains are enriched in uncharged polar amino acids such as asparagine, glutamine, tyrosine, as well as glycine [31]. Of the proteins responsible for protein-protein interactions in the paraspeckle interactome, two thirds contain prion-like domains, indicating that a network of prion-like domain mediated interactions may be extensive within paraspeckles [21]. A combination of in vitro and cell experiments were used to show that two essential paraspeckle proteins, RBM14 and FUS, both require an intact prion-like domain to be targeted to, and effectively build, paraspeckles [13, 21, 23]. In these experiments, mutants of FUS and RBM14 with Y to S substitutions in their prion like domains (mutations that compromise the ability of these domains to form fibrils and hydrogels) were incapable of forming paraspeckles [21]. Additional cellular studies on FUS, TDP43 and HnRNPA1 all attest to the importance of their prion-like domains for liquid phase separation [32-34].

Remarkably, the association of NEAT1_2 with the paraspeckle proteins is so extensive that the RNA is not easily extracted with conventional RNA solubilisation methods such as Trizol. Heating and needle-shearing is required to extract the RNA
fully from the protein phase, leading to an up to 10 fold increase in NEAT1_2 extraction. Indeed, this ‘semi-extractability’ of NEAT1_2 is dependent on the prion like domain of FUS, as cells expressing FUS truncation mutants that lack the prion-like domain have freely extractable NEAT1_2 [23]. It is therefore intriguing to speculate that the structural basis of this semi-extractability of NEAT1_2 may involve some form of FUS-driven phase separation.

**What is the molecular mechanism of paraspeckle function?**
Once paraspeckles form, what is their role in the cell? As yet, no distinct catalytic activity has been demonstrated to occur within the paraspeckle. The main molecular mechanisms identified for paraspeckle function relate to sequestration of molecules, be it RNA or proteins (summarised in Figure 3).

*Paraspeckles trap certain RNA species in the nucleus*
Beyond NEAT1, other types of RNA species can localise to paraspeckles and potentially be regulated by them. The best characterised class of such nucleus-retained RNAs contain double-stranded RNA structures formed through transcribed inverted repeat motifs [35]. This nuclear retention mechanism is driven by NONO and SFPQ binding to the dsRNA structure in the RNA. Recent studies have shown that methylation of the coiled-coil region of NONO is important for regulating its ability to bind these dsRNAs [36]. How might this nuclear retention mechanism be used in cell physiology? One example has emerged for paraspeckles having a role in the post-transcriptional control of gene regulation in a circadian context. Rhythmic expression of paraspeckle/NEAT1_2 levels is observed in pituitary cells and knockdown of NEAT1_2 prevents the rhythmic expression of inverted repeat-fused reporters [37].

*Paraspeckles quantitatively sequester their protein components*
As NEAT1_2 levels increase, paraspeckles elongate ([12], Figure 3). A consequence of increasing paraspeckle length is that a greater amount of constituent paraspeckle protein is required to build the structures, with a concomitant reduction in nucleoplasmic protein availability. This concept has been well defined for the paraspeckle protein SFPQ. As paraspeckle elongate, requiring more SFPQ, the relative levels of SFPQ decrease in the nucleoplasm. The consequence of SFPQ nucleoplasmic depletion is alteration in its direct target genes. For example, SFPQ...
normally represses IL-8, an immune responsive gene, or activates the ADARB2 gene [12, 38]. For paraspeckles, SFPQ is the only molecule to have this detailed study performed, but it is presumed that all paraspeckle proteins will be regulated by changing paraspeckle size in a similar manner. Other nuclear bodies are known to act with a similar subnuclear sequestration mechanism. For example, nuclear bodies seeded by the SPA IncRNA (with a 5' snoRNA cap and 3' polyadenylation) that trap and deplete other specific RNA binding proteins from the nucleoplasm [39].

**New paraspeckle roles in transcription and miRNA processing**

A further possibility exists for paraspeckles to directly dock onto the transcription start sites of actively transcribed genes. A CHART (capture hybridization analysis of RNA targets) experiment showed that NEAT1 RNA crosslinked to the transcription start sites of active genes in a breast cancer cell line, but the probes targeted both NEAT1_1 and NEAT1_2 therefore it is unclear if this is a paraspeckle specific activity, or NEAT1_1 driven activity, being revealed [40]. Further, SFPQ and NONO were recently found to have activity in pri-miRNA processing, enhancing the processing of introns containing pri-miRNA, leading to increasing mature miRNAs [14]. By sequestration of SFPQ and NONO into paraspeckles, along with some of the miRNA processing factors, the miRNA processing is positively influenced [14]. It is thought in this instance that the association of the proteins into the paraspeckle is positively enhancing this pri-miRNA processing capability of the proteins by providing an efficient platform for their interactions (Figure 3).

