Review

Justin M. O’Sullivan*, Dave A. Pai#, Andrew G. Cridge, David R. Engelke and Austen R.D. Ganley

The nucleolus: a raft adrift in the nuclear sea or the keystone in nuclear structure?

Abstract: The nucleolus is a prominent nuclear structure that is the site of ribosomal RNA (rRNA) transcription, and hence ribosome biogenesis. Cellular demand for ribosomes, and hence rRNA, is tightly linked to cell growth and the rRNA makes up the majority of all the RNA within a cell. To fulfill the cellular demand for rRNA, the ribosomal RNA (rDNA) genes are amplified to high copy number and transcribed at very high rates. As such, understanding the rDNA has profound consequences for our comprehension of genome and transcriptional organization in cells. In this review, we address the question of whether the nucleolus is a raft adrift the sea of nuclear DNA, or actively contributes to genome organization. We present evidence supporting the idea that the nucleolus, and the rDNA contained therein, play more roles in the biology of the cell than simply ribosome biogenesis. We propose that the nucleolus and the rDNA are central factors in the spatial organization of the genome, and that rapid alterations in nucleolar structure in response to changing conditions manifest themselves in altered genomic structures that have functional consequences. Finally, we discuss some predictions that result from the nucleolus having a central role in nuclear organization.

Keywords: genome architecture; nucleolus; rDNA.

Introduction

Nucleoli are the largest non-chromosomal structures present within the eukaryotic nucleus. In yeast, the single nucleolus occupies approximately a quarter of the total nuclear volume in a position that is distal to the spindle pole body and in close contact with the nuclear envelope (1–3). In metazoans there can be multiple nucleoli, formed around distinct chromosomal loci, that differ from yeast in details of morphology but retain the dense staining caused by the prodigious production of ribosomes [e.g., reviewed in (4, 5)]. Nucleoli are organized around the core ribosomal RNA (rRNA) gene regions, referred to as nucleolus organizer regions (NORs) (6). NORs can, in some instances, form secondary constrictions on metaphase chromosomes during mitosis.

In eukaryotes, NORs usually consist of rRNA genes that are organized into tandem repeat arrays, collectively known as the rDNA (Figure 1). rDNA gene copy number can vary from a few copies up to tens of thousands of copies, depending on the species [see (7) for a comprehensive table]. For example, the well-characterized single rDNA array in Saccharomyces cerevisiae consists of around 180 copies (8), whereas in humans there are five rDNA arrays (9) that together comprise 300–400 copies per diploid genome (10). There are very few known exceptions to the tandem repeat rule: the intracellular human pathogen Pneumocystis carinii (11) and Tetrahymena (12) both appear to have just a single rDNA locus, although the latter amplifies this copy in the macronucleus (12). Nevertheless, the vast majority of eukaryotes characterized to date have the canonical rDNA organization, in which the polycistronic rRNA coding region, consisting of 18S, 5.8S, and 28S rRNA species (precise nomenclature varies somewhat between species), is interspersed with an intergenic spacer (IGS) region (13). The rDNA genes are the most highly transcribed in the genome, with rRNA accounting for approximately 80% of total RNA in a cell (14, 15). Despite this, the rDNA is a mosaic of transcribed,
typically highly, copies and completely silent copies (16). The organization of active and silent repeats within the linear rDNA array has yet to be determined. Similarly, the role of the silent copies has not been completely resolved, although they are required for efficient DNA repair in budding yeast (17).

The nucleolus is a domain of the nucleus, rather than a body delineated by a membrane or the like. Nevertheless, it has a specific structure that, in mammalian nuclei, consists of an inner fibrillar center, a dense fibrillar component outside of this, and a granular component surrounding this (18, 19), although see (20). Although this is the case in mammalian nuclei, lower eukaryotes, in particular several yeast species, only have two distinctly visible components: a fibrillar component and granules. Furthermore, the fibrillar component in many yeast species is more a collection of strands, rather than a dense body (5). In either case, it has been shown that at least the non-transcribed parts of the rDNA are concentrated in the fibrillar component (FC) (21).

