
MOLECULAR
GENETICS

Monomers of a Satellite DNA Sequence of Chaffinch (*Fringilla coelebs* L., Aves: Passeriformes) Contain Short Clusters of the TTAGGG Repeat

I. A. Liangouzov¹, S. E. Derjusheva¹, A. F. Saifitdinova¹,
A. G. Malykh², and E. R. Gaginskaya

¹ Biological Institute, St. Petersburg State University, Stary Peterhof, St. Petersburg, 198504 Russia;
fax: (812)427-73-10; e-mail: chromas@paloma.spbu.ru

² Fidelity Systems Inc., Gaithersburg, MD 20879, United States; e-mail: fsi@fidelitysystems.com

Received April 12, 2002

Abstract—A novel repeated sequence of chaffinch (*Fringilla coelebs*) designated as GS was isolated from genomic DNA after in vitro amplification of satellite DNA sequences using GSP-PCR technique. The proportion of this repeat in the chaffinch genome constitutes about 0.2%. Monomers are 176 to 199 bp in size and contain a short cluster of the TTAGGG telomeric tandem repeat. The oligomer of the telomeric hexanucleotide is flanked by the sequences that are significantly different in different monomers. The GS sequences are organized as tandemly repeated units and located in a number of chromomycin-positive blocks on the long arms of macrochromosomes 1, 2, 3, 5, and 6, as well as on several microchromosomes. The sequences homologous to the GS satellite of chaffinch were not found in the genomes of redwing (*Turdus iliacus*) and house sparrow (*Passer domesticus*).

INTRODUCTION

Repetitive sequences constitute a substantial part of the eukaryotic genome. At the same time, the data on repetitive sequences in birds are very scarce. Relatively well studied are the satellite DNA sequences specific to sex chromosomes of some members of the order Galliformes [1]. Three different types of repetitive sequences with centromeric localization were also described. These are the PR1 in two pigeon species [2], FCP in chaffinch [3], and a 190-bp repeat found in the representatives of many families, including Falconidae, Gruidae, Psittacidae, Anatidae, and some others [4]. Satellite DNA sequences were found in the genomes of chicken [5], turkey [6], and quail [7]. The proposed localization of these repeats in the centromeric regions of microchromosomes is doubtful and refuted by the results of the in situ hybridization on the giant lampbrush chromosomes from the growing chicken oocytes [8]. Satellite DNA sequences were also isolated from the genomes of various other bird species, but their chromosomal localization was not examined [9–14].

One of the most popular techniques for the isolation of satellite sequences consists in cloning of the fragments forming distinct bands in the gels after electrophoretic separation of the genomic DNA fragments digested with restriction endonucleases. However, this standard technique can be applied only for analysis of sequences comprising more than 0.5% of the genome. Avian genomes are much smaller than the genomes of other vertebrates mostly due to the fact that the proportion of repetitive sequences in them is far lower [15].

For these reasons, isolation of avian satellite DNA sequences by use of standard methods often meets with certain difficulties. Recently, a genome self-priming PCR technique (GSP-PCR) enabling in vitro amplification of tandemly repeated genomic DNA sequences in polymerase chain reaction, where genomic DNA simultaneously serves as a template and as a primer for elongation of the nucleotide chain, has been described [14]. In the present study, this method is applied for the isolation and characterization of a novel GS satellite sequence of chaffinch (*Fringilla coelebs* L., Aves: Passeriformes) genomic DNA.

MATERIALS AND METHODS

Genomic DNA was extracted by standard methods from the whole blood samples of (1) chaffinches (*Fringilla coelebs*, Fringillidae, Passeriformes), caught for banding during spring migration in 2000 at the Kurshskaya spit Ornithological Station (Zoological Institute, Russian Academy of Sciences); (2) house sparrows (*Passer domesticus*, Ploceidae, Passeriformes); and (3) redwings (*Turdus iliacus*, Turdidae, Passeriformes).

Satellite DNA was amplified in vitro using the GSP-PCR technique [14]. The following modification of the method was used. An aliquot of genomic DNA was incubated at 100°C for 30 min. In these conditions DNA degraded with the formation of the fragments less than 1 kb in size. A total of 0.15 µg of the 1 : 1 mixture of degraded and nondegraded genomic DNA was added to a standard polymerase chain reaction mixture.

Amplification conditions included 8 min at 94°C, followed by 35 cycles of 50°C for 3 min; 72°C for 3 min; and 94°C for 1 min; and a final cycle of 50°C for 3 min; and 72°C for 20 min.