**Paraspeckles play a role in a variety of developmental and disease scenarios**

In which biological settings might paraspeckles play a role in regulating gene expression? The paraspeckle proteins can be found both inside and outside paraspeckles, therefore study of paraspeckle protein function cannot be seen as akin to paraspeckle function. In contrast, NEAT1_2 is strictly only found within paraspeckles, and NEAT1_2 levels can be considered a proxy for paraspeckle appearance. Thus observation and manipulation of NEAT1_2 is a route to functional studies of paraspeckles. Here we have distilled functional information about paraspeckles, focusing on studies that have distinguished the NEAT1 isoforms and/or conducted microscopic imaging to detect paraspeckles in different biological settings.
**Paraspeckles are important in female reproduction**

Paraspeckles and Neat1 are not essential for mammalian development under normal conditions - it was originally reported that mice lacking Neat1 were viable [3]. Subsequent analyses revealed that a subset of female Neat1-/- mice had impaired formation of the transient ovarian secretory gland, the corpus luteum, that supports pregnancy (in wildtype mice the corpus luteum has large, prominent paraspeckles). Lack of a formation of pregnant corpus luteum leads to infertility, smaller litters and fewer viable pregnancies for knockout females [41]. There may be an evolutionary conserved function for paraspeckles in secretory structures of the female reproductive system: the marsupial opossum has distinct paraspeckles in the uterine gland, however their functional significance is not known [42]. Neat1 knockout mice also exhibit deficiencies in mammary epithelial cell proliferation leading to a lactation defect in the mothers and low pup survival [43]. Curiously, some Neat1 knockout phenotypes have a stochastic effect, for example the corpus luteum will form in some, but not all female knockouts. These stochastic effects, along with in vitro studies indicating stress induction of paraspeckles through environmental triggers has led to a theme that NEAT1/paraspeckles exert their functional roles when cells are stressed.

**Paraspeckles increase with viral infection**

NEAT1_2 levels have been observed increasing in different cell types following infection with many different RNA viruses, including Japanese encephalitis, rabies [44], HIV [45, 46], influenza [38] and hantaan [47], as well as the DNA-encoded herpes simplex virus [48]. In most cases the evidence suggests that NEAT1_2/paraspeckles increase as a cell defence mechanism. More abundant paraspeckles lead to increased sequestration of the paraspeckle proteins, thereby preventing the hijacking of these proteins by the virus [48, 49], or triggering gene regulatory changes that help the cell [38, 47].

**Are paraspeckles friend or foe in cancer?**

Whilst many studies implicate the molecule NEAT1 in cancer progression, there are few studies directly connecting paraspeckles, or NEAT1_2, to cancer biology. A meta analysis of NEAT1 in cancer revealed a complex and varying pattern of expression
revealing that NEAT1 can be either upregulated, downregulated, or unchanged in tumor compared to normal tissue, depending on different studies and cancer models, with most studies not differentiating between NEAT1 isoforms [50]. Despite this, the possibility for a role for paraspeckles in cancer is gaining more credence. The NEAT1 gene, including the region encoding NEAT1_2, contains a significantly large number of mutations in tumors from liver cancer patients [51], and the NEAT1 promoter was recently found to be a hotspot for mutations in breast cancer [52]. Overall several different cancer subtypes show increasing numbers of mutations in NEAT1_2 [53]. A common theme has emerged of NEAT1/paraspeckles being induced by stress in pre-neoplastic, or tumor cells. In many such cases, however, a function is only apparent when stress is applied to the cell.

In many settings paraspeckles seem to have oncogenic roles. Hypoxia triggers increased NEAT1_2 transcription through hypoxia-inducible factor (HIF) with a resultant increase in paraspeckles in breast cancer cell lines [54]. In prostate cancer, increased levels of NEAT1_1 as well as NEAT1_2/paraspeckles are associated with advanced forms of the disease [55]. In this case, NEAT1 transcription is stimulated by the estrogen receptor (ER) and together NEAT1 and ER promote transcription of a circuit of proliferative genes. Overexpression of NEAT1_1 increased active epigenetic marks at target genes and triggered proliferation, with loss of both NEAT1 isoforms reducing proliferation [55]. In skin fibroblasts, p53 induces NEAT1_2/paraspeckles when pre-neoplastic cells transform into tumors, either in genetic or topical skin cancer mouse models, and loss of NEAT1 prevents this transformation, suggesting NEAT1 is a potent oncogene [56]. In that study NEAT1 and paraspeckles prevented replication stress, aiding the DNA damage response, although the molecular mechanism was not elucidated.