The nucleolus is very protein dense [e.g., reviewed in (4)] and in humans contains at least 700 different proteins (22), while being relatively DNA sparse. The nucleolus emerges from the complex mixture of proteins that associate with the rDNA, such as upstream binding factor (UBF) (23). Creation of the spatial domain of the nucleolus may result from high concentrations of binding sites in a small volume effectively causing retention of these proteins (24) by preventing movement out of the zone, as shown for ribosome movement (25). However, rapid shuttling of proteins between the nucleolus and nucleus has been observed (18, 19), suggesting that the nucleolus is a dynamic structure.

The nucleolus is not just a site of ribosome biogenesis: it functions in a myriad of other nuclear processes, including cell cycle control [reviewed in (4)]. Several proteins are known to localize to the nucleolus in a cell cycle-specific manner, including several that are associated with human disease (26). Furthermore, nucleolar localization of viral proteins involved in viral replication, including HIV, appears to be necessary for replication (19). Additionally, nucleolar structure changes in response to both environmental conditions and the cell cycle (18, 26). Such structural alterations, as well as alterations in the numbers of rDNA repeats, would relieve or exacerbate the retention of proteins sequestered in the nucleolus as a result of changes in the spatial clustering of binding sites. Strikingly, several non-coding RNA transcripts from the rDNA IGS appear to bind and sequester proteins in the nucleolus, and are regulated by stress (27). Given its dynamic nature, and the central role it plays in responding to cellular and environmental challenges, we hypothesize that the nucleolus has a direct role in coordinating nuclear structural organization.

The nucleolus as an organizer of genome structure

The nucleolus can contribute to nuclear organization through the sequestration and release of proteins that then, directly or indirectly, affect the organization of the nucleus. However, for the remainder of this review we are going to consider the issues surrounding the possibility that the nucleolus plays a direct role in the regulation of genome structure and how this might be achieved. In this context, we refer to genome structure as the spatial organization of the genome within the nucleus, thus this form of organization focuses on the DNA, although obviously all the attendant proteins and other factors are also part of this.

There is growing evidence that the genome takes on a specific structural arrangement within the nucleus. In human cells, different chromosomes are found to occupy chromosome ‘territories’, which have different positions in different cell types (28, 29). Genes are also observed to inhabit specific locations in the nucleus (3). Gene loops, that bring linearly distant enhancers in close spatial proximity to promoters, are also thought to be important for regulation of gene expression [e.g., (30–33)]. Recently developed techniques derived from proximity-based ligation (34–36), such as genome conformation capture (GCC) (37) and Hi-C (38), have been developed to experimentally determine global genome structure. Although extremely powerful, these techniques suffer from limitations when it comes to aligning sequences from repetitive elements.
Essentially because repetitive elements cannot be positioned to a unique position, they provide potentially confusing information in proximity-based ligation assays and are typically ignored (38–40). However, the rDNA is a special case and useful information can be obtained by collapsing the rDNA sequences to a single locus (37, 41, 42).

Computational-based approaches, utilizing proximity-ligation data and biophysical characteristics, have been taken to model global genome structure [e.g., (38, 40, 43, 44)]. Interestingly, few restraints are required to impart a crude order on in silico polymer-based reconstructions of the budding yeast nucleus (43, 44). However, one restraint that is required is the positioning of the nucleolus opposite to the spindle pole body (43), suggesting the nucleolus is a significant landmark for spatial organization of the genome.

