For reprint orders, please contact: reprints@futuremedicine.com

Structural genomic changes during mammalian ontogeny: a new dimension

“Evidence is emerging that programmatic modification of the primary structure of the inherited genome may in fact be common, during development and in adult tissues.”

KEYWORDS: DNA methylation ■ epigenetic ■ genome structure instability ■ LINE-1 elements ■ mammalian development ■ mosaicism ■ ontogeny ■ *rDNA* ■ tissue specificity

Differentiated tissues are elaborated during mammalian ontogeny by the coordinated sequential execution of cell type-specific gene expression programs. A common supposition is that basic inherited genomic structure remains constant during this process. This supposition is critically important for the recently developed procedures for obtaining induced pluripotent stem cells from somatic tissues, with great potential for clinical applications.

Is this supposition comprehensively true? Would it not make sense, in terms of energy conservation, to program the necessary changes permanently through alterations in genome structure? Several special-case examples of such purposeful, site-specific manipulation of primary genome structure have long been known to exist [1,2] and have obvious functional consequences during a specific life-cycle stage. These are most common in unicellular organisms but also occur in higher organisms, for example, amplification of specific genes for rapid production of proteins, such as those of the chorion of insect eggs, and rearrangements to produce diversification of variant surface glycoprotein in trypanosomes. Vertebrates also utilize genomic rearrangement to enhance diversity of antigen receptors, through the action of the DNA sequence-specific V(D)J recombinases.

Evidence is emerging that programmatic modification of the primary structure of the inherited genome may in fact be common, during development and in adult tissues. Structural genetic differences are known to occur between cell populations within an individual, including individual humans, constituting somatic mosaicism [3]. These differences, though frequent, have been interpreted as abnormalities, resulting from uncorrected somatic

mutations, happening spontaneously or from environmental insult.

Mammalian brains are somatic mosaics, due to aneuploidy and genomic copy number variations. Recently, it also has been found that active transposable elements, long interspersed nuclear elements 1 (LINE-1) in neuronal tissue, contribute to the mosaicism, with tissue specificity [4,5]. LINE-1 elements influence chromosome integrity and gene expression. Retrotransposition activity for the LINE-1 elements was strong during neuronal differentiation and occurred with high frequency. Activity was also found in the adult brain. Regulatory events leading to LINE-1 activation were elucidated [6]. These observations led to the proposition that LINE-1 element transposition has a purpose: adaptive increase in neuronal variation and plasticity and in brain-controlled phenotypes [7]. However, specific sites of insertion will need to be located to confirm this idea.

Another line of evidence, implicating the occurrence of developmentally-programmed genomic structural changes, comes from a recent study of the gene for ribosomal DNA (*rDNA*) during ontogeny in the mouse [8]. There are hundreds of copies of *rDNA* in each cell on several chromosomes. The *rDNA* of the mice was found to present several SNPs in promoter regions of the gene. The relative percentages of these variant SNPs, indicative of *rDNA* structural status, were determined in sperm, embryos and two differentiated tissues, lung and liver, using a high-throughput quantitative pyrosequencing technique. The percentage of the variants changed in the differentiated tissues: in the males they differed significantly in lung and liver compared with sperm, and in the females they differed in lung compared with liver. Second, within-litter

Yih-Horng Shiao[†]

Laboratory of Comparative Carcinogenesis, Building 538, Room 205A, National Cancer Institute at Frederick PO Box B, Frederick, MD 21702, USA

Lucy M Anderson[†]

