


Review

Short Tandem Repeat-Enriched Architectural RNAs in Nuclear Bodies: Functions and Associated Diseases

Kensuke Ninomiya and Tetsuro Hirose * 

Institute for Genetic Medicine, Hokkaido University, Sapporo 060-0815, Japan; k-ninomiya@igm.hokudai.ac.jp

* Correspondence: hirose@igm.hokudai.ac.jp; Tel.: +81-11-706-5071; Fax: +81-11-706-7540

Received: 10 January 2020; Accepted: 19 February 2020; Published: 20 February 2020



Abstract: Nuclear bodies are membraneless, phase-separated compartments that concentrate specific proteins and RNAs in the nucleus. They are believed to serve as sites for the modification, sequestration, and storage of specific factors, and to act as organizational hubs of chromatin structure to control gene expression and cellular function. Architectural (arc) RNA, a class of long noncoding RNA (lncRNA), plays essential roles in the formation of nuclear bodies. Herein, we focus on specific arcRNAs containing short tandem repeat-enriched sequences and introduce their biological functions and recently elucidated underlying molecular mechanism. In various neurodegenerative diseases, abnormal nuclear and cytoplasmic bodies are built on disease-causing RNAs or toxic RNAs with aberrantly expanded short tandem repeat-enriched sequences. We discuss the possible analogous functions of natural arcRNAs and toxic RNAs with short tandem repeat-enriched sequences. Finally, we describe the technical utility of short tandem repeat-enriched arcRNAs as a model for exploring the structures and functions of nuclear bodies, as well as the pathogenic mechanisms of neurodegenerative diseases.

Keywords: architectural RNA; nuclear body; short tandem repeat-enriched RNA; repeat-expansion disease; toxic RNA

1. Introduction

Nuclear bodies are membraneless compartments in the nucleus, ~0.2 to 3 μm in diameter, which usually exhibit phase-separated features and liquid-like properties. In mammalian cells, there is a wide variety of nuclear bodies, including the nucleolus, Cajal bodies, histone locus bodies (HLBs), nuclear speckles, nuclear stress bodies (nSBs), paraspeckles, the perinucleolar compartment (PNC), polycomb bodies, and promyelocytic leukemia bodies. Nuclear bodies are defined by enrichment of specific marker proteins and RNA molecules, and some are conditionally formed upon experiencing certain stresses [1–4]. A subset of nuclear bodies are built on specific scaffolding long noncoding RNAs (lncRNAs) which include paraspeckles, nSBs, and PNCs formed on nuclear paraspeckle assembly transcript 1 (NEAT1), highly-repetitive satellite III (HSATIII), and pyrimidine-rich noncoding transcript (PNCTR), respectively [5–7]. These lncRNAs with membraneless nuclear body scaffolding functions have been termed architectural RNAs (arcRNAs), since they represent a distinct subcategory among thousands of lncRNAs [8,9]. The existence of multiple types of novel RNA-dependent nuclear bodies has been confirmed in mammalian cells [10,11]. lncRNAs with similar architectural functions have also been identified in *Drosophila* and fission yeast [12–14], suggesting that arcRNAs are widely utilized as structural scaffolds to form membraneless organelles in all eukaryotic species. In the case of the arcRNAs defined above, one specific arcRNA acts as a scaffold for building its cognate membraneless organelle. However, in other membraneless organelles, such as cytoplasmic stress granules, it has recently been reported that multiple RNA species are sequestered to form a scaffold;

therefore, in this case, complex RNA–RNA and RNA–protein interactions are involved in the building of these structures [15,16].

The biological roles of nuclear bodies remain poorly understood, but it has been proposed to be classified them into three functional categories [4,17]. First, they act as a reaction crucible in which specific molecules, such as enzymes and their substrates, accumulate, and thereby catalysis is facilitated. For example, nuclear speckles are widely considered a site for phosphorylation of serine/arginine-rich splicing factors (SRSFs) [18]. HLBs concentrate FLICE-associated huge protein (FLASH), histone cleavage complex (HCC), and U7 small nuclear RNP to promote efficient histone mRNA processing in a cell cycle-dependent manner [19]. Secondly, they act as a molecular sponge in which specific proteins and RNAs are stored and protected from degradation, or sequestered from nucleoplasm, to negatively regulate their nucleoplasmic functions. Paraspeckles are believed to act as molecular sponges; upregulation of NEAT1 leads to an increase in the number and size of paraspeckles, resulting in sequestration of splicing factor proline- and glutamine-rich (SFPQ) into enlarged paraspeckles separate from the nucleoplasm. Consequently, SFPQ-mediated transcription of target genes such as *ADARB2* and *IL8* is downregulated [20]. Thirdly, they act as an organizational hub that anchors multiple chromatin loci, presumably to integrate gene expression at these loci. Super enhancer (SE) is a cluster of enhancers in which high densities of transcription coactivators accumulate to form a phase-separated condensate [21]. Thus, SE is believed to enhance the efficiency of transcription. Paraspeckles are also reported to anchor multiple transcriptionally activated promoters [22,23], raising an intriguing possibility that they can act as an organizational hub for chromatin loci, as well as a molecular sponge as described above.