Nucleolar localization of rDNA has been shown to influence the organization of other genomic loci in the malaria parasite, Plasmodium falciparum (45). Despite this, a structured nucleolus is not essential for nuclear function in yeast, as the rDNA genes can be deleted from their chromosomal locus and replaced with plasmid-encoded copies (46). These extra-chromosomally encoded rDNA genes form multiple, tiny dispersed nucleoli (47), and the growth of these strains is compromised. However, it remains unknown whether the growth defects stem from disruption of nuclear organization, or from attenuated rRNA transcription/processing (46). Nucleolar structure is also disrupted when yeast are forced to transcribe the chromosomal rDNA repeats with RNA polymerase (RNAP) II, rather than RNAP I (48). The entire yeast rDNA array can be shifted to another location within the genome, but in this case only minor phenotypic changes are observed, despite the nucleolus changing its position in the nucleus (49). This is consistent with a limited amount of published data that show that specific rDNA:non-rDNA interactions are sequence specific and independent of the chromosomal position of the non-rDNA locus (42). Thus, more work is required to deduce the effects of changes in nucleolar position on genome structure and function.

If the nucleolus directly regulates nuclear structure then it stands to reason that interactions between the rDNA repeats and other non-nucleolar loci are central to this. This is borne out experimentally in budding yeast where a majority of DNA-DNA interactions involve the rDNA (37). Although it can be argued that this interpretation is simplistic and does not take into account the copy number of the rDNA, any interactions between rDNA and non-rDNA loci are candidates for interactions by which the nucleolus shapes genome organization. These interactions should involve rDNA loci that are directly accessible from the nucleoplasm and are not protected by being internalized within the nucleolar structure.

The division of rDNA units into highly transcribed copies and completely silenced copies may reflect a functional distinction between units buried in the nucleolus interior and those located at the nuclear-nucleolar interface, respectively (T. Kobayashi, personal communication). Although it is almost certain that a main driver for nucleolar organization is the centralization of massive biosynthesis of ribosomes, we speculate that the tandem repeat organization of eukaryotic rDNA genes also enables the conservation of contacts at the nuclear-nucleolar boundary while still maintaining dedicated transcription units within the nucleolus. Such a system would allow the flexible assignment of rDNA repeats to the different functional categories: transcription, repair, replication, and structural associations, the latter having hitherto largely gone unrecognized. Therefore, the maintenance in eukaryotes of rDNA repeats with identical sequences [notably the non-coding regions (50)], at a much greater copy number than is needed for transcription alone, may ultimately stem from the ability of this system to seamlessly replace one repeat with another, ensuring that critical functions are maintained.

Transcription and nucleolus directed organization

The rDNA is not transcriptionally homogeneous; instead, all three classes of RNA polymerase are present in the nucleolus, in at least some organisms. Aside from RNAP I transcription, RNAP II transcription appears to be widespread in eukaryote rDNA (27, 51–55). Furthermore, RNAP III-transcribed 5S rDNA genes are located within the rDNA repeat in several species, including yeast [Figure 1; (56)]. Moreover, around 30 small interspersed nuclear element (SINE) retrotransposons that derive from RNAP III-transcribed genes are found scattered throughout the human rDNA IGS (57). This opens up the question as to the effect of this transcriptional heterogeneity on the spatial organization of the nucleolus/nucleus.

Transcription-induced clustering represents a simple mechanism for spatial genome organization (58–60). Thus, polymerase class-dependent association of active or primed promoters in the rDNA may contribute to the coordination of nuclear-nucleolar structure. In support of this idea, structures consistent with RNAP I transcription factories involving rDNA repeats have been observed in metazoan cells (60). Furthermore, RNAP III forms foci within the nucleoplasm, and not the nucleoli, of human
cells, although it is possible that this is the result of SINE transcription (61). Transcription by all three eukaryotic RNA polymerases on overlapping regions of the rDNA repeat complicates this picture.

The simplest explanation for the overlapping polymerase activities within rDNA repeats is that the different RNA polymerase activities are temporally and spatially separated. This is supported by evidence that suggests a reciprocal relationship between RNAPI and II transcription in the rDNA (62, 63). Thus, the presence of rDNA repeats on the nucleus/nucleolus interface may free them up to be transcribed by RNAPII and/or III. However, in a yeast strain where rDNA repeat number is reduced to the extent that most copies are likely to be transcribed by RNAPI (64), RNAPII transcription is also high (51). This suggests that transcription by these two polymerases is not mutually exclusive.

