Comparative restriction enzymes analysis of rat chromosomal DNA in vitro and in silico

Category: Study of Genome DNA
Hits: 24585

V.A. Chernukhin, M.A. Abdurashitov, V.N. Tomilov, D.A. Gonchar, S.Kh. Degtyarev
Translated from "Ovchinnikov bulletin of biotechnology and physical and chemical biology" V.2, No 3, pp 39-46, 2006

Theoretical diagrams of rat chromosomal DNA cleavage at more than 25 different nucleotide sequences have been plotted based on earlier proposed method of restriction enzymes analysis of mammalian genomes in silico. A map of the rat LINE1 repeats digestions at the same nucleotide sequences has been calculated and a correspondence between the diagrams data and LINE1 repeat cleavage positions has been shown. Restriction enzymes analysis in vitro has been performed in order to obtain patterns of rat chromosomal DNA cleavage by endonucleases with respective recognition sequences. A perfect coincidence has been shown between theoretical DNA cleavage diagrams and experimental patterns of DNA hydrolysis with restriction endonucleases.

More than 90% of the rat chromosomal DNA primary structure has been already determined [1], and these data are constantly updated. In our previous work, we proposed a simple method of restriction enzymes analysis of mammalian DNA *in silico* by construction the distribution diagrams of DNA fragments produced by cleavage of chromosomal DNA at recognition sites of restriction endonucleases. Such cleavage diagrams of rat, mouse and human chromosomal DNAs have been plotted for DNA sequences 5'-CCWGG-3', 5'-GATC-3', 5'-GGCC-3' and 5'-CCGG-3'. The comparison of theoretical calculations with experimental results on chromosomal DNA hydrolysis with restriction endonucleases *Bst2U*I, *Kzo9*I, *Hae*III and *Msp*I which have these recognition sites demonstrated a good correspondence between restriction patterns *in vitro* and *in silico* [2]. In the present paper we study rat chromosomal DNA cleavage at a wider range of restriction endonucleases recognition sites.
The goal of the present work is to construct the fragments distribution diagrams produced by cleavage of rat chromosomal DNA at more than 25 recognition sites, including 4-, 5- and 6-nucleotide sequences, and to compare the calculation data with the results of DNA hydrolysis with respective restriction endonucleases.

MATERIALS AND METHODS

Rat chromosomal DNA

Male Sprague-Dowley rats aged 3-4 months (Breeding Laboratory of Experimental Animals, Institute of Cytology and Genetics, Novosibirsk) were used in the experiments. Genomic DNA from the animal liver was isolated according to [3] with some modifications described below.

Rats were decapitated, liver dissected, and tissue immediately frozen in liquid nitrogen. Pieces of tissue were grinded to a powder under liquid nitrogen using a porcelain mortar and pestle. 0.5 - 1 g of powder were placed into 50 ml cone flask with 5 ml of lysis buffer (100 mM Tris HCl pH 8.0, 10 mM EDTA; 0,5% SDS; 40 µg/ml Proteinase K). The mixture was agitated gently and incubated at 37°C for 3 hrs. 5 ml of phenol (saturated with 10mM Tris HCl pH 8,0) were added. The mixture was incubated for 15 min at room temperature with constant agitation and then centrifuged at 4000 rpm for 5 min at 20°C. The upper aqueous phase was transferred into a cone flask with a pipette using a scissored tip, extracted with a mix of 2,5 ml phenol (saturated with 10mM Tris HCl, pH 8,0) and 2,5 ml chloroform/isoamyl alcohol (24:1) at room temperature for 15 min with constant agitation and centrifuged at 4000 rpm for 5 min at 20°C. The aqueous phase was collected and extracted with 5 ml chloroform/isoamyl alcohol (24:1) as described above with subsequent centrifugation. Then aqueous phase was placed into 50 ml glass and cooled on ice. At slow stirring, isopropanol cooled to 0°C was added drop-wise and DNA precipitated spooled on glass rod, washed in 70% ethanol, gently dried with sterile air (3-6 min) and had been dissolving in 1-1,5 ml of sterile TE buffer (10mM Tris HCl pH 7,0; 1mM EDTA) 24 hours at 4°C.