The products of GSP-PCR reaction were digested with the *Pst*I restriction endonuclease and separated by electrophoresis in 1% agarose gel. A fraction of DNA fragments of 190 to 200 bp in size was isolated from the gel by centrifugation on a spin-column (Sileks-M, Russia) and cloned into the pBluescript II KS (+) plasmid. Clones containing recombinant plasmids were designated GS-1, GS-2, and GS-3.

Recombinant plasmids were sequenced by use of standard pBluescript primers and the BlueDye Terminator Ready Reaction Mix (PE Applied Biosystems, United States) and a protocol recommended by the manufacturer with slight modifications described by us earlier [16]. Consensus sequence of the GS monomer from chaffinch genome was determined by direct genome sequencing method. Synthesized primers (Fidelity Systems, United States) specific for different GS regions, i.e., primers chemically modified in a way to prevent the synthesis of primer dimers, as well as non-specific PCR during sequencing, were as follows:

ZYAB008 GTTTTCGAGATGCTAACCTAGGCGTAACC
 ZYAB009 TAGGCAGGGCAGCCAATCGGCCTAGGGT

(positions of modified nucleotides are boldfaced and underlined).

Genome sequencing was carried out in a final volume of 5 µl, containing 100 ng of genomic DNA; 10 pmol of primer; 0.1 µl of ThermoFidase 2A (Fidelity Systems, United States); and 2 µl of BigDye Terminator Ready Reaction Mix. A panel with the samples was placed in the thermal cycler with the following program: 95°C for 2 min, followed by 400 cycles of 95°C for 5 s; 55°C for 30 s; and 60°C for 2 min. The results were analyzed using ABI377 automated DNA sequencer (PE Applied Biosystems, United States).

A search for homologous sequences in the worldwide GenBank database was performed using the BLASTN 2.2.1 software [17].

Relative quantity of the GS sequences in the genome was estimated by means of dot-blot hybridization. Organization of the GS sequences in the chaffinch genome was determined by use of Southern blotting. Genomic DNA was digested with the *Pvu*II and *Pst*I restriction enzymes. DNA fragments were immobilized onto Hybond-N⁺ membranes (Amersham Pharmacia Biotech, United Kingdom) by use of alkaline vacuum blotting with subsequent irradiation with UV light. Digoxigenin labeled (Digoxigenin-11-dUTP, Roche, Germany) cloned GS fragment was used as a hybridization probe. Labeling was carried out by means of PCR with standard M13 primers. Hybridization, washing, and hybridization signal detection procedures were carried out according to a standard protocol recommended by Roche Company. To prevent hybridization of the GS probe to the genomic DNA telomeric repeat, a prehybridization procedure with an excess of the (TTAGGG)₆ and (CCCTAA)₆ oligonucleotides was performed. The (TTAGGG)_n DNA sample, amplified by means of PCR with telomere-specific primers and without the DNA template, placed on a membrane, was used as control. Comparative densitometry of hybridization signals was carried out by use of the PhotoM 1.2 software designed by A. Chernigovskii (Ioffe Physi-

cotechnical Institute, Russian Academy of Sciences, St. Petersburg, Russia).

Metaphase spreads of chaffinch chromosomes were prepared from primary fibroblast cultures according to a standard method.

Fluorescence in situ hybridization (FISH) was carried out according to a standard technique [18]. The pGS-2 insert biotinylated (Biotin-16-dUTP; Roche, Germany) by PCR was used as the FISH probe. To prevent cross-hybridization of GS with the telomeric repeat, a carrier salmon sperm DNA was added to a hybridization mixture. Detection of hybridization signals was performed with conjugated avidin-FITC or avidin-Cy3. Hybridization signal was amplified with anti-avidin D-antibodies (Vector Laboratories, United Kingdom).

Slides were viewed with the DM RXA fluorescent microscope (Leica Wetzlar, Germany) equipped with white-and-black CCD camera (Cohu). Images were captured and processed with the QFISH software (Leica Cambridge, United Kingdom).

RESULTS AND DISCUSSION

Electrophoregrams of chaffinch genomic DNA digested with the *Bam*HI, *Hind*III, *Pst*I, *Sau*3AI, *Pvu*II, *Msp*I, and *Hpa*II restriction endonucleases were described in our previous paper [16]. Briefly, in addition to the main fraction comprised of 500-bp repeated DNA fragments and detected in the digests of the *Bam*HI, *Hind*III, *Pst*I, and *Sau*3AI restriction endonucleases, the use of the *Pvu*II or *Pst*I endonuclease enabled visualization of a band corresponding to the fraction of 200-bp fragments. The 500-bp repeat designated as FCP, was characterized as centromeric satellite element [3]. The fragments of about 200 bp in size represented another repeat of chaffinch genome. Our study was focused on the examination of this repeat.