In opposition to observations of paraspeckles as oncogenic, other studies find paraspeckles have tumor suppressor roles. P53 induction of NEAT1_1 and NEAT1_2/paraspeckles in cells derived from RAS-driven pancreatic cancer mouse models prevented transformation in vitro and also prevented pancreatic cancer progression in murine models [57]. Interestingly, overexpression of NEAT1_1 in p53-/- cells partially restored the p53-mediated suppression of transformation, suggesting that in this context, NEAT1_1 may be playing a role in tumor suppression, rather than
paraspeckles per se. In different types of blood cancers, NEAT1 levels correlate inversely with proliferation [58]. In colorectal cancer cell lines, inhibition of NEAT1_2 enhanced proliferation, suggesting paraspeckles could be acting as a tumor suppressor in this context [59].

**Paraspeckles in the nervous system**

Some areas of the mouse brain show increased NEAT1_2 levels following seizures [60] and experiments in iPSC-derived neurons indicate that NEAT1 function, but not necessarily paraspeckles per se, may regulate hyperexcitability and have links to epilepsy [61]. Paraspeckles may also have a role in motor neuron disease (MND, also known as amyotrophic lateral sclerosis, or ALS). Several paraspeckle proteins are encoded by genes mutated in familial forms of ALS (Table 1). When ALS patient derived mutants of FUS are expressed in cell models, cytoplasmic aggregates form that can also be enriched for NONO and SFPQ [20]. Motor neurons do not normally express Neat1_2, but its expression is upregulated in early-phase degenerating motor neurons of ALS patients [62]. Overall, however, the physiological significance of the formation of paraspeckles during ALS pathogenesis remains to be elucidated.

**Concluding Remarks and Future Perspectives**

In the 15 years since their discovery, paraspeckles have flourished from obscure nuclear body to important IncRNA-protein model system with novel roles in gene regulation in different disease contexts. By studying how paraspeckles form we are learning more about how nuclear proteins use their RNA binding and protein-protein aggregation propensity to form such membraneless organelles. With many questions still remaining about paraspeckle biology (see ‘outstanding questions’) we are confident that the field will continue to grow, providing a rich source of novel understanding into gene regulation.

**Figure legends**

**Figure 1. Biogenesis of a paraspeckle.** Generation of the overlapping NEAT1 isoforms (A) is first triggered by activation of the NEAT1 promoter and recruitment of RNA polymerase II. Once the sequence corresponding to NEAT1_1 is transcribed, a canonical polyA site at the NEAT1_1 3’ end recruits cleavage and polyadenylation
factors that will generally terminate transcription (B). In some instances, this cleavage and polyadenylation is prevented from occurring by the inhibition of these processes by HNRNPK, which binds to the same region of NEAT1_1 (C). When this happens, RNA Polymerase II continues to transcribe through into the NEAT1_2 region, generating the longer 23kb NEAT1_2 transcript that contains the specific binding sites for NONO and SFPQ (D) that bind and polymerise along the RNA forming NEAT1_2/NONO/SFPQ RNA-protein bundles (E). These bundles are then fused into a microscopically visible mature paraspeckle by FUS, with an activity dependent on the FUS prion like low complexity domain (F). (G) A confocal image of a HeLa cell stained with an antibody to the paraspeckle protein, NONO (green), with the fluorescent signal overlaid on a bright-field image of the cell. Arrows indicate paraspeckles. Scale bar 5 µm.

**Figure 2. Paraspeckle substructure contains core and shell zones.**
Fluorescence micrograph on the left shows a super-resolution/structured illumination image of paraspeckles from murine corpus luteal cells with the shell (green) highlighted by FISH with combined probes recognising the 5’ and 3’ ends of Neat1_2, and the core region (magenta) highlighted by FISH with probes to the middle sequences of Neat1_2. Both single mature paraspeckles, as well as chains of paraspeckle units, are depicted. Scale bar 500 nm. The model on the right summarises super-resolution imaging, depicting the relative locations of different paraspeckle component proteins and RNA regions to either the core or shell zones of the paraspeckle. Image and model adapted from [13].