Chromosomal DNA hydrolysis

The following restriction endonucleases manufactured by SibEnzyme Ltd. were used in the work (the recognition site of respective restriction enzymes is given in the brackets):
1) *Pvu*II (5'-CAGCTG-3'), 2) *Vsp*I (5'-ATTAAT-3'), 3) *Pci*I (5'-ACATGT-3'), 4) *Ksp22*I (5'-TGATCA-3'), 5) *Hind*III (5'-AAGCTT-3'), 6) *EcoR*I (5'-GAATTC-3'), 7) *Bpu10*I (5'-CCTNAGC-3'and 5'-GCTNAGG-3'), 8) *Pct*I (5'-GAATGC-3' and 5'-GCATTC-3'), 9) *Bse3D*I (5'-GCAATG-3'and 5'-CATTGC-3'), 10) *SfaN*I (5'-GCATC-3' and 5'-GATGC-3'), 11) *Bse21*I (5'-CCTNAGG-3'), 12) *BstV2*I (5'-GAAGAC-3' and 5'-GTCTTC-3'), 13) *Bst6*I (5'-CTCTTC-3'and 5'-GAAGAG-3'), 14) *Msp*I (5'-CCGG-3'), 15) *Alu*I (5'- AGCT-3'), 16) *Kzo9*I 5'-GATC-3'), 17) *Taq*I (5'-TCGA-3'), 18) *Rsa*I (5'-GTAC-3'), 19) *Fat*I (5'-CATG-3'), 20) *Tru9*I (5'-TTAA-3'), 21) *Sse9*I (5'-AATT-3'), 22) *BstDE*I (5'-CTNAG-3'), 23) *Hae*III (5'-GGCC-3'), 24) *AspS9*I (5'-GGNCC-3'), 25) *Fsp4H*I (5'-GCNGC-3'), 26) *BspAC*I (5'-CCGC-3' and 5'-GCGG-3'), 27) *Hpa*II (5'-CCGG-3'), 28) *HspA*I (5'-GCGC-3'), 29) *BstFN*I (5'-CGCG-3')
Before hydrolysis reaction, all DNA preparations were treated with ribonuclease A (0.1 mg/ml) for 10 minutes at a room temperature and dialyzed in DispoDialyzer MWCO 50,000 tubes ("Sigma", USA) TE buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA) 100 times of the volume of single DNA preparation at 4°C for 20 hours.
Hydrolysis reactions were performed in 40 µl of the reaction mixture containing 6 µg of DNA, SE-buffers recommended by the manufacturer and 3 µl of restriction enzyme at optimal temperatures for 3 h.

Electrophoresis

Electrophoresis in 8% polyacrylamide gel was used to separate DNA fragments from 40 to 500 bp (6 μg of hydrolyzed DNA was applied on gel in each run); 1.5% of low melting point agarose ("Sigma", USA) was used to separate DNA fragments in the range of 200-2000 bp. Electrophoresis in 1% agarose "Type I-A, Low EEO" ("Sigma", USA) was used to separate DNA fragments of higher molecular weight. 3 μg of hydrolyzed DNA was applied on agarose gel in each run. Tris-acetate buffer was used for electrophoresis in all the cases. After electrophoresis DNA bands were stained with ethidium bromide and photographed in UV light.

Data on the nucleotide sequence

The rat DNA sequence was obtained from the resource ftp://ftp.ensembl.org/pub/ (version of June 2, 2006).

Software

The DNA fragments distribution diagrams were represented in the form of the dependence of the total mass of fragments with fixed length (in base pairs) on the fragments lengths (in base pairs) according to the previously described technique [1].

RESULTS AND DISCUSSION

Previously we have presented fragments distribution diagrams of rat chromosomal DNA cleavage at recognition sites of restriction enzymes *Hae*III (5'-GGCC-3'), *Msp*I (5'-CCGG-3'), *Kzo9*I (5'-GATC-3') and *Bst2U*I (5'-CCWGG-3'). Good correlation between the calculation data and the results of experiments on DNA hydrolysis with these restriction enzymes has also been demonstrated [2]. In this work we plotted the fragments distribution diagrams of rat chromosomal DNA digestion at a wider range of recognition sites including six nucleotides sequences. In accordance with the earlier proposed method [2], the distribution diagrams were constructed in the arctangensoid scale, which imitates the separation of the DNA fragments by electrophoresis in agarose gel. Fig.1a shows the diagrams of DNA cleavage mainly at six nucleotides sequences, which are recognition sites of both well-known restriction enzymes (*Hind*III and *EcoR*I) and ones rarely used (*SfaN*I and some others). These diagrams were selected based on the presence of peaks of 5.5 and more million base pairs, a threshold value for appearance of respective bands in the analysis of the products of DNA cleavage by electrophoresis in agarose gel [2]. Table 1 summarizes data on 26 such peak fragments. As Table 1 shows, in most cases the presence of peaks of such height is explained by the presence of several fragments with sizes differing by 1-2 nucleotides for which, in accordance with the proposed method [1], the total peak height value was calculated. The formation of clusters of such fragments with minimal differences in the sizes is probably associated with deletions and insertions in DNA repeats [4]. Cleavage of these DNA repeats results in formation of the bands observed on agarose and acrylamide gels [2], [5, 6, 7]. Direct analysis

