# The Conserved Fraction of Repetitive Sequences in the Human and Mammalian Genome in Health, in Pathology, and for Long-Term Mutagenic Exposure

**A. I. Potapenko, G. A. Khudolii, and A. P. Akif'ev**

UDC 616-055.5/.7-092:575.224.23]-02-092.9-07

Translated from *Byulleten' Eksperimental'noi Biologii i Meditsiny*, Vol. 117, № 1, pp. 85-88, January, 1994
Original article submitted July 14, 1993

The percentage and the size distribution of weakly divergent sites of repetitive sequences of human chromosomal DNA were determined in health and in hereditary genome diseases which are characterized by chromosome instability and predisposition to malignant tumor.

**Key Words:** *DNA; repetitive sequences; homology; nuclease S1*

The presence of an appreciable (20-80%) fraction of repetitive sequences is typical of the eukaryotic genome. The functions of the majority of repeats (except for genes encoding ribosomal RNA, tRNA, and histones) have not been elucidated even though they have been intensively studied for 25 years. For a number of reasons, specifically due to its comparatively slow pace, sequencing of the human genome cannot yet solve the problem as to the function of repeats.

We think that the biological significance of repetitive sequences may be indirectly judged by the homology among the members of individual families. When a pronounced homology is absent, and the repeats themselves are "trash" in the genome [9], the sequences comprising the new family of repeats start to diverge immediately after the new family "emerges" and are eventually eliminated from the fraction of repeats. On the other hand, when the homology within families, notably within sequences interspersed over the genome, is preserved, we are justified in wondering what its functional significance is and are prompted to begin probing the genetic process which maintains this homology. Currently, just R repeats, which are involved in the process of mitotic crossing-over, have been described from this standpoint [7], and the reason that they preserve an extremely high homology within them has been explained [1].

In the present study we aimed to determine the share and the size distribution of precise, i.e., slightly divergent, sites of repetitive sequences of human chromosomal DNA in health and in hereditary genome diseases, which are characterized by chromosome instability and predisposition to malignant tumors, as well as of DNA from hepatocytes and murine Ehrlich ascitic carcinoma (EAC).

## MATERIALS AND METHODS

Total cellular DNA was isolated by the phenol-detergent method [8] from leukocytes of healthy persons and patients with Down, Recklinghausen, Louis-Barr, and Fanconi's syndromes and tuberous sclerosis, from hepatocytes and brain cells of young and old mice, as well as from EAC cells, either intact or subjected to long-term mutagenic exposure to γ-radiation and thiophosphamide [4]. The

Department of Chemical and Biological Processes, Institute of Chemical Physics, Russian Academy of Sciences, Moscow. (Presented by A. I. Archakov, Member of the Russian Academy of Medical Sciences)

0007-4888/94/0001-0089$12.50 

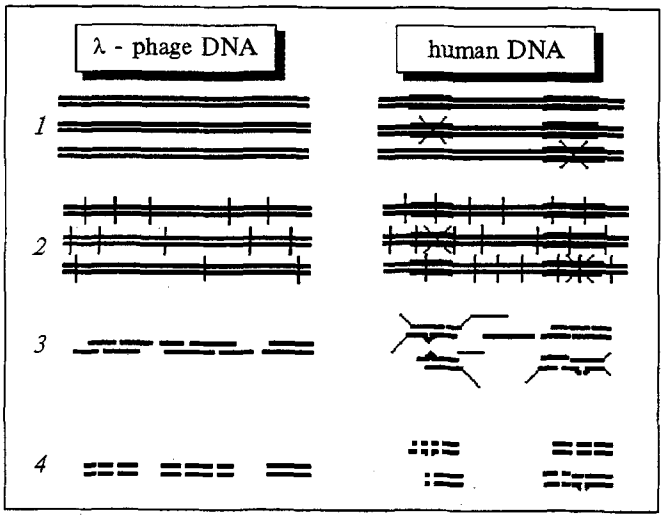


**Fig. 1.** Schematic of experiment. *1*) initial; *2*) after sonication; *3*) after denaturing and renaturing; *4*) after digestion with S1 nuclease. X denotes differences between the repeats within an individual family, leading to disturbances of secondary structure of DNA after renaturing.