Author for correspondence: Laboratory of Comparative Carcinogenesis, Building 538, Room 205A, National Cancer Institute at Frederick PO Box B, Frederick, MD 21702, USA

lmmandersonphd@gmail.com

[†]Authors contributed equally

Future
Medicine part of

fsg

variances were calculated. These variances should be constant throughout development, if the genomic ratio of *rDNA* SNPs is established at fertilization. By contrast, for each of the five *rDNA* SNP variants, within-litter variances were significantly ($p < 0.0001$) greater for lung, liver and sperm than for embryos [8]. Thus, there was tissue-specific developmental adjustment in the structure of the *rDNA*. This phenomenon is uniquely different from previously reported specific structural genomic changes noted above. Rather than pertaining only to a particular stage of development or class of proteins, or a noncoding repetitive element, *rDNA* is an ubiquitously expressed gene that is of central importance to the growth and energy economy of all cells [9]; rRNA biogenesis is rate-limiting for ribosome production and protein synthesis.

“There was tissue-specific developmental adjustment in the structure of the *rDNA* ... a ubiquitously expressed gene that is of central importance to the growth and energy economy of all cells.”

Are the nonrandom shifting ratios of *rDNA* variants during differentiation of mouse tissues of functional significance? Information on this possibility was limited but suggestive. The *rDNA* SNP sites in the main promoter all have potential involvement in control of expression of the gene. The -218 and -178 sites are located within the repetitive enhancer elements. The -178 site is only a few bases downstream from the termination of the RNA transcript initiated at the spacer promoter, and a few bases upstream from the binding site for transcription termination factor-1, a critical regulator of *rDNA* silencing and expression. The -104 SNP is within the sequence for the regulatory promoter RNA (pRNA) [10].

Methylation of specific CpG sites is a critical aspect of *rDNA* expression and regulation [11]. In the mouse tissues, the degree of methylation of a concentration of CpG sites in one of the *rDNA* promoters showed ontogenetic regulation, with the highest methylation in sperm, and was higher in lung than in liver tissues [8]. There was a significant negative or positive correlation between the percentage of CpG methylation and the percentage of one of the *rDNA* variants in the majority of tissues. Interestingly, the percentage of the variant, which showed the most prominent ontogeny-related shift in relative amounts, was strongly correlated with the percentage of CpG methylation in all differentiated tissues, but less so in embryos. In an *in vitro*

transcription assay [8], the mouse *rDNA* ontogenetic variants were differentially methylated, and differentially transcribed. Thus, genomic and epigenetic modulations may work hand-in-hand during tissue differentiation.

These results suggest that the structural shifting of the *rDNA* SNP copies during ontogeny is functionally purposeful; more research will be required to confirm this. Published reports indicate that functional differences among *rDNA* variants are plausible. Recently, polymorphisms in the size of the hypervariable region, just upstream from the main promoter region of *rDNA*, have revealed that the intergenic rRNA, transcribed from the spacer promoter, is derived from a subclass of variants [12]. Several mouse *rDNA* variants with SNPs in the transcribed region had tissue specific expression [13], and SNP profiles in the *rDNA* promoter region were associated with the binding of specific regulatory factors [14].

The mechanism of the variant- and tissue-specific genomic shifts in *rDNA* is an intriguing question. Expression of the V(D)J recombinase, RAG, has recently been described in the brain, leading to speculation that site-specific recombination could have functional significance beyond the immune system [15]. There is one perfect CACAGTG and four CACTGTG sites in murine *rDNA* (GenBank BK000964) for potential recognition by the RAG recombinase. Nonconsensus substitutions at positions four or five of the RAG recognition sequence destabilized but did not prevent RAG binding, with retention of potential for recombination [16]. The RAG complex can target DNA without the canonical heptamer nearby, by engaging HMGB1/2 activity [17], such as might be provided by upstream binding factor, the master regulator of *rDNA* chromatin conformation. In addition, sequence-specific recombination in *rDNA* may occur as a result of p53 binding [18], and sequence-specific cleavage of *rDNA* by topoisomerases has been noted [19,20].

“...the mouse *rDNA* ontogenetic variants were differentially methylated, and differentially transcribed. Thus, genomic and epigenetic modulations may work hand-in-hand during tissue differentiation.”