of DNA repeats seems to be a simpler method of obtaining a digestion pattern *in silico*, however, its use may be incorrect due to rather high variability of the nucleotide sequence in these DNA regions. For comparison, we carried out analysis of the averaged primary structure of the most complete (6914 bp) LINE1 repeat of rat DNA [8]. Data on the length of restriction fragments of this consensus LINE1 repeat, which corresponds to the data of the column 3, are presented in the fourth column of Table 1. As the Table shows, a sufficiently high correlation of genomic DNA and consensus LINE1 repeat cleavage is observed. However, for a number of sites, the calculation data obtained with our method differ from those of LINE1 repeat analysis. In particular, one of *Vsp*I sites was omitted in the latter case; a difference is observed in the sizes of *Pvu*II fragment and upper fragments of *Ksp22*I and *SfaN*I patterns of DNA digestion. This difference is associated with extra 3 bp and 133 bp length fragments, which are inserted in consensus LINE1 repeat fragment. A comparative study of consensus LINE1 repeat nucleotide sequence and some other ones [9] shows that these short DNA fragments indeed present in some LINE1 repeats, but an analysis of the whole DNA by our method as well as experimental data don't reveal these fragments. A revised version of consensus LINE1 repeat nucleotide sequence [10] includes 1) deletion of 133 bp fragment at positions 1294-1426, 2) deletion of 3 bp fragment at positions 1278-1280 and 3) replacement of A for T in position 1438 with formation of the second *Vsp*I site. Data on cleavage of new consensus LINE1 repeat proposed in this work [10] are given in 5th column. Comparison of 3rd and 5th columns shows a perfect coincidence of data provided and confirms a correct DNA sequence of proposed LINE1 repeat if compare to published one.

In general, however, the analysis of consensus LINE1 repeat does not allow to evaluate the number of obtained DNA fragments of the given length and thus to predict their visualization in the experiment. For example, low-molecular weight fragments in *EcoR*I, *SfaN*I, *BstV2*I and some other restriction enzymes digestions deduced from the analysis of LINE1 repeat structure give low peaks in the diagrams and, as it will be shown further, can't be seen in the gel photographs.

Restriction endonuclease	Recognition site	Lengths of the peak fragments*	Lengths of fragments* produced by cleavage of L1_RN [17]	Lengths of fragments* produced by cleavage of consensus L1 repeat (this work)
PvuII	CAGCTG	4549,4550	4685	**4549**
VspI	ATTAAT	3361,3362	-	**3362**
PciI	ACATGT	1171,1172	**1172**	**1172**
Ksp 22 I	TGATCA	2678,2679 2094,2095	2814 **2095**	**2678** **2095**
HindIII	AAGCTT	3626,3627 179	**3627** **179**	**3627** **179**
EcoRI	GAATTC	2316,2317,2318 1372,1373	**2318** **1373**	**2318** **1373**
Bpu 10 I	CCTNAGC	2116,2117	**2117**	**2117**
PctI	GAATGC	2105,2106 641, 619 413 236	**2106** **641** , **625** ** **413** **236**	**2106** **641, 625**** **413** **236**
Bse3DI	GCAATG	3908,3909	**3917** **	**3917****
SfaNI	GCATC	2786,2787 1151 437	2922 **1170** ** **418** **	**2786** **1170**** **418****
Bse21I	CCTNAGG	2140,2142 1960,1961,1962	**2142** **1962**	**2142** **1962**
BstV2I	GAAGAC	1279,1280 924 747 327	**1266** ** **924** **761** ** **327**	**1266**** **924** **761**** **327**
Bst 6 I	CTCTTC	2656,2657 2115,2116	**2657** **2116**	**2657** **2116**

Table 1. Predicted lengths of DNA fragments, which are produced by cleavage of rat genome, published and proposed consensus sequences of LINE1 repeat.

* - included only fragments less than 500 bp with peak values more than 4 millions bp and fragments more than 500 bp with peak values more than 5.5 millions bp

The lengths of fragments in fourth and fifth columns, which are the same as the length in third column, are shown in bold.

** - the length difference of indicated DNA fragments, which are produced by cleavage of IIS type restriction endonucleases PctI, Bse3DI, SfaNI and BstV2I, is caused by a location of the restriction enzymes cleavage positions outside recognition sequences

Fig.1b presents experimental data on rat chromosomal DNA cleavage with restriction endonucleases. The comparison of calculation data in the third column of Table 1 with experimental results shows that all the indicated fragments are present in the form of bands in agarose gel photographs.

Fig. 1.

a) Distribution diagrams of total DNA fragments lengths (in bp) depending on the fragment size for restriction endonucleases recognition sites. M1 - calculated DNA fragment length marker corresponding to SE 1 kb ladder. Detailed diagrams are shown on the Fig 4.
b) b) Electrophoregram of rat genomic DNA digestion with restriction endonucleases in 1% agarose gel. M1 - DNA fragment length marker SE 1 kb ladder.
c) c) Electrophoregram of rat genomic DNA digestion with restriction endonucleases in polyacrylamide gel. M2 - DNA fragment length marker pUC19/MspI.