Since repeated DNA fragments of about 200 bp in size constituted only a small proportion of chaffinch genomic DNA, to obtain preparative amounts of this DNA fraction we used the GSP-PCR technique, which

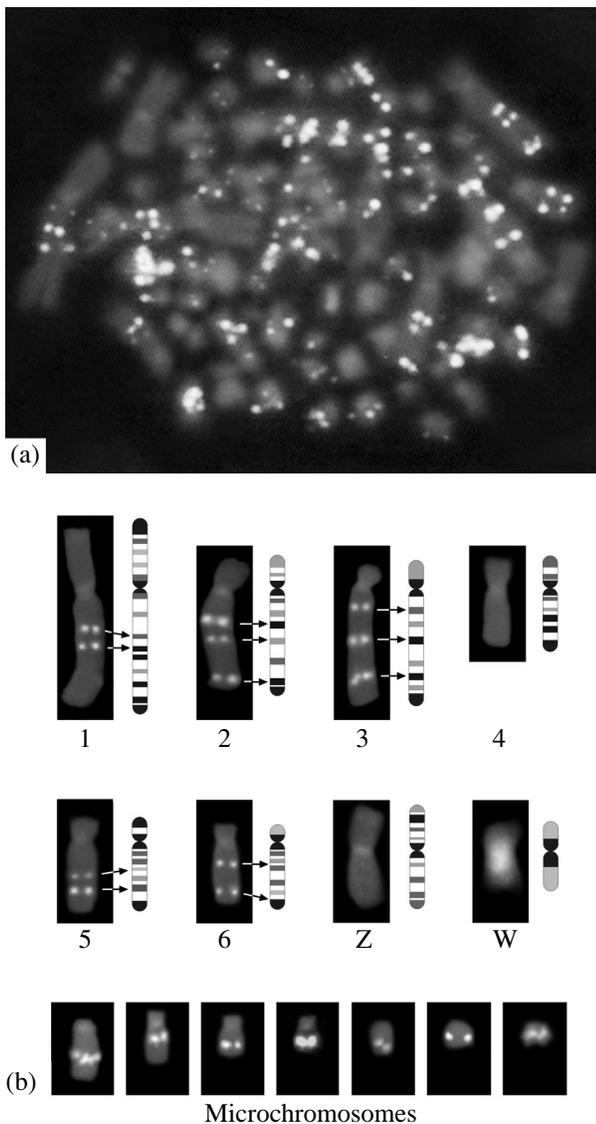


Fig. 3. (a) Metaphase spread; (b) individual chromosomes (macrochromosomes 1 to 6, Z and W sex chromosomes, and some microchromosomes) from chaffinch fibroblasts after FISH with cloned GS sequence as a probe. Arrows on ideograms of macrochromosomes indicate the positions of chromomycin-positive blocks containing the GS sequence.

is explained by high conservatism of this region, containing the recognition sites for the restriction enzymes used (Fig. 2). It is tempting to assume that these regions are functionally important and that natural selection acts against mutation in them. On the other hand, the repeat described can be evolutionary young: the monomers have already amplified, but have not yet diverged largely. The increased polymorphism level characteristic of the regions containing the TAACCC arrays can be associated with the presence of this simple repetitive sequence in the monomer. This repeat cause the increased mutation rate by possible strain slippage dur-

ing replication, repair and conversion, as well as because of recombination and other factors affecting the high rate of microsatellite divergence [19].

Relative content of the sequences homologous to the cloned GS fragments in the chaffinch genome was assessed by use of dot-blot hybridization technique. Comparative densitometry of the signals obtained after hybridization of different amounts of genomic DNA and pGS-1 reference DNA to digoxigenin labeled GS-1 fragment showed that proportion of the GS repeat in the chaffinch genome constitutes 0.2%, which approximately corresponds to 2×10^4 copies per diploid genome.

Dot-blot hybridization technique was also used for searching the GS homologous sequences in the genomes of other passerine birds: redwing and house sparrow. The hybridization data showed that genomes of these two species lacked the sequences homologous to the GS repeat of chaffinch. These data are consistent with the current views that tandemly organized satellite DNA sequences evolve very rapidly and usually are different even in phylogenetically close species [5–7, 20].