In the case of RNAPI and RNAPIII, it is clear that the transcription units can co-exist. Not only are the 5S rRNA genes and 35S rRNA genes (transcribed by RNAPIII and I, respectively) interspersed in the linear repeats but there is also substantial evidence in the literature that 5S rRNA genes are associated with nucleoli even when located at distant sites in the linear genome (see below). Thus, the dynamics of rDNA repeat transcription is an important area for future research.

The nucleolus and RNAPIII decoded genes

The spatial organization of the 5S rDNA genes is one of the clearest examples of the nucleolus affecting nuclear organization. Whereas in S. cerevisiae the 5S rDNA are located with the large rDNA repeats, in most eukaryotes they are not, and instead are present either as one or more clusters of repeats (e.g., Drosophila melanogaster, chicken, Arabidopsis thaliana, and human), in other repeat clusters (e.g., crustaceans and dinoflagellates), or entirely linearly dispersed (e.g., Neurospora crassa and Schizosaccharomyces pombe) (56, 65–70). However, these differences in the linear organization of the 5S genes between species belies commonalities in their spatial localization. For example, in mice ectopic 5S rDNA gene sequences have been shown to promote nucleolar localization (71). Similarly, in humans, one of the transcribed, linear clusters of 5S genes on chromosome 1 was shown to localize to a perinucleolar compartment (72). Moreover, the linearly dispersed 5S genes in many other eukaryotes have been shown to co-localize with nucleoli in three dimensions (73), suggesting that there are benefits to co-localizing the 5S genes with the other ribosomal genes. This is strong evidence for the nucleolus playing a direct role in the spatial organization of the nucleus.

The co-localization of RNAPIII decoded loci with nucleoli is not restricted to the 5S rDNA – tRNA genes also show interesting patterns of spatial organization. Eukaryotic tRNA genes are generally dispersed throughout the linear genomes, although in rare cases there are isolated linear clusters of tRNA genes. Xenopus laevis oocytes have developmentally regulated tRNA genes that are found in clusters (74), and multiple clusters of tRNA genes in S. pombe are located within the centromeric heterochromatin (73, 75, 76). They are also frequent sites of genomic rearrangements (77, 78). In S. cerevisiae, both microscopy and crosslinking proximity analysis show that tRNA genes cluster together and co-localize with the nucleolus (37, 42, 79–81). In addition, a smaller cluster of tRNA genes has also been identified at the centromere of S. cerevisiae (40, 43), consistent with the observation that the tRNA genes in S. pombe are primarily clustered at the centromere at a position offset from the nucleolus (82). As previously noted for 5S rDNA sequences, yeast tRNA coding regions confer interaction specificity with the nucleolus (42), indicating that position alone is insufficient to explain this phenomenon.

Little is known about the three-dimensional organization of tRNA genes in most eukaryotes, however, and whether they co-localize with nucleoli. This is important to determine, as metazoan nuclei can be 100 times larger than yeast but have only 2–3 times as many tRNA genes (83). Thus, there is a significantly greater structural problem to overcome, and the relative effect of tRNA gene clustering on overall genome organization will be much less. In this context, if RNAPIII transcription units are key components for spatial organization, a significantly more frequent DNA element would be needed in complex eukaryotes. In this context, it is interesting to consider that SINEs, retrotransposons derived from RNAPIII transcripts (usually tRNA and 7SL RNA), are found in great quantities in large eukaryotic genomes (84–86). There is evidence that SINEs can form clusters in mammalian nuclei (87, 88) and substantial evidence that at least some SINEs bind RNAPII clusters in vivo (89). It will be interesting to test whether some subset of these SINE clusters co-localize with nucleoli, especially in light of the finding that Alu SINEs are processed in the nucleolus (90).