In particular, the presence of bands of respective molecular weight at rat DNA hydrolysis with restriction enzymes *Pvu*II, *Vsp*I, *Pci*I, *Ksp22*I, *Hind*III, *EcoR*I, *Bpu10*I, *Pct*I, *Bse3D*I, SfaNI, *Bse21*I, *BstV2*I and *Bst6*I is experimentally observed. The patterns of chromosomal DNA hydrolysis with restriction enzymes *BamH*I and *Pce*I, respectively, are presented in Fig.1b on runs 4 and 15. The names of these enzymes are absent in Table 1, as cleavage at recognition sites *BamH*I and *Pce*I does not provide diagrams with high peak values. However, as it is seen from Figure 1b, DNA hydrolysis with these restriction enzymes also results in the appearance of DNA band of 4500-5000-bp length, which is especially noticeable in the case of *Pce*I. Besides, DNA hydrolysis with *Bst6*I provides an additional band of 4500-5000 base pairs length presented as a small peak in Fig.1a. Detailed analysis of DNA cleavage diagrams at recognition sites of restriction enzymes *BamH*I, *Bst6*I and *Pce*I shows that in all three cases several fragments, which differ in size in the range of 200 bp and have low peak values (data are not presented), are located in the region of 4500-5000 base pairs. The appearance of the band on the gel is probably associated with low resolution of electrophoresis for fragments of 4500-5000 base pairs length and can result from several small peaks overlapping. Another consequence of such overlap in peaks is the thickness of the upper band of DNA hydrolysis with *Ksp22*I which corresponds to 2678-2679-bp fragments length.
From Fig.1a and Table 1 it can be seen that in accordance with the obtained diagrams, the products of rat DNA cleavage with restriction enzymes *Hind*III, *Pct*I, *SfaN*I and *BstV2*I include low-molecular weight fragments. To analyze them, electrophoresis in PAAG was performed. This electrophoresis visualizes

fragments with a smaller peak value because a larger amount of DNA is loaded on the gel run. Fig.1c shows the patterns of DNA hydrolysis with restriction endonucleases *Hind*III, *Pct*I, *SfaN*I and *BstV2*I in PAAG. DNA cleavage with restriction enzyme *Hind*III results in the formation of the 179-bp fragment, which is present in Table 1, and an additional fragment of 370 base pairs. The last result is attributed to a hydrolysis of satellite DNA, which contains a recognition site of restriction enzyme *Hind*III [6]. In Fig.1c, one can also see all the other low molecular weight fragments, which are present in the third column of Table 1. *Pct*I forms a double of 641 and 619 as well as fragments of 413 and 236 base pairs, and *SfaN*I cleaves off the 437-bp fragment and *BstV2*I - the 327-bp fragment. The other low molecular weight fragments, revealed in the analysis of LINE1 repeat for recognition sites of restriction enzymes are not detected in the experiment. This fact does not count in favor of using the repeats data in construction of DNA cleavage diagrams.

Fig. 2.

a) Distribution diagrams of total DNA fragments lengths (in bp) depending on the fragment size for restriction endonucleases recognition sites. M1 - calculated DNA fragment length marker corresponding to SE 1 kb ladder. Detailed diagrams are shown on the Fig 5.
b) b) Electrophoregram of rat genomic DNA digestion with restriction endonucleases in 1% agarose gel. M1 - DNA fragment length marker SE 1 kb ladder.
c) c) Electrophoregrams of rat genomic DNA digestion with restriction endonucleases RsaI and AspS9I in polyacrylamide gel. M2 - DNA fragment length marker pUC19/MspI.

Fig.2a shows diagrams of DNA cleavage mainly at four nucleotides sequences; the numbers shown on them indicate the sizes of fragments with peak values of more than 5.5 million bp. As compared with diagrams in Fig.1, the formation of a considerably larger number of DNA fragments is observed in this case, which results in the appearance of the basic curve of the fragments distribution and peaks against this background. Data on these fragments peaks are collected and compared with corresponding results of simulation of LINE1 consensus fragment cleavage in Table 2.

Restriction endonuclease	Recognition site	Lengths of peak fragments	Lengths of fragments* formed by L1_RN [17] cleavage	Lengths of fragments* formed by cleavage of consensus L1 repeat sequence proposed in this work
MspI	CCGG	5537, 5538 404, 405	5673 **405**	**5537** **404**
AluI	AGCT	1127	**1127**	**1127**
Kzo9I	GATC	768 752 449, 447 337 321	**768** **752** **449** **337** **321**	**768** **752** **449** **337** **321**
TaqI	TCGA	1233 1187	**1233** **1187**	**1233** **1187**
RsaI	GTAC	1650 802 555 330 197	**1650** **802** **555** **330** **197**	**1650** **802** **555** **330** **197**
FatI	CATG	1007 832 648 634	**1007** **832** **648** **634**	**1007** **832** **648** **634**
Tru9I	TTAA	703 573, 574, 575 468 447	**703** **574** **468** **447**	**703** **573** **468** **447**
BstDEI	CTNAG	 624	**624**	 **624**
HaeIII	GGCC	1173 863 697 588 575	**1173** **863** **697** **588** **575**	**1173** **863** **697** **588** **575**

AspS9I	GGNCC	1110 1025 554 534 456 374 353 202 155	**1110** **1025** **554** **534** **456** **374** **353** **202** **155**	**1110** **1025** **554** **534** **456** **374** **353** **202** **155**
Fsp4HI	GCNGC	1613 1035 825 543	**1613** **1035** **825** **543**	**1613** **1035** **825** **543**
Bst2UI	CCWGG	1690, 1691 1017 713 462 341 313 291 215 177 146	**1691** **1017** **713** 465 **341** **313** **291** **215** **177** **146**	**1691** **1017** **713** **461** **341** **313** **291** **215** **177** **146**

Table 2. Comparison of predicted lengths of fragments formed due to cleavage of rat genome with several restriction enzymes with calculated lengths of fragments formed by cleavage of consensus sequences od LINE1-repeats.
* - only fragments <500 b.p. in length with peak values more than 4 billions b.p. and fragments >500 b.p. in length with peak values more than 5.5 billions b.p. are included.
Fragments in the fourth and fith columns which are of the same length as those in the third column are shown in bold.