**Figure 3. Paraspeckles regulate gene expression by sequestration and enhancing the efficiency of paraspeckle protein function.** The steady state is depicted in the centre, with paraspeckle proteins and targeted structured RNAs found both inside paraspeckles, as well as elsewhere in the nucleus, including regulating gene transcription at promoters. When the paraspeckle proteins are organised inside paraspeckles there is evidence for enhanced processing of pri-miRNA transcripts. The right panel depicts that happens when cells become stressed, or NEAT1_2 is increased. In this instance, paraspeckles elongate, increasing the sequestration of paraspeckle proteins, with the pool of nucleoplasmic paraspeckle protein diminishing,
resulting in a change in transcription of target genes. The left panel shows what happens when NEAT1_2 levels decrease, either in a particular stage of the circadian cycle, in cells that do not express NEAT1_2, or when NEAT1_2 is artificially reduced. In this instance, paraspeckle proteins and regulated RNAs are free to bind to target gene promoters, or, in the case of regulated RNAs, to be exported from the nucleus and used as templates for translation. The efficiency of pri-miRNA processing by paraspeckle proteins is diminished.

**Acknowledgements**

We are grateful to the members of our laboratories for helpful reading of this manuscript and apologize to those whose work we could not cite due to space limitations. This work was supported by the Australian Research Council grant DP160102435 to AHF and CSB.
Table 1 – Paraspeckle proteins and RNAs