2. Short Tandem Repeat-Enriched ArcRNAs

To implement these functions, especially the reaction crucible and molecular sponge functions, a subset of arcRNAs possess repeat sequences to accumulate multiple copies of specific proteins and/or RNAs. For example, the cytoplasmic noncoding RNA activated by DNA damage (NORAD) lncRNA sequesters PUMILIO via at least 17 binding sites with an eight nucleotide (nt) unit (UGUAUAUA) to repress its function, which is essential for maintenance of genomic stability and mitochondrial function [24–26]. Xist lncRNA, which plays a critical role in forming the Xist cloud during X chromosome inactivation, contains several functional repeats (e.g., an A repeat unit with 25 nt and a C repeat unit with ~120 nt) [27,28]. Two 5′ SnoRNA capped and 3′ polyadenylated (SPA) lncRNAs sequester TDP43, RNA binding Fox-1 homolog 2 (RBFOX2), and heterogeneous nuclear ribonucleoprotein M (HNRNPM) through the repetitive UG-rich sequence to regulate splicing [29,30]. The nuclear lncRNA functional intergenic repeating RNA element (Firre) interacts with nuclear matrix protein heterogeneous nuclear ribonucleoprotein U (HNRNPU) via 156 nt repeats (16 in mouse and eight in human) to control nuclear architecture [31]. The *Drosophila* arcRNA heat-shock RNA-omega (Hsr omega) contains tandem repeats of 280 nt in a stretch of ~10 kb that contribute to the recruitment of various RNA-binding proteins to omega speckles (thermal stress-induced nuclear bodies) [12,14]. The yeast arcRNA meiRNA forms nuclear foci called Mei2 dots at its own genomic locus and contains at least 25 UNAAAC repeats that sequester the RNA-binding protein (RBP) Mmi1 to inhibit its function and stimulate the progression of meiosis [13].

Since most canonical RNA-binding domains of RBPs recognize five to six nt motifs for specific binding [32,33], arcRNAs possessing multiple tandem repeats of short sequence stretches, representing a small subcategory of arcRNAs, are more advantageous for efficiently sequestering a large number of specific RBPs with high affinity to an RNA molecule with multiple repeat sequences (Table 1). YAP et al. (2018) screened RNA sequencing (RNA-seq) reads consisting of short-tandem repeats that had been excluded from conventional transcriptome analyses because they were non-uniquely mapped. By collating with the RNA-binding motif database, functional short-tandem repeat-enriched RNAs were predicted, eventually yielding five RNAs that are physiologically expressed in human cells as candidate functional short-tandem repeat-enriched RNAs [6]. Among them, pyrimidine-rich

noncoding transcript (PNCTR) RNA contains multivalent CUCUCU-hexamers that bind to pyrimidine tract-binding protein 1 (PTBP1). PNCTR acts as a functional arcRNA by serving as an essential scaffold for PNC and sequestering PTBP1 within PNCs separate from the nucleoplasm to inhibit its splicing regulatory function and promote cell survival.

Table 1. List of short tandem repeat-enriched RNAs.

ArcRNA	Enriched Repeat	Nuclear Body	Binding Proteins	Functions and Roles	References
PNCTR	CUCUCU	PNC	PTBP1	Control of pre-mRNA splicing through sequestration of PTBP1	[6]
HSATIII	GGAAU	nSB-S and/or nSB-M	More than 100 proteins including SRSF1, SRSF9, CLK1, TDP43, FUS, SFPQ, CREBBP, BRD4, HSF1	Induced upon thermal stress. Control of pre-mRNA splicing through	[5,34–38]
		nSB-S nSB-M	SAFB, SLTM, NCOA5 HNRNPM, HNRNPA1, HNRNPH1	re-phosphorylation of SRSFs during recovery after stress removal	
SCA31 repeat	UGGAA	SCA31 foci	SRSF1, SRSF9, TDP43, HNRNPA2/B1, FUS, SFPQ	Expanded in SCA31 patients. Unusual production of toxic pentapeptides	[39,40]
SPA	UG-rich sequence	PWS region body	TDP43, RBFOX, HNRNPM	Deleted in PWS patients. Sequestration of RBPs to regulate splicing	[29,30]
meiRNA	UNAAAC	Mei2 dot	Mmi1	Promoting the progression of meiosis by repressing Mmi1	[13]

3. Nuclear Stress Bodies Formed by HSATIII ArcRNAs

HSATIII is an additional example of a short tandem repeat-enriched arcRNA that forms nuclear stress bodies (nSBs), consisting mainly of highly-repetitive (GGAAU)_n sequences [41]. The primate-specific pericentromeric satellite III regions are transcriptionally silent under normal conditions, but are transcribed to synthesize HSATIII arcRNAs under thermal stress conditions [5,42,43]. HSATIII arcRNAs remain stable in nuclei, but form membraneless nuclear stress bodies (nSBs) upon recruitment of specific RNA-binding proteins such as Scaffold attachment factor B (SAFB), specific sets of SRSFs, transcription factors HSF1 and CREBBP, and bromodomain protein BRD4 [5,36,37,44]. In our recent study, comprehensive proteomics analysis of HSATIII ribonucleoprotein complexes, which are most likely nSB complexes, using the chromatin isolation by RNA purification-mass spectrometry (ChIRP-MS) technique identified 141 previously unreported proteins, many of which were nuclear RNA-binding proteins involved in pre-mRNA splicing and processing [34]. In addition, cell division cycle (CDC)-like kinase 1 (CLK1), a nuclear protein kinase that phosphorylates SRSFs, is recruited to nSBs specifically during the recovery phase after thermal stress (Figure 1). During thermal stress, SRSF1 and SRSF9 are dephosphorylated and recruited to nSBs. After thermal stress, they are rapidly rephosphorylated by CLK1 kinase, and consequently, which enables rapid changes in pre-mRNA splicing [34]. Thus, nSBs are dynamic nuclear bodies, the composition of which changes in response to temperature shifts that serve as a reaction crucible for efficient SRSF phosphorylation, which is required for temperature-dependent rapid control of pre-mRNA splicing (Figure 1).