The Table 2 data show that the peaks formation is associated with the cleavage of LINE1 repeats, whereas the presence of the basic curve is the result of cleavage of the rest of genomic DNA. The form of this curve varies depending on the analyzed nucleotide sequence, e.g., for recognition sites of restriction enzymes AluI, *Fat*I, *Tru9*I , *Sse9*I and *BstDE*I the number of fragments in the length range from 50 to 500 bp is considerably higher than for the others. Fig.2b presents the pattern of rat DNA hydrolysis with restriction endonucleases possessing recognition sites presented in Fig. 2a. The comparison of Figures 2a and 2b shows that the presence of the basic curve of the fragments distribution in the diagrams correlates with the location of DNA spot in the corresponding photographs of gels. The formation of such spot complicates the visualization of individual DNA fragments, which have peak values in the diagrams. Because of this, the numbers on the diagrams in Fig.2a indicate the sizes of only those fragments, which are outside these dark areas. A comparison of data presented in Figures 2a and 2b reveals a correspondence between more than 25 peak values of the fragments

lengths obtained in the diagrams and the bands seen in the gel photographs. In particular, as was shown previously [2], hydrolysis with restriction enzymes *Msp*I and *Hae*III reveals clearly distinguishable bands, which correspond to peaks of 5537+5538 and 404+405 bp in the first case, and 1173, 863, 697, 575+588 as well as 370 bp (satellite DNA) in the second case. The appearance of bands corresponding to the 1127-bp fragment, the double fragment (752 and 768 bp) and the double peak (1187 and 1233 bp) is clearly observed when DNA was treated with restriction enzymes AluI, *Kzo9*I and *Taq*I, respectively. DNA hydrolysis with enzymes *Rsa*I, *Fat*I and *Tru9*I provides several weak bands corresponding to the peaks at 1650, 989, 802 and 555 in the first case, 1007 and 832 in the second case and 703, 573-575 and 447-468 bp in the third case. DNA hydrolysis with restriction enzymes *Sse9*I and *BstDE*I does not provide clearly defined bands due to their overlapping with the basic curve of low molecular weight fragments. Cleavage with restriction enzyme *AspS9*I results in the appearance of discrete bands corresponding to the peaks at 1025+1100 bp and 534+554 bp, while restriction enzyme *Fsp4H*I provides the 1613 bp fragment and less noticeable fragments (1463, 1035, 825 and 543 bp). As seen in Fig.2a, low values of the basic curve in the region of less than 600 bp, which are required for visualization of fragments in PAAG, are observed in the case of DNA cleavage at recognition sites of restriction enzymes *AspS9*I, *Fsp4H*I, *Hae*III, *Kzo9*I, *Msp*I, *Rsa*I and *Taq*I. The results of electrophoresis of the products of rat chromosomal DNA cleavage with restriction enzymes *Hae*III, *Kzo9*I and *Msp*I in PAAG were discussed previously [2]. As for the other restriction enzymes the corresponding peaks exist in diagrams of DNA cleavage with enzymes *AspS9*I and *Rsa*I only. Fig.2c presents data of electrophoresis of the products of rat DNA cleavage with restriction enzymes *AspS9*I and *Rsa*I in PAAG. The first case reveals bands corresponding to the 456, 374, 353, 202 and 155 bp fragments and the second case shows bands which correspond to 581, 555 bp fragments plus an additional 370 bp fragment of satellite DNA [6].

From data presented in Fig.2b, it can be seen that DNA cleavage with restriction endonucleases *Taq*I, *Rsa*I and *Fat*I , which all have recognition sites with the same GC composition, results in different depth of hydrolysis and different location of the DNA spot. Correspondingly, basic curves for recognition sites of these restriction enzymes, as shown in Fig.2a, also considerably differ from each other. One can point out several peculiarities of rat chromosomal DNA, that influence the depth of hydrolysis by restriction endonucleases with the same set of nucleotides in the recognition site.

Fig. 3.
Rat genomic DNA digestion with restriction endonucleases containing CG-dinucleotide in their recognition sites. M1 - DNA fragment length marker SE 1 kb ladder.