Mapping of the GS repeat to chaffinch metaphase chromosomes was carried out using FISH. The GS satellite DNA was localized in a number of blocks on the long arms of macrochromosomes 1, 2, 3, 5, and 6 (Fig. 3). Chromosomal locations of the examined satellite coincided with the chromosome regions that are intensively stained with the GC-specific fluorochrom, chromomycin A₃. These data are in a good agreement with the data on the enrichment of the repeat with GC. Depending on the structure of the monomer variable part, total G + C content of the GS sequence varies from 56 to 58%. After FISH, multiple hybridization signals were also revealed on microchromosomes. However, precise localization of the repeat on these chromosomes is impossible because of their small size. The Z and W sex chromosomes of chaffinch, as well as macrochromosome 4, lack the repeat clusters large enough to be detected by FISH.

In eukaryotic chromosomes satellite DNA sequences are located predominantly in the blocks of constitutive heterochromatin, which are usually detected by means of C-banding. This method, however, do not permit obtaining of reproducible banding patterns in case of avian chromosomes. Some authors suggest that the blocks not detectable by means of C-banding technique but displaying intense fluorescence after staining with chromomycin A₃ fluorochrom, can represent interstitial heterochromatin [21]. Until now, this proposal was not confirmed experimentally and seemed rather speculative. However, the coincidence of the regions of the chaffinch GS satellite localization with the chromomycin-positive blocks on metaphase chromosomes of this species suggests that avian karyotypes, and chaffinch karyotype in particular, contain the blocks of interstitial heterochromatin, which are not detectable by use of standard C-banding.

The presence of the tandemly repeated TTAGGG telomeric hexanucleotide in the GS satellite monomers deserves special interest. Analysis of telomeric repeats localization in the members of different orders of mammals, birds, reptiles, amphibia, and fish showed that the (TTAGGG)_n repeats can be also found in the internal chromosomal loci [22–24], mostly in the regions of constitutive heterochromatin [22]. The origin of nontelomeric sites of the TTAGGG repeat is unknown. It can be associated with such mechanisms as telomere–telomere fusion of ancestral chromosomes and/or amplification of the short regions of tandem repeat, present in the ancestral karyotypes as “cryptic telomeres.” Indeed, interstitial (TTAGGG)_n sites mark the regions of chromosome fusion in karyotypes of humans [25], Indian muntjac [26], okapi, and some other species [23, 27]. At the same time, interspersed TTAGGG repeats are found in the nucleolus organizer regions, i.e., the chromosome regions containing blocks of tandemly organized rRNA genes [28, 29]. Species-specific interspersed telomeric repeats within the pericentric chromosomal regions, not associated with the centric fusions was also observed [24, 30]. In this case a distinct correlation between the size of the heterochromatin blocks and the intensity of hybridization signals in them after FISH with telomeric probe was revealed. It, however, remained unclear whether the telomeric motif detected was a component of the repetitive unit of satellite DNA or it formed independent blocks alternated with satellite monomers. Analysis of the position of satellite DNA and the (TTAGGG)_n repeat relative each other in the region of their colocalization in centromeric heterochromatin showed that these sequences formed two independent compartments [31]. Our findings demonstrate the possibility of telomeric hexanucleotide expansion directly into monomers of satellite DNA. Short clusters of TTAGGG telomeric motif were earlier found in satellite DNA monomers from cat [32] and from one of the fish species [20]. At present, it remains obscure, how telomeric sequences appear in satellite DNA. It can be hypothesized that a certain role in this processes is played by long-lived single-stranded DNA breaks typical of the regions of tandem repeats [33].

In conclusion, it should be noted that the GS sequence is a tandem repeat with monomer size identical to that of the known centromeric 190-bp repeat found in genomes of the representatives of many avian families [4]. However, the GS sequence lacks conservative regions typical of the centromeric repeat, has non-centromeric location, and is species-specific.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (grant nos. 00-04-55013 and 02-00-49116), the Russian Ministry of Education (grant no. E00-6.0-218), the Russian State Program “Russian Universities” (UR.07.01.010), and the Rus-

sian Ministry of Industry, Science, and Technologies (The Program of Support for Unique Instruments, KOMUS-1, state registration no. 06-03).