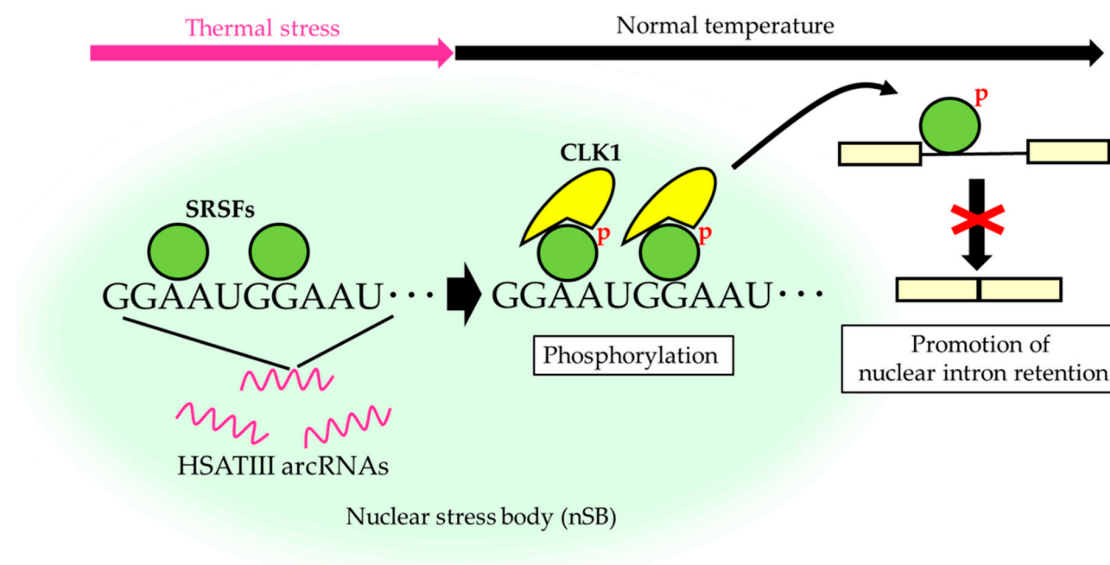


Figure 1. The reaction crucible model of nuclear stress bodies (nSBs) for serine/arginine-rich splicing factors (SRSF) phosphorylation. SRSFs are globally dephosphorylated upon thermal stress exposure. Specific SRSFs, including SRSF1, SRSF7, and SRSF9, are concentrated within nSBs through association with highly-repetitive satellite III (HSATIII) arcRNAs. Cell division cycle (CDC)-like kinase 1 (CLK1) is specifically recruited to nSBs during recovery after thermal stress cessation to rapidly rephosphorylate nSB-localized SRSFs, which represses splicing of hundreds of specific introns.

Among nSB-localized SRSFs and SR-related proteins, SRSF1 and SRSF9 are particularly highly enriched in nSBs, and share the common GGARG binding motif (R is A or G) [32], which is similar to the repeats in HSATIII arcRNAs (Figure 1). This strongly suggests that SRSF1 and SRSF9 bind directly to the enriched GGAAU repeat sequence in HSATIII, resulting in a high local concentration in nSBs.

HSATIII arcRNAs also serve as a scaffold for a minor subclass of nSBs (named nSB-M) including HNRNPM, which is distinct from the major population of HSATIII-dependent bodies characterized by SAFB enrichment (named nSB-S) [35]. This suggests that nSBs are heterologous nuclear bodies with distinct structures and functions. Intriguingly, nSB-S and nSB-M are formed on common HSATIII arcRNAs with different sets of RNA-binding proteins. The molecular mechanism of the formation of the two distinct nSBs with HSATIII arcRNAs remains to be investigated.

4. Comparison of nSBs and Spinocerebellar Ataxia Type 31 (SCA31) Foci Formed by Analogous Short Tandem Repeat-Enriched ArcRNAs

Spinocerebellar ataxia type 31 (SCA31) is a neurological disease caused by insertion of a 2.5 to 3.8 kb repeat that includes the long TGGAA stretch present in a common intronic region of BEAN1 (brain expressed associated with NEDD4-1) and thymidine kinase 2 (TK2) genes [39,40,45]. BEAN1 transcripts containing expanded UGGAA repeats are mainly retained in nuclei where they form RNA foci in Purkinje cells of the cerebellum of SCA31 patients. These RNA foci sequester a number of RNA-binding proteins including TDP43 (Table 1). This is possible because the UGGAA tandem repeat contains multiple copies of GAAUG, which is identical to the high-affinity binding motif of TDP43 [32]. A subpopulation of SCA31 repeat RNA is believed to be transported to the cytoplasm and translated into a toxic pentapeptide repeat (PPR), which is a potential cause of the disorder [39]. Moreover, phenotypic analysis using *Drosophila* models of SCA31 expressing long UGGAA repeat RNAs revealed toxic effects that are suppressed by the RNA chaperone-like function of TDP43, FUS, and HNRNPA2/B1. Additionally, nontoxic short UGGAA repeat RNA suppresses the toxic effects of the ALS-causing TDP43 mutation. This raises an intriguing possibility that SCA31 repeat RNAs and ALS-causing factors such as TDP43, FUS, and HNRNPA2B1 could interact to neutralize the respective toxic effects, suggesting that the disruption induced by aberrant expansion of the SCA31