DNA preparations were thrice sheared by sonication (UZDN-A) for one minute at one-minute intervals to attain a mean molecular weight of 300-500 base pairs (bp). The DNA fraction with a mode size from 200 to 1000 bp was isolated by means of preparative electrophoresis in 1.5% agarose gel (100 V, 1.5 h) [8]. This procedure enabled us to eliminate small DNA fragments, which would make it difficult to perform the subsequent analysis in the experiments with renaturation.

The size-fractionated DNA samples were denatured and renatured to cot=40 under standard conditions for isolation of the total fraction of repeats [5]. Following extensive digestion with nuclease S1 (DIG=0.99) [5], the remaining double-stranded fragments of repetitive sequences of DNA were precipitated with ethanol and analyzed by gel electrophoresis. Beforehand, conditions were chosen such that nuclease S1 recognized the unpaired nucleotide sequences and single-stranded DNA breaks and transformed them into double-stranded breaks [2]. This enzyme is known to recognize, under similar conditions, the single-stranded sequences characterized by mismatching in 2 bp [6]. The samples of EAC DNA were investigated by the method of thermal elution from hydroxyapatite (HAP) [3] after sonication, renaturation, and S1 digestion.

In different stages of the experiments, the mode size of the fragments with respect to molecular weight was determined in all DNA preparations by electrophoresis in 12% polyacrylamide gel [8] (2 h, 100 V + 16 h, 40 V). The λ-phage

**TABLE 1.** Characteristics of DNA Samples after Sonication (I) and after Denaturing, Renaturing, and S1 Digestion (II)

| Cellular DNA | Mean-numerical molecular weight | | Share of DNA with molecular weight | | | |
|---|---|---|---|---|---|---|
| | | | lower than 70 bp | | lower than 200 bp | |
| | I | II | I | II | I | II |
| *DNA of EAC and mice* | | | | | | |
| EAC | 274 | 110 | 4 | 16 | 17 | 53 |
| EAC (thiophosphamide) | 305 | 107 | 0 | 15 | 10 | 51 |
| EAC (γ) | 284 | 154 | 0 | 5 | 15 | 41 |
| from liver (y) | 285 | 118 | 0 | 10 | 16 | 55 |
| from liver (s) | 301 | 110 | 0 | 13 | 8 | 47 |
| from liver (s) | 264 | 120 | 3 | 11 | 15 | 43 |
| from brain (y) | 423 | 146 | 0 | 8 | 12 | 36 |
| from brain (s) | 610 | 134 | 0 | 10 | 0 | 38 |
| from λ−phage | 416 | 228 | 0 | 1 | 0 | 21 |
| *DNA of human lymphocytes* | | | | | | |
| Control I | 532 | 87 | 0 | 23 | 0 | 49 |
| Control II | 542 | 17 | 0 | 8 | 0 | 31 |
| Control III | 747 | 133 | 0 | 11 | 0 | 45 |
| Louis−Barr | 95 | 67 | 17 | 34 | 58 | 66 |
| Down (I) | 407 | 129 | 0 | 10 | 0 | 36 |
| Down (II) | 508 | 162 | 0 | 17 | 8 | 35 |
| Fanconi syndrome | 462 | 123 | 0 | 13 | 1 | 42 |
| Fanconi syndrome* | 506 | 125 | 0 | 14 | 0 | 40 |
| Recklinghausen I* | 441 | 90 | 0 | 21 | 2 | 57 |
| Recklinghausen II* | 268 | 99 | 6 | 18 | 9 | 55 |
| Tuberous sclerosis* | 455 | 104 | 0 | 19 | 0 | 58 |

**Note.** DNA of young (2 months) and senescent (31 months) animals are denoted as y and s, respectively, in parentheses. Asterisk: heterozygote carriers.

DNA, treated in accordance with the same scheme as other samples, was used as the control yielding ideal duplexes after annealing. Figure 1 is schematic of the experiments; typical molecular weight distributions of the fragments are presented in Fig. 2.

