**ABSTRACT**

A program for constructing nucleosome formation potential profile was applied for investigation of exons, introns, and repetitive sequences. The program is available at http://wwwmgs.bionet.nsc.ru/mgs/programs/recon/. We have demonstrated that introns and repetitive sequences exhibit higher nucleosome formation potentials than exons. This fact may be explained by functional saturation of exons with genetic code, hindering the localisation of efficient nucleosome positioning sites.

**INTRODUCTION**

Chromatin organization into nucleosomes is the basal process of DNA packaging in eukaryotes. Many nucleosome positioning specificities are related to the functional patterns of genomic sequences (Trifonov, 1997). There is a suggested relationship between nucleosome packaging of DNA and intron–exon organization of the eukaryotic genes (Denisov et al., 1997; Levitsky et al., 1999). A role in regulation of nucleosome positioning is also ascribed to repetitive DNA sequences (Fitzgerald et al., 1994; Englander and Howard, 1995; Thastrom et al., 1999). We have applied a method for nucleosome formation potential construction (Levitsky et al., 2001) to studying exons, introns, and repetitive sequences. The sequence sets analyzed are characterized in Table 1. The fragments of human exons and introns (sequences of 160 bp in length) and human donor and acceptor splice sites (([-200; +200]) relative to the site centre) were extracted from EMBL. The human Alu repeats sequences (with the length between 160–1956 bp) were extracted from the database REPBASE ftp://ftp.ebi.ac.uk/pub/databases/repbase.

**RESULTS**

Figure 1a shows the nucleosome formation potential \( \phi(X) \) distributions for sets of human exon and intron sequences and nucleosome sites. The distribution for exons is shifted leftward relative to that of introns and has the pronounced tail at its left flank. The intron distribution is close to that of nucleosome sites. The \( \phi(X) \) profiles of the donor and acceptor splice sites (Figure 1b) demonstrate that \( \phi(X) \) values for exons are essentially more distant from unity than that for introns. Typical of donor and acceptor splice site regions are linear trends of \( \phi(X) \)-positive from exon to intron and negative from intron to exon, correspondingly.

The nucleosome positioning sites detected in the neighbourhood of splice sites support the hypothesis on evolutionary rationale of the origin of introns (Solovyev and Kolchanov, 1985; Csordas, 1989). According to it, the problem of DNA packaging of a gene lacking the nucleosome positioning signal due to restrictions stemming from the primary structure of the protein can be solved through inserting an intron carrying a nucleosome positioning signal. Conceivably, the nucleosome code is determined by a specific pattern of alternating DNA regions with differing nucleosome formation potentials. The regions with high potential are localized predominantly to introns, whereas those with low potential, to exons. As for 5′ gene regulatory regions, localization of the low-potential regions is determined by the gene expression pattern and remoteness from the transcription start (Levitsky et al., 2001). These conclusions comply with the concept of mosaic arrangement for chromatin organization (Liu and Stein, 1997), experimental data (Liu et al., 1995;...
Nucleosome formation potential of exons, introns, and Alu repeats

Fig. 1. (a) Histogram of nucleosome formation potential \( \varphi(X) \) distributions for introns, exons and Alu repeats (for comparison, see the \( \varphi(X) \) distribution for the set of nucleosome sites); (b) nucleosome formation potential \( \varphi(X) \) profiles of human donor and acceptor splice sites.

Lauderdale and Stein, 1992), and other observations (Denisov et al., 1997; Levitsky et al., 1999).

The nucleosome formation potential distribution of human Alu repeats is shown in Figure 1a together with the distributions of the sets of nucleosome sites, introns, and exons. It is evident that the distribution of Alu repeats is similar to those of introns and nucleosome sites and differs significantly from that of exons. The other analyzed families of dispersed repeats displayed similar patterns (data not shown). Note in this connection the experimental data (Englander and Howard, 1995) on the ability of Alu repeats to assist nucleosome rotational positioning in the genomic DNA. Both this ability and the abundance of Alu repeats (about 10^6 per genome) suggest that the arrangement of nucleosome positioning signals is among their functions. The available data indicate that the satellite DNA also possess an increased nucleosome formation potential (Fitzgerald et al., 1994; Thastrom et al., 1999).
ACKNOWLEDGEMENTS
The work was supported by the Russian Foundation for Basic Research, Russian National Human Genome Program, Integrated Program of Siberian Branch of the Russian Academy of Sciences. The authors thank A.V.Katokhin for useful discussions and to G.B.Chirikova for paper translating.

REFERENCES