<table>
<thead>
<tr>
<th>gene name</th>
<th>importance in paraspeckle formation</th>
<th>prion like domain</th>
<th>ALS2-mutation link</th>
<th>paraspeckle zone</th>
<th>reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>AHDC1</td>
<td>dispensible</td>
<td>no</td>
<td></td>
<td></td>
<td>[5]</td>
</tr>
<tr>
<td>AKAP8L</td>
<td>dispensible</td>
<td>yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CELF6</td>
<td>n. d.</td>
<td>no</td>
<td></td>
<td></td>
<td>[5]</td>
</tr>
<tr>
<td>CIRBP</td>
<td>dispensible</td>
<td>no</td>
<td></td>
<td></td>
<td>[5]</td>
</tr>
<tr>
<td>CPSF5</td>
<td>dispensible</td>
<td>no</td>
<td></td>
<td></td>
<td>[5]</td>
</tr>
<tr>
<td>CPSF6</td>
<td>dispensible</td>
<td>no</td>
<td></td>
<td></td>
<td>[5]</td>
</tr>
<tr>
<td>CPSF7</td>
<td>important</td>
<td>no</td>
<td></td>
<td></td>
<td>[5]</td>
</tr>
<tr>
<td>DAZAP1</td>
<td>essential</td>
<td>yes</td>
<td></td>
<td></td>
<td>[5]</td>
</tr>
<tr>
<td>DLX3</td>
<td>n. d.</td>
<td>yes</td>
<td></td>
<td></td>
<td>[5]</td>
</tr>
<tr>
<td>EWSR1</td>
<td>dispensible</td>
<td>yes</td>
<td>yes</td>
<td></td>
<td>[5] [63]</td>
</tr>
<tr>
<td>FAM98A</td>
<td>important</td>
<td>yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FIGN</td>
<td>important</td>
<td>yes</td>
<td></td>
<td></td>
<td>[5]</td>
</tr>
<tr>
<td>FUS</td>
<td>essential</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>core</td>
</tr>
<tr>
<td>HNRNPA1</td>
<td>important</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>HNRNPA1L2</td>
<td>n. d.</td>
<td>yes</td>
<td></td>
<td></td>
<td>[5]</td>
</tr>
<tr>
<td>HNRNPF</td>
<td>n. d.</td>
<td>no</td>
<td></td>
<td></td>
<td>[5]</td>
</tr>
<tr>
<td>HNRNPH1</td>
<td>n. d.</td>
<td>yes</td>
<td></td>
<td></td>
<td>[5]</td>
</tr>
<tr>
<td>HNRNPH3</td>
<td>essential</td>
<td>yes</td>
<td></td>
<td></td>
<td>[5]</td>
</tr>
<tr>
<td>HNRNPK</td>
<td>essential</td>
<td>no</td>
<td></td>
<td></td>
<td>[5]</td>
</tr>
<tr>
<td>HNRNPR</td>
<td>important</td>
<td>yes</td>
<td></td>
<td></td>
<td>[5]</td>
</tr>
<tr>
<td>HNRNPUL1</td>
<td>important</td>
<td>yes</td>
<td></td>
<td></td>
<td>[5]</td>
</tr>
<tr>
<td>MEX3A</td>
<td>n. d.</td>
<td>no</td>
<td></td>
<td></td>
<td>[5]</td>
</tr>
<tr>
<td>NONO</td>
<td>essential</td>
<td>yes</td>
<td></td>
<td>core</td>
<td>[15]</td>
</tr>
<tr>
<td>PCED1A</td>
<td>important</td>
<td>yes</td>
<td></td>
<td></td>
<td>[5]</td>
</tr>
<tr>
<td>PSPC1</td>
<td>dispensible</td>
<td>yes</td>
<td></td>
<td>core</td>
<td>[5]</td>
</tr>
<tr>
<td>RBM3</td>
<td>dispensible</td>
<td>yes</td>
<td></td>
<td></td>
<td>[5]</td>
</tr>
<tr>
<td>RBM4B</td>
<td>dispensible</td>
<td>no</td>
<td></td>
<td></td>
<td>[5]</td>
</tr>
<tr>
<td>RBM7</td>
<td>dispensible</td>
<td>no</td>
<td></td>
<td></td>
<td>[5]</td>
</tr>
<tr>
<td>RBM12</td>
<td>important</td>
<td>yes</td>
<td></td>
<td></td>
<td>[5]</td>
</tr>
<tr>
<td>RBM14</td>
<td>essential</td>
<td>yes</td>
<td>yes</td>
<td>patch</td>
<td>[15]</td>
</tr>
<tr>
<td>RBMX</td>
<td>dispensible</td>
<td>no</td>
<td></td>
<td></td>
<td>[5]</td>
</tr>
<tr>
<td>RUNX3</td>
<td>dispensible</td>
<td>yes</td>
<td></td>
<td></td>
<td>[5]</td>
</tr>
<tr>
<td>SFPQ</td>
<td>essential</td>
<td>yes</td>
<td>yes</td>
<td>core</td>
<td></td>
</tr>
<tr>
<td>SMARCA4</td>
<td>essential</td>
<td>no</td>
<td></td>
<td>patch</td>
<td></td>
</tr>
<tr>
<td>(BRG1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SRSF10</td>
<td>important</td>
<td>no</td>
<td></td>
<td></td>
<td>[5]</td>
</tr>
<tr>
<td>SS18L1</td>
<td>n. d.</td>
<td>yes</td>
<td>yes</td>
<td>[67] [5]</td>
<td></td>
</tr>
<tr>
<td>TAF15</td>
<td>important</td>
<td>yes</td>
<td>yes</td>
<td>[63] [5]</td>
<td></td>
</tr>
<tr>
<td>TDP43</td>
<td>n. d.</td>
<td>yes</td>
<td>yes</td>
<td>shell</td>
<td>[62]</td>
</tr>
<tr>
<td>UBAP2L</td>
<td>dispensible</td>
<td>yes</td>
<td></td>
<td></td>
<td>[5]</td>
</tr>
<tr>
<td>ZC3H6</td>
<td>dispensible</td>
<td>yes</td>
<td></td>
<td></td>
<td>[5]</td>
</tr>
</tbody>
</table>

Paraspeckle RNAs

<table>
<thead>
<tr>
<th>gene name</th>
<th>importance</th>
<th>N/A</th>
<th>5’+3’ shell, middle core</th>
<th>reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>NEAT1</td>
<td>essential</td>
<td>N/A</td>
<td></td>
<td>[2, 4, 6, 7]</td>
</tr>
<tr>
<td>IR-containing</td>
<td>dispensible</td>
<td>N/A</td>
<td></td>
<td>[35]</td>
</tr>
<tr>
<td>RNAs</td>
<td>AG rich RNAs</td>
<td>dispensible</td>
<td>N/A</td>
<td>shell</td>
</tr>
<tr>
<td>-----------------</td>
<td>--------------</td>
<td>-------------</td>
<td>-------</td>
<td>-------</td>
</tr>
</tbody>
</table>

n.d. – not determined  
N/A – not applicable  
1 A type of low complexity domain rich in polar and small amino acids (Gly, Ala, Ser, Pro, Asn, Gln, Tyr) implicated in forming fibrillar higher-order aggregates.  
2 Amyotrophic lateral sclerosis, also known as Motor Neuron Disease  
3 Partition of components of molecular mixtures into distinct demixed phases, e.g. oil and water. In the cell, many membraneless organelle display liquid behaviour suggesting that they are demix liquids.  
4 The paraspeckle zones relate to the super-resolution imaging of paraspeckles as performed in [13]
References

11. Li, R. et al. (2017) Functional dissection of NEAT1 using genome editing reveals substantial localization of the NEAT1_1 isoform outside paraspeckles. RNA 23 (6), 872-881.


