ISSN 1022-7954, Russian Journal of Genetics, 2010, Vol. 46, No. 9, pp. 1052–1054. © Pleiades Publishing, Inc., 2010.
Original Russian Text © J.A. Makarova, D.A. Kramerov, 2010, published in Genetika, 2010, Vol. 46, No. 9, pp. 1185–1187.

Small Nucleolar RNAs and Their Genes in Vertebrates

J. A. Makarova and D. A. Kramerov

Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, 119991 Russia;
e-mail: kramerov@eimb.ru

Received January 28, 2010

Abstract—Genes of box C/D small nucleolar RNAs (snoRNAs) were searched for in the genomes of members of all classes of vertebrates that do not belong to placental mammals. A tendency for an increase in the number of copies of snoRNA genes was observed in such vertebrates. This trend was most pronounced in anamnia (amphibians and fish). Box C mutations were found in 14 snoRNAs in all gene copies among all species studied. The role of the described events is discussed.

DOI: 10.1134/S1022795410090097

Small nucleolar RNAs (snoRNAs) are among the largest groups of protein-noncoding RNAs. In complex with proteins, they produce two most widespread rRNA modifications: pseudouridylation and 2'-O-methylation of ribose (approximately 100 nucleotides are subject to one of the two types of modification in human rRNA [1]). Owing to the complementary interaction with rRNA, snoRNAs determine the site of modification, and proteins associated with them realize the modification. In accordance to the type of modification, snoRNAs are classified into two families: H/ACA snoRNAs guiding pseudouridylation and C/D snoRNAs guiding 2'-O-methylation of ribose. Most of known snoRNAs of both families are involved in rRNA modification. For instance, 93 of 121 snoRNAs of the C/D family participate in methylation of human rRNA. C/D RNAs owe their name to the conserved sequences C (UGAUGA), D' (CUGA), and D (CUGA) that are present in them. The C and D sequences are located close together due to the complementary interactions between the snoRNA ends. Located downstream from D and/or D' are so-called antisense elements (regions of 9–20 nucleotides), which are complementary to the corresponding rRNA sequences and able to interact with them. Modification affects the remaining part that is included in the RNA/RNA helix and four nucleotides apart of the D and/or D' sequence. Modifications can stabilize the spatial structure of rRNA and are necessary for the normal ribosome function [2, 3].

Almost all snoRNAs of vertebrates are encoded in an unusual way: their genes reside in introns of other genes, so-called host genes [4]. This work continues our study [5] of the number and location of snoRNA genes of the C/D famly in vertebrate genomes.

Genes of 93 box C/D snoRNAs involved in rRNA modifications were searched for using the algorithm WU-BLAST 2.0 (http://www.ensembl.org/Multi/blastview) in genomes of the following species: opossum (*Monodelphis domestica*), duckbill (*Ornithorhynchus anatinus*), chicken (*Gallus gallus*), lizard (*Anolis carolinensis*), frog (*Xenopus tropicalis*), and zebrtafish (*Danio rerio*). At the first stage of the search, nucleotide sequences of human snoRNA genes from the snoRNA database [6] were used. Then, using the found nucleotide sequences of vertebrate snoRNA genes, their possible homologs were searched for in genomes of the corresponding species with the use of the same algorithm. The found candidate sequences with conserved C, D/D' boxes and antisense elements were aligned with the vertebrate genomes using the BLAST (http://genome.ucsc.edu) algorithm. To elucidate whether the candidate sequences are located in the introns of host genes, the mRNA and EST databases were used. In case the identified sequences contained the C, D/D' boxes and an antisense element and were flanked by short inverted repeats and resided in introns of another gene, they were considered as snoRNA genes, and other copies were additionally manually searched for in the host gene introns.

Most (about 70%) of C/D snoRNA genes in placental mammals are represented by a single copy, and the remaining genes have two or three copies. In representatives of some other classes of vertebrates, the number of multicopy C/D snoRNA genes increases [5]. In this work, we performed a search for 93 C/D snoRNA genes in genomes of representatives of all classes of vertebrates that do not belong to placental mammals.

The number of multicopy snoRNA genes increases already in primitive mammals, such as opossum and duckbill (Fig. 1). The same tendency seems to be preserved in chickens and lizards, but the lack of data does not permit unambigous conclusions to be made. In anamnia (frog and zebrafish), the number of genes with mulitple copies increases so that it exceeds the number of single-copy genes. Note that not only the number of multicopy genes, but also the number of

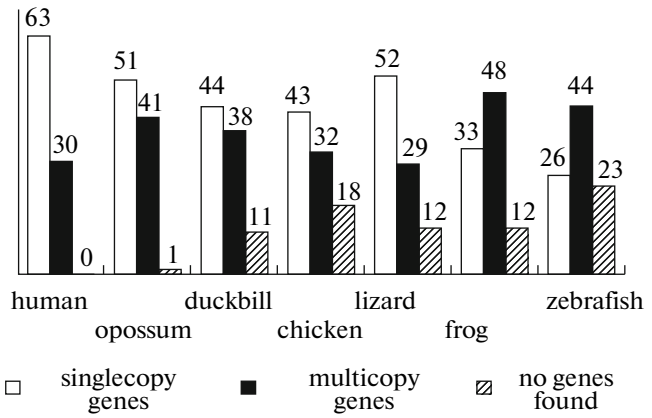

Fig. 1. The number of single- and multicopy C/D snoRNA genes in representatives of different classes of vertebrates.