REFERENCES

1. Mizuno, S. and Macgregor, H., The ZW Lampbrush Chromosomes of Birds: A Unique Opportunity to Look at the Molecular Cytogenetics of Sex Chromosomes, *Cytogenet. Cell Genet.*, 1998, vol. 80, pp. 149–157.
2. Solovei, I.V., Joffe, B.I., Gaginskaya, E.R., and Macgregor, H.C., Transcription on Lampbrush Chromosomes of a Centromerically Localized Highly Repeated DNA in Pigeon (*Columba*) Relates to Sequence Arrangement, *Chromosome Res.*, 1996, vol. 4, pp. 588–603.
3. Saifitdinova, A.F., Derjusheva, S.E., Malykh, A.G., *et al.*, Centromeric Tandem Repeat from the Chaffinch Genome: Isolation and Molecular Characterization, *Genome*, 2001, vol. 44, pp. 96–103.
4. Madsen, C.S., Brooks, J.E., de Kloet, E., and de Kloet, S.R., Sequence Conservation of an Avian Centromeric Repeated DNA Component, *Genome*, 1994, vol. 37, pp. 351–355.
5. Matzke, M.A., Varga, F., Berger, H., *et al.*, A 41–42 bp Tandemly Repeated Sequence Isolated from Nuclear Envelopes of Chicken Erythrocytes Is Located Predominantly on Microchromosomes, *Chromosoma*, 1990, vol. 99, pp. 131–137.
6. Matzke, A.J.M., Varga, F., Gruendler, P., *et al.*, Characterization of a New Repetitive Sequence That Is Enriched on Microchromosomes of Turkey, *Chromosoma*, 1992, vol. 102, pp. 9–14.
7. Tanaka, K., Suzuki, T., Nojiri, T., *et al.*, Characterization and Chromosomal Distribution of a Novel Satellite DNA Sequence of Japanese Quail (*Coturnix coturnix japonica*), *J. Hered.*, 2000, vol. 91, pp. 412–415.
8. Evteev, A.V., Andreeva, T.F., and Gaginskaya, E.R., Transcriptional Analysis of the CNM Subtelomeric Repetitive Sequence in Chicken Microbivalents during Oogenesis, *Tsitologiya*, 2000, vol. 42, pp. 278–279.
9. Zhu, X., Skinner, J., and Burgoyne, L.A., The Macrosatellites of the Toulouse Goose: The Major Tandemly Repetitive DNA in the Toulouse Goose Genome, *Genome*, 1990, vol. 33, pp. 641–645.
10. Zhu, X., Skinner, J., and Burgoyne, L.A., The Interspersed DNA Repeat Conserved in Water Birds’ Genomes, *Genome*, 1991, vol. 34, pp. 493–494.
11. Zhu, X.Y., Testori, A., Oh, S., *et al.*, Avian Tandem Repeats: Nuclear Protein-Binding Sites and the Origins of Short Tandem Repetitive DNAs, *Biochem. Int.*, 1991, vol. 25, pp. 191–198.
12. Zhu, X., Testori, A., Oh, S., *et al.*, Avian Species-Specific Tandem Repeats Contain Nuclear Protein-Binding Sites, *Biochem. Biophys. Res. Commun.*, 1992, vol. 182, pp. 447–451.
13. Keyser, C., Montagnon, D., Schlee, M., *et al.*, First Isolation of Tandemly Repeated DNA Sequences in New-World Vultures and Phylogenetic Implications, *Genome*, 1996, vol. 39, pp. 31–39.
14. Buntjer, J.B. and Lenstra, J.A., Self-Amplification of Satellite DNA in vitro, *Genome*, 1998, vol. 41, pp. 429–434.