By definition, rDNA:non-rDNA interactions must involve interplay between different loci, but it need not be direct and may involve RNA, proteins or other factors (e.g., epigenetic modifications) that facilitate either directed or self-assembled interactions. Irrespective of how the associations are stabilized, they must be flexible enough to
allow reassignment of the rDNA repeat to another function without interfering with the primary function of the nucleolus – ribosome production. A simple model to explain the origin of these interactions is that the act of transcription or transcriptional regulation is responsible for interaction formation and/or maintenance. This is consistent with polymerase class-dependent association of different regions of the rDNA, and more generally with the idea of transcription factories. However, in the yeast strain where all the rDNA repeats are transcriptionally active due to enforced reductions in copy number, little or no phenotype is observed (17). It is possible that interactions with the rDNA may function to position non-nucleolar loci during nuclear division (91), when the rDNA are transcriptionally or replicatively inactive and accessible to other factors. In this case, transcription would not be the sole driver of interactions that involve the rDNA repeats.

The nucleolus and heterochromatin

The nucleolus appears to influence the chromatin structure of the DNA that surrounds it. In metazoans, the nucleolus is commonly observed to be surrounded by shell of late-replicating heterochromatin. Similarly, tephritids (fruit flies) and other dipterans (true flies) also exhibit preferential associations of the rDNA with heterochromatin-rich chromosomes (92). In Drosophila, there appears to be a direct relationship between the nucleolus and non-rDNA heterochromatin (93, 94). Furthermore, nucleolar association seems to be an important factor to maintain the heterochromatic state of the inactive human X chromosome (95), with the Barr body originally being known as the ‘nucleolar satellite’ (96). Analysis of nucleolus-associated chromatin domains (NADs) in two human cell lines (i.e., HeLa and HT1080) identified satellite repeats as being the major components of the NADs (97). Repetitive elements have also been implicated as forming part of the NAD in yeast (98).

Overlap between some metazoan NADs and reported lamina-associated domains suggests that specific genomic regions could alternate between associating with the nucleolus and the nuclear periphery, either in different cells or at different times (97, 99). The regulation of this recruitment would necessarily affect the organization of the remainder of the genome too. However, large-scale relocations are not a necessity if relative long-range positioning can be maintained through alterations to the compaction levels of intervening regions, rather than simply by physical relocation of the DNA. In effect, some contacts can be broken whereas others are maintained. However, direct recruitment of non-rDNA loci to the nucleolar boundary remains to be demonstrated. Therefore, regulation of these interactions in response to specific signals or pathways is still a hypothesis that requires testing.

Do bacteria have nucleoli and do they also function to organize the nucleoid?

It has traditionally been thought that bacteria lack the equivalent of a nucleolus as their repetitive ribosomal DNA genes are organized as dispersed repeats. However, it is clear that the bacterial nucleoid is structured (100–108), and recent evidence suggests that the rRNA genes in Escherichia coli may be transcribed in specific foci in the cell, opening up the idea that bacteria contain a nucleolus-like structure (109, 110) to facilitate recycling of RNA polymerase and coordination of ribosome assembly (111).

The different copies of bacterial ribosomal RNA genes, including the spacer regions, have high levels of sequence similarity. This finding was unexpected given the apparent dispersal of these genes in the genome. It has been proposed that sequence similarity is maintained through a process of gene conversion (112). Therefore, putative bacterial nucleoli may serve not only to optimize rRNA transcription and hence growth (111) but also to juxtapose ribosomal DNA genes to facilitate gene conversion between the disparate copies.

Whether the bacterial equivalent of a nucleolus actually exists is an important area for future study as it will shed light on critical aspects of bacterial growth rate regulation (111).