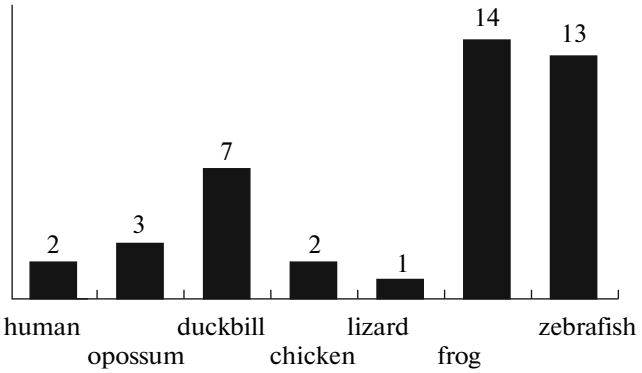

Fig. 2. The number of C/D snoRNA genes represented by four and more copies.

gene copies increases. For instance, in placental mammals almost all multicopy genes are represented by two or three copies, and the copy number in anamnia is often four and more (up to 13) (Fig. 2).

The numbers of multicopy genes in vertebrates, other than placental mammals, are actually higher than our present estimates. This is due to the fact that more than half of snoRNA host genes in vertebrates of this group (192 of 370) contain unsequenced regions, whereas host genes of placental mammals (human and mouse) lack such defects. With rare exceptions, all snoRNA gene copies are harbored in introns of the same host gene. Therefore, the discovery of extra snoRNA gene copies in unsequenced regions is probable almost in half of cases. Not a single gene was found for some C/D box snoRNAs (Fig. 1). This is partly due to differences in the rRNA methylation patterns in different classes of vertebrates, but mainly due to the fact that vertebrate genomes contain as yet unsequenced regions.

An increase in the number of copies of C/D snoRNA genes in vertebrates, except placental mammals, is not associated with the genome size: among all studied species, the largest genomes are in humans, mice, and opossum [7]. The trend for an increase in the number of copies of snoRNA genes is most pronounced in representatives of anamnia (frog and zebrafish). The observed phenomenon is as yet difficult to interprete conclusively. It is quite probable, however, that a decrease in the number of copies of C/D snoRNA genes in placental mammals (or an increase in their copy number in other vertebrates) can be explained by the fact that egg cells in placental mammals are smaller in size and thus contain less ribosomes.

The nucleotide sequence of the box C (UGAUGA) is important for the normal snoRNA function. About 30% of vertebrate snoRNA genes found by us (618 of 922) have one nucleotide substitution in this box. In most cases, there are other copies of the same gene containing an intact C box. Box C mutations interfere with binding of specific C/D RNP proteins and, correspondingly, disturb the C/D RNP function [8]. We found that in 14 snoRNAs mutations in the C box were present in all gene copies in all studied species. Apparently, this feature of the C-box structure is maintained by selection. It seems that in this way different levels of methylation of rRNA sites can be achieved. The methylated state of some sites of modification may, probably, be more important than of others. Evidence in favor of existence of such differences was reported in a number of publications (e. g., see [9]). The data obtained by us indicate the way by which differential methylation of rRNA sites could be realized: alteration in the number of gene copies and changes in the nucleotide sequences of binding sites of C/D snoRNP proteins.

ACKNOWLEDGMENTS

The work was supported by the Russian Foundation for Basic Research (grant no. 08-04-00350) and the Program of the Russian Academy of Sciences "Molecular and Cell Biology."

REFERENCES

1. Maden, B.E., The Numerous Modified Nucleotides in Eukaryotic Ribosomal RNA, *Prog. Nucleic Acid Res. Mol. Biol.*, 1990, vol. 39, pp. 241–303.
2. Kiss-Laszlo, Z., Henry, Y., Bachellerie, J.P., et al., Site-Specific Ribose Methylation of Preribosomal RNA: A Novel Function for Small Nucleolar RNAs, *Cell*, 1996, vol. 85, pp. 1077–1088.
3. Decatur, W.A. and Fournier, M.J., rRNA Modifications and Ribosome Function, *Trends Biochem. Sci.*, 2002, vol. 27, pp. 344–351.
4. Makarova, Yu.A. and Kramerov, D.A., Small Nucleolar RNA Genes, *Russ. J. Genet.*, 2007, vol. 43, no. 2, pp. 149–158.

5. Makarova, J.A. and Kramerov, D.A., Analysis of C/D Box snoRNA Genes in Vertebrates: The Number of Copies Decreases in Placental Mammals, *Genomics*, 2009, vol. 94, pp. 11–19.
6. Lestrade, L. and Weber, M.J., snoRNA-LBME-db, a Comprehensive Database of Human H/ACA and C/D Box snoRNAs, *Nucl. Acids Res.*, 2006, vol. 34, pp. 158–162.
7. Gregory, T.R., Nicol, J.A., Tamm, H., et al., Eukaryotic Genome Size Databases, *Nucl. Acids Res.*, 2007, vol. 35, pp. 332–338.
8. Watkins, N.J., Dickmanns, A., and Luhrmann, R., Conserved Stem II of the Box C/D Motif Is Essential for Nucleolar Localization and Is Required, along with the 15.5K Protein, for the Hierarchical Assembly of the Box C/D snoRNA, *Mol. Cell. Biol.*, 2002, vol. 22, pp. 8342–8352.
9. Esguerra, J., Warringer, J., and Blomberg, A., Functional Importance of Individual rRNA 2'-O-Ribose Methylations Revealed by High-Resolution Phenotyping, *RNA*, 2008, vol. 14, pp. 649–656.