15. Tiersch, T.R. and Wachtel, S.S., On the Evolution of Genome Size of Birds, *J. Hered.*, 1991, vol. 82, pp. 363–368.
16. Malykh, A.G., Zhurov, V.G., Saifitdinova, A.F., *et al.*, Structural and Functional Characterization of a Centromeric Highly Repetitive DNA Sequence of *Fringilla coelebs* L. (Aves: Passeriformes), *Mol. Biol. (Moscow)*, 2001, vol. 35, no. 3, pp. 331–335.
17. Altschul, S.F., Madden, T.L., Schaffer, A.A., *et al.*, Gapped BLAST and PSI-BLAST: A New Generation of Protein Database Search Programs, *Nucleic Acids Res.*, 1997, vol. 25, pp. 3389–3402.
18. Lichter, P. and Cremer, T., Chromosome Analysis by Non-Isotopic *in situ* Hybridization, *Human Cytogenetics: A Practical Approach*, Oxford: IRL, 1992, pp. 157–192.
19. Toth, G., Gaspari, Z., and Jurka, J., Microsatellites in Different Eukaryotic Genomes: Survey and Analysis, *Genome Res.*, 2000, vol. 10, pp. 967–981.
20. Garrido-Ramos, M.A., de la Herran, R., Ruiz Rejon, C., and Ruiz Rejon, M., A Satellite DNA of the Sparidae Family (Pisces, Perciformes) Associated with Telomeric Sequences, *Cytogenet. Cell Genet.*, 1998, vol. 83, pp. 3–9.
21. Rodionov, A.V., GT-Reach Chromomycin-Positive (CMA/DA) Blocks in the Karyotypes of Chicken *Gallus gallus domesticus*: An Ideogram and Intraspecific Pattern Variation, *Genetika (Moscow)*, 1998, vol. 34, no. 11, pp. 1523–1527.
22. Meyne, J., Baker, R.J., Hobart, H.H., *et al.*, Distribution of Non-Telomeric Sites of the (TTAGGG)_n Telomeric Sequence in Vertebrate Chromosomes, *Chromosoma*, 1990, vol. 99, pp. 3–10.
23. Fagundes, V., Vianna-Morgante, A.M., and Yonenaga-Yassuda, T., Telomeric Sequence Localization and G-Banding Patterns in the Identification of a Polymorphic Chromosomal Rearrangement in the Rodent *Akodon cursor* (2n = 14, 15, and 16), *Chromosome Res.*, 1997, vol. 5, pp. 228–232.
24. Garagna, S., Ronchetti, E., Mascheretti, S., *et al.*, Non-Telomeric Chromosome Localization of (TTAGGG)_n Repeats in the Genus *Eulemur*, *Chromosome Res.*, 1997, vol. 5, pp. 487–491.
25. Ijdo, J.W., Baldini, A., Ward, D.C., *et al.*, Origin of Human Chromosome 2: An Ancestral Telomere–Telomere Fusion, *Proc. Natl. Acad. Sci. USA*, 1991, vol. 88, pp. 9051–9055.
26. Lee, C., Sasi, R., and Lin, C.C., Interstitial Localization of Telomeric DNA Sequences in the Indian Muntjac Chromosomes: Further Evidence for Tandem Chromosome Fusions in the Karyotypic Evolution of the Asian Muntjacs, *Cytogenet. Cell Genet.*, 1993, vol. 63, pp. 156–159.
27. Vermeesch, J.R., De Meurichy, W., Van den Berghe, H., *et al.*, Differences in the Distribution and Nature of the Interstitial Telomeric (TTAGGG)_n Sequences in the Chromosomes of the Giraffidae, Okapi (*Okapia johnstoni*), and Giraffe (*Giraffa camelopardalis*): Evidence for Ancestral Telomeres at the Okapi Polymorphic rob (5; 26) Fusion Site, *Cytogenet. Cell Genet.*, 1996, vol. 72, pp. 310–315.
28. Salvadori, S., Deiana, A., Coluccia, E., *et al.*, Colocalization of (TTAGGG)_n Telomeric Sequences and Ribosomal Genes in Atlantic Eels, *Chromosome Res.*, 1995, vol. 3, pp. 54–58.
29. Liu, W.-S. and Fredga, K., Telomeric (TTAGGG)_n Sequences Are Associated with Nucleolus Organizer Regions (NORs) in the Wood Lemming, *Chromosome Res.*, 1999, vol. 7, pp. 235–240.
30. Pagnozzi, J.M., de Jesus Silva, M.J., and Yonenaga-Yassuda, Y., Intraspecific Variation in the Distribution of the Interstitial Telomeric (TTAGGG)_n Sequences in *Microureus demerarae* (Marsupialia: Didelphidae), *Chromosome Res.*, 2000, vol. 8, pp. 585–591.
31. Faravelli, M., Moralli, D., Bertoni, L., *et al.*, Two Extended Arrays of a Satellite DNA Sequence at the Centromere and at the Short-Arm Telomere of Chinese Hamster Chromosome 5, *Cytogenet. Cell Genet.*, 1998, vol. 83, pp. 281–286.
32. Fanning, T.G., Origin and Evolution of a Major Feline Satellite DNA, *J. Mol. Biol.*, 1987, vol. 197, pp. 627–634.
33. Gruzdev, A.D., Heterochromatin and Single-Strand DNA Breaks: A Hypothesis, *Genetika (Moscow)*, 1999, vol. 35, no. 7, pp. 869–872.