repeat and the TDP43 mutation could be responsible for the neurodegenerative disease phenotypes [39]. Interestingly, the SCA31 UGGAA repeat is essentially equivalent to the GGAAU repeat in HSATIII arcRNAs. It is possible that SCA31 foci control pre-mRNA splicing, which is analogous to the recently discovered function of HSATIII in nSBs described above. Indeed, the HSATIII-interacting proteins identified by ChIRP-MS partially overlap with those of the UGGAA repeat RNA identified by in vitro binding experiments (Table 1) [34,35,39,40], suggesting that at least the overlapping proteins, such as TDP43 and FUS, interact directly or indirectly with GGAAU repeats rather than with non-repeat regions. It should be noted that HSATIII complexes are nSBs that are specifically formed under thermal stress conditions, but SCA31 RNA foci are formed under normal temperature conditions. Thus, RNAs possessing analogous repeat sequences potentially form distinct RNP complexes with partially overlapping protein components, and play distinct or partially overlapping roles depending on the context, environment, and tissue type (Figure 2).

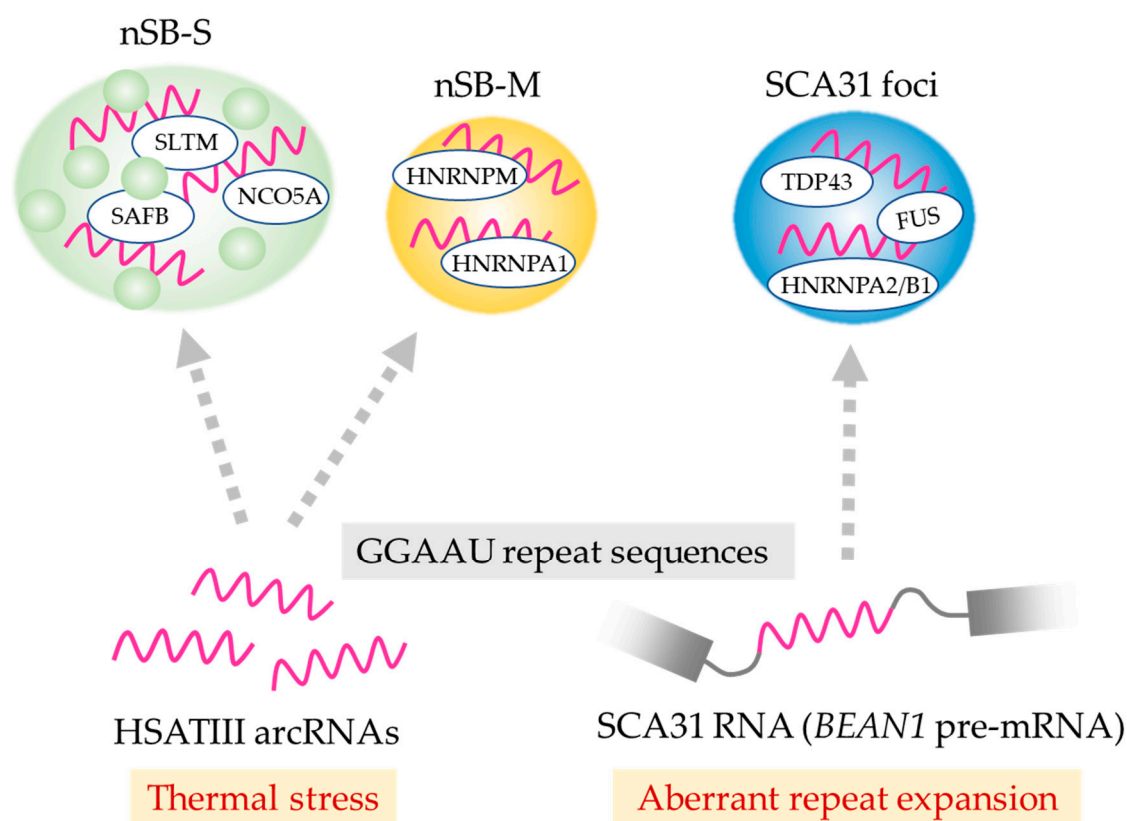


Figure 2. Distinct nuclear bodies assemble on a common repeat sequence of arcRNAs. nSB-S and nSB-M are mutually exclusive subnuclear bodies formed on HSATIII arcRNAs upon thermal stress exposure. SCA31 foci are formed on the expanded UGGAA repeat region of BEAN1 transcripts in the nuclei of Purkinje cells in the brains of SCA31 patients. The identified resident proteins are indicated in each nuclear body. GGAAU repeat-enriched regions are colored magenta.