The ontogenetic shifts in ratios of *rDNA* variants and the putative functional consequences of these shifts obviously may be a component of the required close regulation of rRNA production, in the context of cell economy and

growth. In addition, there are hints that *rDNA* and rRNA may have wider roles. It has been proposed that *rDNA*, due to its high intrinsic instability, plays a global role in maintenance of overall genome stability [21]. In *Drosophila*, *rDNA* copy number influenced degree of heterochromatin for unlinked genes [22]. Biogenesis of rRNA has been implicated in some diverse human outcomes, including suicide [23] and Down's syndrome [24]. In humans, sizes of *rDNA* clusters showed a very high degree of variability and essentially complete heterozygosity in white blood cells [25]. There was a high rate of meiotic recombination and evidence for somatic recombination in approximately a third of individuals. Furthermore, human *rDNA* gene clusters have been discovered to be recombinational hotspots in cancer: half of lung or colorectal cancers had *rDNA* rearrangements relative to surrounding tissues or blood [26]. rRNA is strongly implicated in cancer [27] and is currently a target in anti-cancer therapeutic design [28]. In human lung cancer cell lines, there was a negative correlation between a SNP in the 5' leader of rRNA and levels of a ncRNA [29]. Might there be disease-related variant – and/or cell-type specificity – in these *rDNA* polymorphisms and recombination phenomena?

A final question is whether these observations of *rDNA* are specific to this particular gene, or might other ubiquitously expressed genes also undergo specific, functionally relevant genomic structural evolutions during ontogeny? This could especially be likely for those present in multiple copies with a high likelihood for recombination, such as repetitive elements that have acquired regulatory roles as promoters, enhancers, splice sites and insulators. But the special case of V(D)J recombinase in the immune system illustrates that any gene could, in theory, be subject to ontogenetic structural processing. This possible new dimension for developmental genetics deserves further exploration, particularly in view of its implications for stem cell usage and for cancer development.

Financial & competing interests disclosure

This research was supported in part by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