Glossary

Paraspeckle: micron-sized ribonucleoprotein granule found in the mammalian cell nucleus

Long noncoding RNA: An RNA transcript over 200 nucleotides in lengths which has no coding potential. New functional IncRNAs with roles in gene regulation and/or structure (architectural IncRNA) are constantly being discovered.

Membraneless organelle: a functional subcellular body which is not bound by a lipid membrane. Without a membrane boundary, the organelle must be held together by interactions between its component molecules.

Liquid-liquid phase separation: partition of components of molecular mixtures into distinct demixed phases, e.g. oil and water. In the cell, many membraneless organelle display liquid behaviour suggesting that they are demix liquids.

Core-shell structure: A three-dimensional object with regions of varying composition. The core forms the centre of the object which is surrounded by the shell (e.g. watermelon).

Protein low complexity domains: Regions of a protein containing a small representation of amino acid types, sometimes as polymers (e.g. polyproline, polyalanine), as tandem-repeats (e.g. Tyr-Xaa-Xaa-Gln repeats) or with identifiable functional motifs (e.g. Arg-Gly-Gly motifs for nucleic acid binding).

Prion-like domains: A type of low complexity domain rich in polar and small amino acids (Gly, Ala, Ser, Pro, Asn, Gln, Tyr) implicated in forming fibrillar higher-order aggregates.
Outstanding questions

• What are the elements in the NEAT1_2 RNA that the paraspeckle proteins recognise and bind? Given relatively low sequence homology for NEAT1_2 in mammals, is it possible these protein-binding elements are set by secondary or tertiary structure of NEAT1_2? How do mutations in NEAT1_2 in cancer cells impact protein binding and paraspeckle formation?

• NONO and SFPQ function in miRNA processing is enhanced by their organisation into paraspeckles – will this principle hold true for other paraspeckle protein activities that could act in trans? Many paraspeckle proteins have roles in alternative splicing, could such activities also be enhanced by paraspeckle organisation? It is interesting to speculate as paraspeckles are found adjacent to nuclear speckles, bodies rich in splicing factors.

• What mechanism regulates individual paraspeckle length? Which factors allow the single paraspeckle units to form a longer chain?

• What is the interrelationship between NEAT1_1 and NEAT1_2? Is it possible that one role for NEAT1_2/paraspeckle production is to regulate through negative feedback, the otherwise independent NEAT1_1 molecule that may have powerful oncogenic activity?

• Paraspeckles generally form close to the NEAT1 gene and then can move elsewhere in the nucleus. But why do even distally located, mature paraspeckles depend on active transcription of NEAT1 for their integrity? How does the signal that transcription has ceased get transmitted to such paraspeckles?

• Are there additional proteins targeted to paraspeckles that are yet to be discovered?

• What are the key signatures in prion like domains that target them to paraspeckles, are there other nuclear proteins with such domains that do not get targeted there? How distinct is one prion like domain from another?
Trends box

- Paraspeckles have become an important model system for studying long noncoding RNA-protein interactions in the context of gene regulation
- Nascent 23,000 nucleotide NEAT1 long noncoding RNA transcripts act as a seed to recruit nuclear RNA binding proteins and build a paraspeckle
- Protein domains that mediate Liquid liquid phase separation are essential for many aspects of paraspeckle formation, including glueing together individual ribonucleoprotein bundles into a mature paraspeckle
- Paraspeckle formation is dynamic and triggered by many different cell stress scenarios including infection and transformation
Neat1_5' + 3'/Neat1_m

500 nm

Shell
- Neat1_5'
- Neat1_3'
- Tardbp

Core
- Neat1_m
- Sfpq
- Nono
- Pspc1
- Fus

Patch
- Rbm14
- Brg1
Steady State (paraspeckle)

\text{NEAT1}_2 (paraspeckle)

eg. circadian cycling

\text{NEAT1}_2 (paraspeckle)

eg. normal

\text{NEAT1}_2 (paraspeckle)

eg. stress

\text{paraspeckle protein} \quad \text{paraspeckle-targeted RNA} \quad \text{pri-miRNA}