5. Future Perspectives

In addition to SCA31 RNA described above, recent studies have demonstrated that several kinds of short-tandem repeat-enriched arcRNAs are also synthesized from specific gene loci containing pathogenic expansion of three to six nt repeats, which are believed to cause various neurological and neuromuscular disorders such as myotonic dystrophy 1 and 2, fragile X tremor ataxia syndrome (FXTAS), SCA8, SCA10, and SCA36 [46–48]. Hence, these repeat RNAs are termed toxic RNAs. Since the disease-associated repeat regions are, in most cases, located within introns or 5'- or 3'-UTRs of host genes, they should function as part of pre- or spliced mRNAs (or possibly excised introns) transcribed

from host genes. The proposed pathogenic mechanism involves toxic RNAs sequestering specific sets of RNA-binding proteins to form abnormal aggregates, which likely impairs cellular functions and generates toxic repeat polypeptides by AUG-dependent or repeat-associated non-AUG (RAN) translation [49,50]. However, in the case of most disease associated repeat RNAs, a comprehensive overview of their binding proteins and the precise pathogenic mechanisms has not been performed.

Functional analysis of short tandem repeat-enriched arcRNAs has provided important insight into the structures and functions of membraneless nuclear bodies formed on specific arcRNAs. Additionally, this class of arcRNAs can serve as models for structural and functional analysis of arcRNAs that are difficult to investigate directly due to technical reasons. Firstly, short tandem repeat-enriched arcRNAs are highly specific for particular nuclear bodies, hence they are the most reliable markers of nuclear bodies, whereas most associated proteins are typically located in the nucleoplasm outside nuclear bodies. In most cases, these arcRNAs are efficiently knocked down by nucleofection of a single antisense oligonucleotide (ASO) that hybridizes with multiple sites of the tandem repeat stretch, which enables the specific disruption of nuclear bodies and the ability to investigate their functions. Hybridization-based techniques for RNP precipitation, such as ChIRP, capture hybridization analysis of RNA targets (CHART), and RNA antisense purification (RAP) [51], have been developed and utilized to purify the components of RNA-based complexes. These methods are powerful tools for the comprehensive identification of factors associated with specific RNAs *in vivo*, but it is usually difficult to design appropriate antisense oligonucleotides (ASOs) that can act as probes for efficient capture of specific RNP complexes, presumably due to inefficient association of ASOs that target RNAs, which are largely masked by proteins. In the case of repeat RNAs, ASOs can more efficiently capture entire RNP complexes, presumably through multivalent hybridization with some unmasked regions of target RNAs. Furthermore, efficient and stable purification of substantial amounts of RNPs allows us to perform biochemical analysis of time course- and signal-dependent changes within RNP complexes, such as exploring changes in phosphorylation states of specific components of RNPs [34]. For the same reasons described above, this class of arcRNAs are relatively easily detected by fluorescent *in situ* hybridization (FISH). Furthermore, since databases that assort binding sequence motifs of multiple RBPs have been constructed [32,33], proteins that bind the specific repeat region can be predicted using bioinformatics. A subset of novel arcRNAs, whose binding proteins remain identifiable, also possess repeat sequences [10].

Therefore, this approach is applicable for unveiling the structures and functions of novel arcRNA-dependent cellular bodies and disease-related RNA foci. Currently, small compounds that specifically bind to disease-related repeat RNAs have been explored as candidate therapeutic drugs [52–54]. Further research should be directed to evaluate such candidate drugs in disease model cells and investigate their modes of action. Intensive investigation of short tandem repeat-enriched arcRNAs provides important insights into the pathogenic mechanisms of RNA repeat expansion-related neurological and neuromuscular diseases, and reveals some basic concepts pertinent to therapeutic development.

Funding: This research was supported by MEXT KAKENHI grants to T.H. (JP26113002, JP16H06279, JP17H03630, and JP17K19335) and N.K. (JP19K06478)], and by the Tokyo Biochemical Research Foundation (to TH).

Acknowledgments: The authors thank Shungo Adachi and Tohru Natsume for proteomic analysis of nSBs; Junichi Iwakiri, Goro Terai, and Kiyoshi Asai for bioinformatic and transcriptomic analysis using HSATIII KD cells, and members of the Hirose laboratory for valuable discussions.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Huang, S. Review: Perinucleolar structures. *J. Struct. Biol.* **2000**, *129*, 233–240. [[CrossRef](#)] [[PubMed](#)]
2. Nizami, Z.; Deryusheva, S.; Gall, J.G. The Cajal body and histone locus body. *Cold Spring Harb. Perspect. Biol.* **2010**, *2*, a000653. [[CrossRef](#)] [[PubMed](#)]