Conclusion

Accepting that the nucleolus is not simply a raft adrift the nuclear landscape, what advantage is there in the nucleolus controlling nuclear structure? We contend that the answer lies in the central position that ribosomes have within cellular metabolism (Figure 2). Stresses of all kinds affect ribosome activity [e.g., reviewed in (113)], the production of ribosomes, and consequently the nucleolus itself. Responses to stress [e.g., reviewed in (114)] may sometimes involve gross alterations to nucleolar structure [e.g., (115)]. These alterations have been related to the release and stabilization of
proteins from the nucleolus [e.g., (115) and reviewed in (114)]; therefore, it is likely that alterations to the NADs associated with the nucleolar boundary also occur during stress response, but direct evidence for this is lacking. Our hypothesis predicts that such alterations occur and cause stress-related alterations to the associated genes, and these events are part of how the stress response is relayed to appropriate transcriptional networks outside the nucleolus (Figure 2). Thus, nucleolar structure acts as an intermediary between the genomic structural network that coordinates transcription and the cytoplasmic translational network (Figure 2). The fact that regions of the nucleolus are acted on by the three different polymerases supports the sensory role of the rDNA. This model is conceptually similar to the rDNA theory of aging proposed by Kobayashi (116). In this theory, the repetitive nature of the rDNA makes it uniquely prone to instability, and this instability acts as an early warning system for general genomic instability, triggering the aging pathway. Therefore, we propose that nucleolar structure is the keystone that synchronizes expression and cellular responses by linking the distinct genomic and cytoplasmic protein networks within cells.

Acknowledgements: J.M.O.S. and A.G.C. are supported by the Marsden Fund. J.M.O.S. was also funded by Gravida: National Centre for Growth and Development. D.R.E. and D.A.P. were funded by the National Institutes of Health grant GM082875. D.A.P. was also funded by the National Institutes of Health University of Michigan Genetics Pre-doctoral Training Grant (T32 GM07544) and a Rackham Merit Fellowship. A.R.D.G. is supported by the Marsden Fund.

Received October 15, 2012; accepted December 21, 2012

References


64. French SL, Osheim YN, Cioci F, Nomura M, Beyer AL. In exponentially growing Saccharomyces cerevisiae cells, rRNA synthesis is determined by the summed RNA polymerase I loading rate rather than the number of active genes. Mol Cell Biol 2003; 23: 1558–68.
75. 2006; 170: 639–41.

Justin M. O’Sullivan studied Cellular and Molecular Biology at Canterbury University, obtained his PhD at Otago University, and completed post-doctoral work at the University of Kent and the University of Oxford. He was appointed a Senior Lecturer at Massey University, New Zealand. He is currently a Senior Research Fellow at the Liggins Institute at the University of Auckland.

Dave Pai studied Biophysics at Johns Hopkins University before attending the University of Michigan, obtaining a PhD in Biological Chemistry in the group of David Engelke.

Andrew Cridge graduated in New Zealand from Lincoln University (B.Sc., Hons.) and Otago University (PhD). He completed post-doctoral studies in post-transcription regulation of gene expression at the University of Manchester (UK) and Massey University (New Zealand). Additional research at Massey University focused on examining the spatial organization of gene structure in its role in gene regulation. Currently, he is a post-doctoral fellow at the Laboratory for Evolution and Development, at the University of Otago.

David Engelke trained in Biochemistry as an undergraduate at the University of Wisconsin, obtained his PhD in Molecular Biology from Washington University in St. Louis, and completed post-doctoral work at the University of California, San Diego and the California Institute of Technology. He is a Professor of Biological Chemistry at the University of Michigan, where for 30 years his group has studied the biochemistry, genetics, and cell biology of small RNA biosynthesis in eukaryotic nuclei.

Austen Ganley has a PhD in Molecular Genetics from Massey University, New Zealand. He was a Clark Postdoctoral Fellow in Molecular Evolution and Comparative Genomics at Duke University, NC, USA. He also worked in the National Institute for Basic Biology and the National Institute of Genetics in Japan with Takehiko Kobayashi before returning to Massey University (Albany) in New Zealand to establish his own group in 2008. His main interests are in understanding the biology and evolution of the ribosomal DNA repeats.