## RESULTS

The majority of families of repetitive sequences in the genomes of higher eukaryotes are known to exhibit a high degree of divergence. Specifically, the human Alu family has a 10-15% divergence. The mean divergence within the fraction of repeats in EAC cells was estimated by thermoelution from HAP, in order to assess the difference between the melting temperatures of the murine DNA renatured to cot=40 ($T°_m$=80°C) and of the native DNA sheared by sonication to a mode size of 3000 bp ($T°_m$=80°C). This difference constituted 6°C. Taking into account that the length of the l-phage DNA changed after annealing and S1 digestion (Table 1), the melting point of repetitive duplexes must presumably drop 1.5°C due to the length reduction. A 1% divergence is known [5] to reduce the melting point of duplexes by 1°C, on average, and it is likely that the fraction of repeats in the murine genome is 4-6% divergent. The melting point of the fraction of repeats after extensive digestion with S1 is 84°C. Its difference from the $T°_m$ of the fragmented native DNA entirely depends on the reduction of the mean length of the fragments (Table 1). This is evidence that just ideal duplexes remain double-stranded after S1 digestion. If this is so, electrophoresis of the fraction of repeats in polyacrylamide gel after annealing and S1 digestion (Fig. 2, *c*) will yield the length distribution of distances between the neighboring mutant substitutions.

When mutational events randomly appear within the repeats, their number over a given length $X$ is described by Poisson's distribution with the mean number of events per nucleotide ($D$), which, in our case, is the same as the share of unpaired bases ($D$=0.05) in the fraction of repeats. The distance between the neighboring alterations is a random value with the probability density:

$$f(X) = D \times e^{-Dx}$$

and with the expectation $1/D$. Therefore, at a 5% divergence within the fraction of repeats the mean length of DNA fragments after annealing and digestion with S1 nuclease must be $1/D$=20 bp.

As is seen from Table 1, the actual values are several times higher. This apparently indicates that certain regions of repeats are controlled by genetic

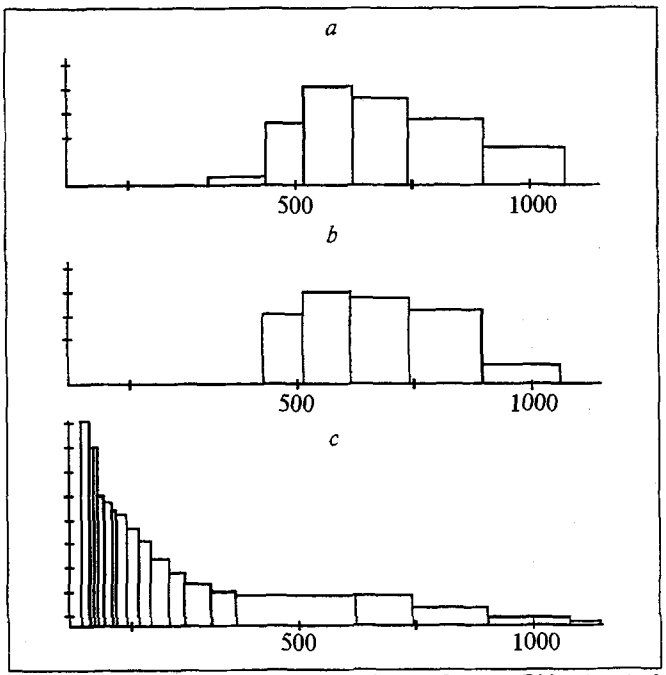


Fig. 2. Molecular weight distribution of human DNA (control I) after sonication (*a*), sonication and S1 digestion (*b*), and sonication, renaturing, and S1 digestion (*c*). Abscissa: molecular weight, bp; ordinate: share of DNA.

mechanisms, and mutational events are nonrandomly distributed over the length of the repeats.

Table 1 also shows additional characteristics of the actual distributions. For instance, the share estimate of DNA fragments shorter than 70 and 200 bp constitutes 97 and 99.99%, respectively. As is seen from the table, the actual values are several times lower than the theoretically predicted ones.