References

- Bostock CJ. Chromosomal changes associated with changes in development. *J. Embryol. Exp. Morph.* 83(Suppl.), S7–S40 (1984).
- Borst P, Greaves DR. Programmed gene rearrangements altering gene expression. *Science* 235(4789), 658–667 (1987).
- De S. Somatic mosaicism in healthy human tissues. *Trends Genet.* 27(6), 217–223 (2011).
- Muori AR, Chu VT, Marchetto MC, Deng W, Moran JV, Gage FH. Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. *Nature* 435(7044), 903–910 (2005).
- Coufal NG, Garcia-Perez JL, Peng GE *et al.* L1 retrotransposition in human neural progenitor cells. *Nature* 460(7259), 1127–1131 (2009).
- Moutri AR, Marchetto MC, Coufal NG *et al.* L1 retrotransposition in neurons is modulated by MeCP2. *Nature* 468(7322), 443–446 (2010).
- Singer T, McConnell MJ, Marchetto MC, Coufal NG, Gage FH. LINE-1 retrotransposons: mediators of somatic variation in neuronal genomes? *Trends Neurosci.* 33(8), 345–354 (2010).
- Shiao YH, Leighty RM, Wang C *et al.* Ontogeny-driven *rDNA* rearrangement, methylation, and transcription, and paternal influence. *PLoS ONE* 6(7), e22266 (2011).
- Moss T. At the crossroads of growth control; making ribosomal RNA. *Curr. Opin. Genet. Develop.* 14(2), 210–217 (2004).
- Mayer C, Neubert N, Grummt I. The structure of NoRC-associated RNA is crucial for targeting the chromatin remodelling complex NoRC to the nucleolus. *EMBO Rep.* 9(8), 774–780 (2008).
- McStay B, Grummt I. The epigenetics of rRNA genes: from molecular to chromosome biology. *Ann. Rev. Cell Dev. Biol.* 24, 131–157 (2008).
- Santoro R, Schmitz K, Sandoval H, Grummt I. Intergenic transcripts originating from a subclass of ribosomal DNA repeats silence ribosomal RNA genes in trans. *EMBO Rep.* 11(1), 52–58 (2010).
- Tseng H, Chou W, Wang J, Zhang X, Zhang S, Schultz RM. Mouse ribosomal RNA genes contain multiple differentially regulated variants. *PLoS ONE* 3(3), e1843 (2008).
- Zhang S, Wang J, Tseng H. Basonuclin regulates a subset of ribosomal RNA genes in HaCaT cells. *PLoS ONE* 2(9), e902 (2007).
- Gericke GS. An integrative view of dynamic genomic elements influencing human brain evolution and individual neurodevelopment. *Med. Hypo.* 71(3), 360–373 (2008).
- Arnal SM, Holub AJ, Salus SS, Roth DB. Non-consensus heptamer sequences destabilize the RAG post-cleavage complex, making ends available to alternative DNA repair pathways. *Nucleic Acids Res.* 38(9), 2944–2954 (2010).
- Zhang M, Swanson PC. HMGB1/2 can target DNA for illegitimate cleavage by the RAG1/2 complex. *BMC Mol. Biol.* 10, 24 (2009).
- Boehden GS, Baumann C, Siehler S, Wiesmuller L. Wild-type p53 stimulates homologous recombination upon sequence-specific binding to the ribosomal gene cluster repeat. *Oncogene* 24(26), 4183–4192 (2005).
- Vogelauer M, Camilloni G. Site-specific *in vivo* cleavages by DNA topoisomerases I in the regulatory region of 35S rRNA in *Saccharomyces cerevisiae* are transcription independent. *J. Mol. Biol.* 293(1), 19–28 (1999).
- Halligan BD, Davis JL, Edwards KA, Liu LF. Intra- and intermolecular strand transfer by HeLa DNA topoisomerase I. *J. Biol. Chem.* 257(7), 3995–4000 (1982).

- 21 Kobayashi T. Regulation of ribosomal RNA gene copy number and its role in modulating genome integrity and evolutionary adaptability in yeast. *Cell Mol. Life Sci.* 68(8), 1395–1403 (2011).
- 22 Paredes S, Maggert KA. Ribosomal DNA contributes to global chromatin regulation. *Proc. Natl Acad. Sci. USA* 106(42), 17829–17834 (2009).
- 23 McGowan PO, Sasaki A, Huang TC *et al.* Promoter-wide hypermethylation of the ribosomal RNA gene promoter in the suicide brain. *PLoS ONE* 3(5), e2085 (2008).
- 24 Demirtas H. AgNOR status in Down's syndrome infants and a plausible phenotype formation hypothesis. *Micron* 40(5–6), 511–518 (2009).
- 25 Stults DM, Killen MS, Pierce HH, Pierce AJ. Genomic architecture and inheritance of human ribosomal RNA gene clusters. *Genome. Res.* 18(1), 13–18 (2008).
- 26 Stults DM, Killen MW, Williamson EP *et al.* Human rRNA gene clusters are recombinational hotspots in cancer. *Cancer Res.* 69(23), 9096–9104 (2009).
- 27 Montanaro L, Trete D, Derenzini M. Nucleolus, ribosomes, and cancer. *Am. J. Pathol.* 173(2), 301–310 (2008).
- 28 Drygin D, Rice WG, Grummt I. The RNA polymerase I transcription machinery: an emerging target for the treatment of cancer. *Ann. Rev. Pharmacol. Toxicol.* 50, 131–156 (2010).
- 29 Shiao YH, Lupascu ST, Gu YD *et al.* An intergenic non-coding rRNA correlated with expression of the rRNA and frequency of an rRNA single nucleotide polymorphism in lung cancer cells. *PLoS ONE* 4(10), e7505 (2009).