3. Biamonti, G.; Vourc'h, C. Nuclear stress bodies. *Cold Spring Harb. Perspect. Biol.* **2010**, *2*, a000695. [[CrossRef](#)] [[PubMed](#)]
4. Mao, Y.S.; Zhang, B.; Spector, D.L. Biogenesis and function of nuclear bodies. *Trends Genet.* **2011**, *27*, 295–306. [[CrossRef](#)] [[PubMed](#)]
5. Jolly, C.; Metz, A.; Govin, J.; Vigneron, M.; Turner, B.M.; Khochbin, S.; Vourc'h, C. Stress-induced transcription of satellite III repeats. *J. Cell Biol.* **2004**, *164*, 25–33. [[CrossRef](#)] [[PubMed](#)]
6. Yap, K.; Mukhina, S.; Zhang, G.; Tan, J.S.C.; Ong, H.S.; Makeyev, E.V. A Short Tandem Repeat-Enriched RNA Assembles a Nuclear Compartment to Control Alternative Splicing and Promote Cell Survival. *Mol. Cell* **2018**, *72*, 525–540.e13. [[CrossRef](#)]
7. Chujo, T.; Hirose, T. Nuclear Bodies Built on Architectural Long Noncoding RNAs: Unifying Principles of Their Construction and Function. *Mol. Cells* **2017**, *40*, 889–896. [[CrossRef](#)]
8. Hirose, T.; Mishima, Y.; Tomari, Y. Elements and machinery of non-coding RNAs: Toward their taxonomy. *EMBO Rep.* **2014**, *15*, 489–507. [[CrossRef](#)]
9. Chujo, T.; Yamazaki, T.; Hirose, T. Architectural RNAs (arcRNAs): A class of long noncoding RNAs that function as the scaffold of nuclear bodies. *Biochim. Biophys. Acta* **2016**, *1859*, 139–146. [[CrossRef](#)]
10. Chujo, T.; Yamazaki, T.; Kawaguchi, T.; Kurosaka, S.; Takumi, T.; Nakagawa, S.; Hirose, T. Unusual semi-extractability as a hallmark of nuclear body-associated architectural noncoding RNAs. *EMBO J.* **2017**, *36*, 1447–1462. [[CrossRef](#)]
11. Mannen, T.; Yamashita, S.; Tomita, K.; Goshima, N.; Hirose, T. The Sam68 nuclear body is composed of two RNase-sensitive substructures joined by the adaptor HNRNPL. *J. Cell Biol.* **2016**, *214*, 45–59. [[CrossRef](#)] [[PubMed](#)]
12. Prasanth, K.V.; Rajendra, T.K.; Lal, A.K.; Lakhota, S.C. Omega speckles - a novel class of nuclear speckles containing hnRNPs associated with noncoding hsr-omega RNA in Drosophila. *J. Cell Sci.* **2000**, *113*, 3485–3497.
13. Shichino, Y.; Yamashita, A.; Yamamoto, M. Meiotic long non-coding meiRNA accumulates as a dot at its genetic locus facilitated by Mmi1 and plays as a decoy to lure Mmi1. *Open Biol.* **2014**, *4*, 140022. [[CrossRef](#)] [[PubMed](#)]
14. Singh, A.K.; Lakhota, S.C. Dynamics of hnRNPs and omega speckles in normal and heat shocked live cell nuclei of Drosophila melanogaster. *Chromosoma* **2015**, *124*, 367–383. [[CrossRef](#)]
15. Van Treeck, B.; Parker, R. Emerging Roles for Intermolecular RNA-RNA Interactions in RNP Assemblies. *Cell* **2018**, *174*, 791–802. [[CrossRef](#)] [[PubMed](#)]
16. Krause, H.M. New and Prospective Roles for lncRNAs in Organelle Formation and Function. *Trends Genet.* **2018**, *34*, 736–745. [[CrossRef](#)]
17. Shin, Y.; Brangwynne, C.P. Liquid phase condensation in cell physiology and disease. *Science* **2017**, *357*. [[CrossRef](#)] [[PubMed](#)]
18. Colwill, K.; Pawson, T.; Andrews, B.; Prasad, J.; Manley, J.L.; Bell, J.C.; Duncan, P.I. The Clk/Sty protein kinase phosphorylates SR splicing factors and regulates their intranuclear distribution. *EMBO J.* **1996**, *15*, 265–275. [[CrossRef](#)]
19. Tatomer, D.C.; Terzo, E.; Curry, K.P.; Salzler, H.; Sabath, I.; Zapotoczny, G.; McKay, D.J.; Dominski, Z.; Marzluff, W.F.; Duronio, R.J. Concentrating pre-mRNA processing factors in the histone locus body facilitates efficient histone mRNA biogenesis. *J. Cell Biol.* **2016**, *213*, 557–570. [[CrossRef](#)]
20. Hirose, T.; Virnicchi, G.; Tanigawa, A.; Naganuma, T.; Li, R.; Kimura, H.; Yokoi, T.; Nakagawa, S.; Benard, M.; Fox, A.H.; et al. NEAT1 long noncoding RNA regulates transcription via protein sequestration within subnuclear bodies. *Mol. Biol. Cell* **2014**, *25*, 169–183. [[CrossRef](#)]
21. Sabari, B.R.; Dall'Agnesse, A.; Boija, A.; Klein, I.A.; Coffey, E.L.; Shrinivas, K.; Abraham, B.J.; Hannett, N.M.; Zamudio, A.V.; Manteiga, J.C.; et al. Coactivator condensation at super-enhancers links phase separation and gene control. *Science* **2018**, *361*. [[CrossRef](#)] [[PubMed](#)]
22. Sridhar, B.; Rivas-Astroza, M.; Nguyen, T.C.; Chen, W.; Yan, Z.; Cao, X.; Hebert, L.; Zhong, S. Systematic Mapping of RNA-Chromatin Interactions In Vivo. *Curr. Biol.* **2017**, *27*, 602–609. [[CrossRef](#)] [[PubMed](#)]
23. West, J.A.; Davis, C.P.; Sunwoo, H.; Simon, M.D.; Sadreyev, R.I.; Wang, P.I.; Tolstorukov, M.Y.; Kingston, R.E. The long noncoding RNAs NEAT1 and MALAT1 bind active chromatin sites. *Mol. Cell* **2014**, *55*, 791–802. [[CrossRef](#)] [[PubMed](#)]