Thus, our findings provide evidence that not only individual families of repetitive sequences (R repeats) are conserved in the course of evolution, but also the internal regions of the majority of repeats. In the human genome the fraction of precise duplexes, after annealing to cot=40 and extensive digestion with S1 nuclease, amounts to 12%.

Table 1 shows that the parameters of the length distributions of precise duplexes in the fraction of repeats are unaltered for a number of hereditary diseases which are characterized by an increased incidence of spontaneous mutations and predisposition to malignant tumors. This also attests to the important role of the described regions of the repeats in maintaining the vital activity of cells.

The experiments previously carried out in our laboratory demonstrated that about 50% of the fraction of human precise repeats yielded ideal duplexes after hybridization with murine repeats [4]. We also obtained tentative results showing that in the bee (*Apis mellifera*) genome at the larval stage the same sequences are detected as in the fraction of human

precise duplexes. Therefore, at least some of the highly conservative repeats persist for an extremely long time in the eukaryotic genome.

The next task is to create a clone library of the conservative fraction of repetitive sequences of the human genome and to analyze in this fraction the presence of repeats common to the genomes of diverse eukaryotic species.

The authors would like to express their profound gratitude to their assistants T. A. Ketova and S. S. Degtyareva.

The research was partially supported by the Russian Federation "Human Genome" Research Project and "Priority Trends in Genetics" Program.

## REFERENCES

1. A. P. Akif'ev and A. I. Potapenko, *Izv. Sib. Otd. Akad. Nauk SSSR*, № 2, 29 (1990).
2. A. I. Potapenko and L. K. Obukhova, *Izv. Ros. Akad. Nauk, Ser. Biol.*, № 6, 937 (1992).
3. G. A. Khudolii, Kh. A. Khakimov, T. V. Gorelova, and A. P. Akif'ev, *Dokl. Akad. Nauk SSSR*, **282**, 996 (1985).
4. G. A. Khudolii and A. P. Akif'ev, *Dokl. Ros. Akad. Nauk*, **329**, № 3, 372 (1993).
5. R. J. Britten *et al.*, *J. Mol. Evol.*, **9**, 1 (1976).
6. M. C. Burdon and S. H. Lees, *Bioscience Reports*, **5**, 627 (1985).
7. H. Stern, *Chromosoma*, **89**, 127 (1984).
8. N. Maniatis *et al.*, *Molecular Cloning*, Cold Spring Harbor Laboratory (1982).
9. E. Zuckerland, *J. Mol. Evol.*, **34**, 259 (1992).

# EXPERIMENTAL BIOLOGY

# Individual Differences in Responses to Acute Stress Associated with Type of Behavior (Prediction of Stress Resistance)

**K. Yu. Sarkisova and M. A. Kulikov**

UDC 612.821+591.51+616-008.61

Translated from *Byulleten' Eksperimental'noi Biologii i Meditsiny*, Vol. 117, № 1, pp. 89-92, January, 1994
Original article submitted September 13, 1993

On the basis of the initial parameters of behavior in the "open field," "forced swimming," and "emotional resonance" tests, the main behavioral parameters - the number of squares crossed, the number of standing postures, and the time of passive swimming - are shown to be predictable for stress in rats with different types of behavior.

**Key Words:** *acute stress; specificities of behavior; prediction; stress resistance*

Prediction of resistance to stress and to other pathogenic effects, based on specificities of behavior, is of not only theoretical, but also practical importance. Specifically, this makes it possible to preliminarily choose individuals with different degrees of resistance to certain factors, which is necessary both for studying the mechanisms of individual resistance to these factors and for developing methods of individual prophylaxis and treatment of disturbances caused by these factors. For instance, a correlation between the resistance of the cardiovascular system to emotional stress and

Group of Experimental Pathology and Therapy of Higher Nervous Activity, Laboratory of Mathematical Neurobiology of Learning, Institute of Higher Nervous Activity and Neurophysiology, Russian Academy of Medical Sciences, Moscow. (Presented by P. V. Simonov, Member of the Russian Academy of Medical Sciences)

0007-4888/94/0001-0092$12.50 