1. ***Hindrance of genomic DNA hydrolysis due to CG-dinucleotides methylation.*** According to the available estimates, 70-80% of all CG pairs in mammalian genomes are methylated [11]. Fig.3 shows the results of DNA cleavage with restriction endonucleases containing dinucleotide CG in the recognition site, in particular, with restriction enzymes *BspAC*I : (5'-CCGC-3' and 5'GCGG-3'), *BstFN*I (5'-CGCG-3'), *HspA*I (5'-GCGC-3'), *Hpa*II and *Msp*I (5'-CCGG-3'). As seen in Figure 3, enzymes *BspAC*I , *BstFN*I, *HspA*I and *Hpa*II only slightly cleave genomic DNA, whereas *Msp*I, which is able to cleave the site 5'-CmCGG-3', hydrolyzes DNA very well. Thus, out of all analyzed restriction enzymes with CG dinucleotide in the recognition site, chromosomal DNA is effectively cleaved by *Msp*I and *Taq*I, with the latter also being insensitive to cytosine methylation in the recognition site (Fig. 2b and 3).
In addition, DNA hydrolysis with some restriction enzymes can be complicated by recognition sites overlapping with CG-methylated dinucleotide. When compared with the location of the basic curve in corresponding diagrams, the difference in dark area positions on photographs of DNA products separation after cleavage by *Hae*III and *Fsp4H*I may be explained by presence of 5mCG dinucleotide (Fig 2b). Part of recognition sites of restriction enzymes HinfI (5'-GANTC-3') and *EcoR*I (5'-GAATTC-3') in rat satellite DNA contains methylated cytosine and these sites are hydrolyzed at a considerably lower rate than the non-modified recognition sites [5].

2. ***GC composition.*** GC composition of rat genomic DNA is 40% [12] and restriction enzymes with AT recognition sites (*Tru9*I and *Sse9*I) should produce a deeper hydrolysis than restriction endonucleases which have GC pairs only in the recognition site (for example, *Hae*III and *AspS9*I). Data provided in Fig.2 confirm this consideration.

3. ***The presence of high-frequency and low-frequency dinucleotides in the recognition site.*** The occurrence frequency of some dinucleotide pairs considerably differs from the average statistical expected values. This can account for the difference in the average length of the restriction fragments at DNA hydrolysis with enzymes whose recognition sites have the same GC composition.
Thus, DNA hydrolysis with restriction endonuclease *Taq*I (recognition site TCGA) produce the area of chromosomal DNA fragments spot is considerably higher than that at hydrolysis with restriction enzymes Kzo9I (GATC) and *Rsa*I (GTAC) though GC composition of recognition sites of all these enzymes is the same.
This result can be accounted for by the fact that in chromosomal DNA of eukaryotes the occurrence frequency of dinucleotide CG is five times lower than the average statistical expected frequency [13]. Low occurrence frequency of the CG pair can be explained by a high rate of transformation of 5-methylcytosine in this dinucleotide into thymine which occurs as a result of the spontaneous deamination reaction [14].
Similarly, one can explain the difference in the depth of DNA hydrolysis with restriction enzymes *Hae*III (recognition site 5'-GGCC-3') and *Msp*I (recognition site 5'-CCGG-3'), as shown in Fig.1b, although recognition sequences of both enzymes consist of only GC pairs. *Msp*I recognition site contains low-frequency dinucleotide CG and enzyme's action results in a smaller depth of DNA hydrolysis if compared to that of *Hae*III hydrolysis.
A contrary situation is observed for dinucleotides TG, CA, AG and CT. These dinucleotides in

chromosomal DNA of eukaryotes are represented with higher frequency compared to the average statistical value of occurrence [15, 16]. This can explain the deep hydrolysis, observed in Fig.2b for restriction enzymes *Fat*I, (recognition site CATG), *Alu*I(recognition site AGCT) and *BstDE*I (recognition site CTNAG). The recognition site of these restriction endonucleases contains two high-frequency dinucleotides (CA and TG for *Fat*I, AG and CT for *Alu*Iand *BstDE*I) at a time.

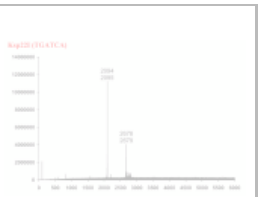
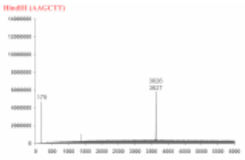
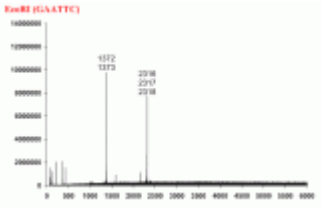
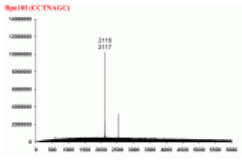
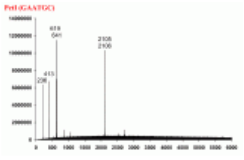
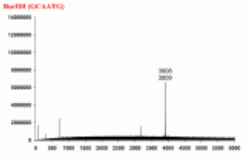
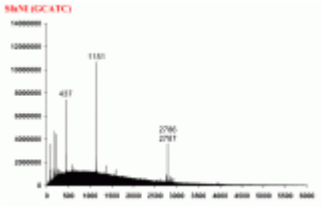
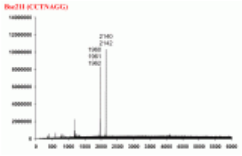
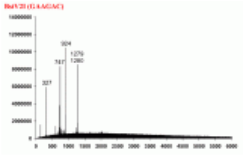
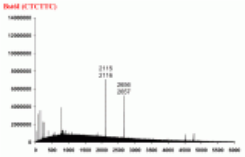
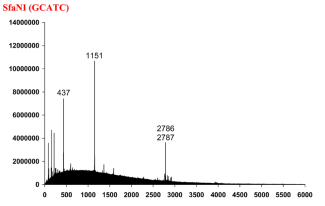
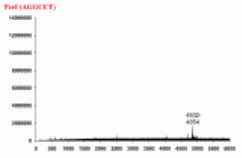
			