24. Lee, S.; Kopp, F.; Chang, T.C.; Sataluri, A.; Chen, B.; Sivakumar, S.; Yu, H.; Xie, Y.; Mendell, J.T. Noncoding RNA NORAD Regulates Genomic Stability by Sequestering PUMILIO Proteins. *Cell* **2016**, *164*, 69–80. [[CrossRef](#)] [[PubMed](#)]
25. Tichon, A.; Gil, N.; Lubelsky, Y.; Havkin Solomon, T.; Lemze, D.; Itzkovitz, S.; Stern-Ginossar, N.; Ulitsky, I. A conserved abundant cytoplasmic long noncoding RNA modulates repression by Pumilio proteins in human cells. *Nat. Commun.* **2016**, *7*, 12209. [[CrossRef](#)]
26. Kopp, F.; Elguindy, M.M.; Yalvac, M.E.; Zhang, H.; Chen, B.; Gillett, F.A.; Lee, S.; Sivakumar, S.; Yu, H.; Xie, Y.; et al. PUMILIO hyperactivity drives premature aging of Norad-deficient mice. *eLife* **2019**, *8*. [[CrossRef](#)]
27. Zhao, J.; Sun, B.K.; Erwin, J.A.; Song, J.J.; Lee, J.T. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science* **2008**, *322*, 750–756. [[CrossRef](#)]
28. Liu, F.; Somarowthu, S.; Pyle, A.M. Visualizing the secondary and tertiary architectural domains of lncRNA RepA. *Nat. Chem. Biol.* **2017**, *13*, 282–289. [[CrossRef](#)]
29. Yin, Q.F.; Yang, L.; Zhang, Y.; Xiang, J.F.; Wu, Y.W.; Carmichael, G.G.; Chen, L.L. Long noncoding RNAs with snoRNA ends. *Mol. Cell* **2012**, *48*, 219–230. [[CrossRef](#)]
30. Wu, H.; Yin, Q.F.; Luo, Z.; Yao, R.W.; Zheng, C.C.; Zhang, J.; Xiang, J.F.; Yang, L.; Chen, L.L. Unusual Processing Generates SPA lncRNAs that Sequester Multiple RNA Binding Proteins. *Mol. Cell* **2016**, *64*, 534–548. [[CrossRef](#)]
31. Hacisuleyman, E.; Goff, L.A.; Trapnell, C.; Williams, A.; Henao-Mejia, J.; Sun, L.; McClanahan, P.; Hendrickson, D.G.; Sauvageau, M.; Kelley, D.R.; et al. Topological organization of multichromosomal regions by the long intergenic noncoding RNA Firre. *Nat. Struct. Mol. Biol.* **2014**, *21*, 198–206. [[CrossRef](#)] [[PubMed](#)]
32. Ray, D.; Kazan, H.; Cook, K.B.; Weirauch, M.T.; Najafabadi, H.S.; Li, X.; Gueroussov, S.; Albu, M.; Zheng, H.; Yang, A.; et al. A compendium of RNA-binding motifs for decoding gene regulation. *Nature* **2013**, *499*, 172–177. [[CrossRef](#)] [[PubMed](#)]
33. Dominguez, D.; Freese, P.; Alexis, M.S.; Su, A.; Hochman, M.; Palden, T.; Bazile, C.; Lambert, N.J.; Van Nostrand, E.L.; Pratt, G.A.; et al. Sequence, Structure, and Context Preferences of Human RNA Binding Proteins. *Mol. Cell* **2018**, *70*, 854–867. [[CrossRef](#)] [[PubMed](#)]
34. Ninomiya, K.; Adachi, S.; Natsume, T.; Iwakiri, J.; Terai, G.; Asai, K.; Hirose, T. lncRNA-dependent nuclear stress bodies promote intron retention through SR protein phosphorylation. *EMBO J.* **2019**, e102729. [[CrossRef](#)]
35. Aly, M.K.; Ninomiya, K.; Adachi, S.; Natsume, T.; Hirose, T. 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.* **2019**, *516*, 419–423. [[CrossRef](#)] [[PubMed](#)]
36. Hussong, M.; Kaehler, C.; Kerick, M.; Grimm, C.; Franz, A.; Timmermann, B.; Welzel, F.; Isensee, J.; Hucho, T.; Krobtsch, S.; et al. The bromodomain protein BRD4 regulates splicing during heat shock. *Nucleic Acids Res.* **2017**, *45*, 382–394. [[CrossRef](#)]
37. Metz, A.; Soret, J.; Vourc'h, C.; Tazi, J.; Jolly, C. A key role for stress-induced satellite III transcripts in the relocalization of splicing factors into nuclear stress granules. *J. Cell Sci.* **2004**, *117*, 4551–4558. [[CrossRef](#)]
38. Jolly, C.; Morimoto, R.; Robert-Nicoud, M.; Vourc'h, C. HSF1 transcription factor concentrates in nuclear foci during heat shock: Relationship with transcription sites. *J. Cell Sci.* **1997**, *110*, 2935–2941.
39. Ishiguro, T.; Sato, N.; Ueyama, M.; Fujikake, N.; Sellier, C.; Kanegami, A.; Tokuda, E.; Zamiri, B.; Gall-Duncan, T.; Mirceta, M.; et al. Regulatory Role of RNA Chaperone TDP-43 for RNA Misfolding and Repeat-Associated Translation in SCA31. *Neuron* **2017**, *94*, 108–124. [[CrossRef](#)]
40. Sato, N.; Amino, T.; Kobayashi, K.; Asakawa, S.; Ishiguro, T.; Tsunemi, T.; Takahashi, M.; Matsuura, T.; Flanigan, K.M.; Iwasaki, S.; et al. Spinocerebellar ataxia type 31 is associated with “inserted” penta-nucleotide repeats containing (TGGA)n. *Am. J. Hum. Genet.* **2009**, *85*, 544–557. [[CrossRef](#)]
41. Valgardsdottir, R.; Chiodi, I.; Giordano, M.; Cobianchi, F.; Riva, S.; Biamonti, G. Structural and functional characterization of noncoding repetitive RNAs transcribed in stressed human cells. *Mol. Biol. Cell* **2005**, *16*, 2597–2604. [[CrossRef](#)] [[PubMed](#)]
42. Denegri, M.; Moralli, D.; Rocchi, M.; Biggiogera, M.; Raimondi, E.; Cobianchi, F.; De Carli, L.; Riva, S.; Biamonti, G. Human chromosomes 9, 12, and 15 contain the nucleation sites of stress-induced nuclear bodies. *Mol. Biol. Cell* **2002**, *13*, 2069–2079. [[CrossRef](#)] [[PubMed](#)]