***Pvu*II**	***Vsp*I**	***Pci*I**	***Kzo22*I**
			
***Hind*III**	***EcoR*I**	***Bpu10*I**	***Pct*I**
			
***Bse3D*II**	***SfaN*I**	***Bse21*I**	***BstV2*I**
			
***BstV6*II**	***BamH*I**	***Pce*I**	

Fig. 4. Detailed distribution diagrams of total DNA fragments lengths depending on the fragment size for restriction endonucleases recognition sites.

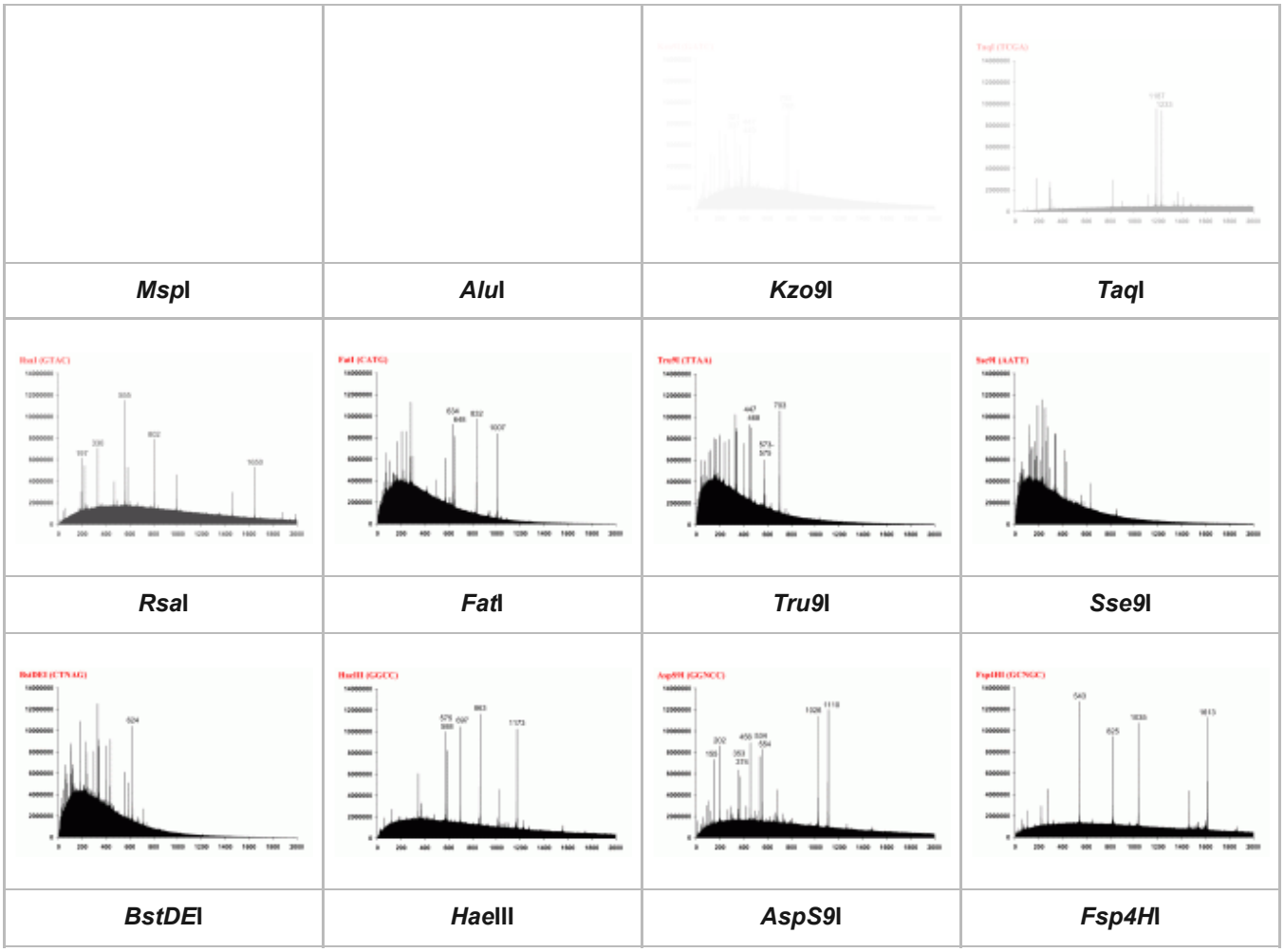


Fig. 5. Detailed distribution diagrams of total DNA fragments lengths depending on the fragment size for restriction endonucleases recognition sites.

CONCLUSION

Diagrams of rat chromosomal DNA cleavage at 25 recognition sequences of restriction endonucleases have been plotted and discussed. The presence of more than fifty significant peaks of DNA fragments (or clusters of fragments) has been revealed for a cleavage at these nucleotide sequences. A map of LINE1 consensus repeat digestion with restriction enzymes has been produced and a high correspondence between the experimental data displayed in diagrams and LINE1 repeat's cleavage positions has been shown. Based on restriction enzymes analysis a revised version of consensus LINE1 repeat sequence has been suggested. Experiments on chromosomal DNA hydrolysis with restriction endonucleases and subsequent electrophoresis of reaction products in agarose and/or polyacrilamide gels have been performed. A comparison of the constructed distribution diagrams of rat chromosomal DNA fragments and the patterns of DNA digestions with the corresponding restriction endonucleases has revealed a good correspondence between theoretical and experimental data.

The authors thank Dr. V. I. Kaledin Dr. and Dr. G. V. Vasilyev for assistance in animal experiments and DNA isolation.

REFERENCES

1. Rat Genome Sequencing Project Consortium 2004. Genome sequence of the Brown Norway rat yields insights into mammalian evolution. // Nature - 2004 - Vol.428. - P.493 - 521.
2. Tomilov V.N., Chernukhin V.A., Abdurashitov M.A., Gonchar D.A., Degtyarev S.Kh. The method of restriction analysis of mammalian genomes *in silico*. // The Yu.A. Ovchinnikov Bulletin of Biotechnology and Physico-chemical Biology. - 2007 - Vol.3 -No.1 - P . (in Russian , **English online version**)
3. Dashkevich V.S., Dymshits G.M., Salganik R.I., Ulanov B.P. About the destabilization of the secondary structure of the DNA of regenerating rat liver in the process of replication // Molecular Biology. - 1972. - Vol.6 -No.5. - P. 689-697 (in Russian)
4. Southern E.M., 1970, Nature (Lond), 227, 794-798; Walker P,M.B. // Progr. Biophys. Mol. Biol. 1971 - Vol.23. - P.145-190
5. Philippsen P., Streeck R.E., Zachau H.G. Defined Fragments of Calf, Human, and Rat DNA Produced by Restriction Nucleases. // European Journal of Biochemistry - 1974 - Vol.45., - P.479-488.
6. Pech M., Igo-Kemenes T., Zachau H.G. Nucleotide sequence of highly repetitive component of rat DNA. // Nucl. Acids Res. - 1979. - P.417-432.
7. Horz W., Hess I., Zachau H.G. Highly regular arrangement of restriction nuclease sensitivity site in rodent satellite DNAs. // European Journal of Biochemistry 1974 - Vol.45. - P. 501-512.
8. Soares, M.B., Schon, E., Efstratiadis A. Rat LINE1: The origin and evolution of long interspersed middle repetitive DNA elements. // Journal of Molecular Evol. - 1985, - Vol.22. - P.117-133
9. Penzkofer T., Dandekar T., Zemojtel T. L1Base: from functional annotation to prediction of active LINE-1 elements. Nucleic Acids Research, 2005, Vol. 33, D498-D500.
10. SibEnzyme Database site - **LINE1 DNA repeat text file**
11. Ehrlich, M., Gama S.M., Huang L.H., Midgett, R.M., Kuo, K.C., McCune, R.A., and Gehrke, C. Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cell. // Nucl. Acids Res. - 1982 - Vol.10 - P.2709-2721
12. Kalinin F.L., Lobov V.P., Zhidkov V.A. Biochemistry Reference Book. "Naukova Dumka", Kiev 1971 (in Russian)
13. Schorderest D.F., and Gartler S.M. Analysis of CpG suppression in methylated and nonmethylated species. // Proc. Natl. Acad. Sci. USA - 1992 - Vol.89 - P.957-961.
14. Lindahl T. Instability and dacay of the primary structure of DNA. // Nature - 1993 -Vol.362. - P.709-715;
15. Ohno S. Universal rule for coding sequence construction: TA/CG deficiency-TG/CT excess. // Proc. Natl. Acad. Sci. USA - 1988 - Vol.85 - P.9630-9634.
16. Yomo T., Ohno S. Concordant evolution of coding and noncoding regions of DNA made possible by the universal rule of TA/CG deficiency-TG/CT excess.// Proc. Natl. Acad. Sci. USA - 1989 - Vol.86 - P.8452-8456.
17. Jurka J., Kapitonov V.V., Pavlicek A., Klonowski P., Kohany O., Walichiewicz J. (2005) Repbase Update, a database of eukaryotic repetitive elements. Cytogenetic and Genome Research 110: 462-467.

ACCEPTED ABBREVIATIONS

RE - restriction endonuclease
bp - base pair
EDTA - ethylene diamine tetraacetic acid
PAAG - polyacrylamide gel