43. Valgardsdottir, R.; Chiodi, I.; Giordano, M.; Rossi, A.; Bazzini, S.; Ghigna, C.; Riva, S.; Biamonti, G. Transcription of Satellite III non-coding RNAs is a general stress response in human cells. *Nucleic Acids Res.* **2008**, *36*, 423–434. [[CrossRef](#)] [[PubMed](#)]
44. Kawaguchi, T.; Tanigawa, A.; Naganuma, T.; Ohkawa, Y.; Souquere, S.; Pierron, G.; Hirose, T. SWI/SNF chromatin-remodeling complexes function in noncoding RNA-dependent assembly of nuclear bodies. *Proc. Natl. Acad. Sci.* **2015**, *112*, 4304–4309. [[CrossRef](#)] [[PubMed](#)]
45. Niimi, Y.; Takahashi, M.; Sugawara, E.; Umeda, S.; Obayashi, M.; Sato, N.; Ishiguro, T.; Higashi, M.; Eishi, Y.; Mizusawa, H.; et al. Abnormal RNA structures (RNA foci) containing a penta-nucleotide repeat (UGGAA)_n in the Purkinje cell nucleus is associated with spinocerebellar ataxia type 31 pathogenesis. *Neuropathology* **2013**, *33*, 600–611. [[CrossRef](#)]
46. Jain, A.; Vale, R.D. RNA phase transitions in repeat expansion disorders. *Nature* **2017**, *546*, 243–247. [[CrossRef](#)]
47. Swinnen, B.; Robberecht, W.; Van Den Bosch, L. RNA toxicity in non-coding repeat expansion disorders. *EMBO J.* **2019**, e101112. [[CrossRef](#)]
48. Todd, P.K.; Paulson, H.L. RNA-mediated neurodegeneration in repeat expansion disorders. *Ann. Neurol.* **2010**, *67*, 291–300. [[CrossRef](#)]
49. Zu, T.; Gibbens, B.; Doty, N.S.; Gomes-Pereira, M.; Huguet, A.; Stone, M.D.; Margolis, J.; Peterson, M.; Markowski, T.W.; Ingram, M.A.; et al. Non-ATG-initiated translation directed by microsatellite expansions. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 260–265. [[CrossRef](#)]
50. Mori, K.; Weng, S.M.; Arzberger, T.; May, S.; Rentzsch, K.; Kremmer, E.; Schmid, B.; Kretzschmar, H.A.; Cruts, M.; Van Broeckhoven, C.; et al. The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTL/ALS. *Science* **2013**, *339*, 1335–1338. [[CrossRef](#)]
51. Cao, M.; Zhao, J.; Hu, G. Genome-wide methods for investigating long noncoding RNAs. *Biomed. Pharm.* **2019**, *111*, 395–401. [[CrossRef](#)] [[PubMed](#)]
52. Rzuczek, S.G.; Park, H.; Disney, M.D. A toxic RNA catalyzes the in cellulo synthesis of its own inhibitor. *Angew. Chem. Int. Ed. Engl.* **2014**, *53*, 10956–10959. [[CrossRef](#)] [[PubMed](#)]
53. Rzuczek, S.G.; Colgan, L.A.; Nakai, Y.; Cameron, M.D.; Furling, D.; Yasuda, R.; Disney, M.D. Precise small-molecule recognition of a toxic CUG RNA repeat expansion. *Nat. Chem. Biol.* **2017**, *13*, 188–193. [[CrossRef](#)] [[PubMed](#)]
54. Nguyen, L.; Lee, J.; Wong, C.H.; Zimmerman, S.C. Small molecules that target the toxic RNA in myotonic dystrophy type 2. *ChemMedChem* **2014**, *9*, 2455–2462. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).