Biochimica et Biophysica Acta, 740 (1983) 339–341
Elsevier

BBA Report

BBA 90022

THE DISTRIBUTION OF CR1, AN *Alu*-LIKE FAMILY OF INTERSPERSED REPEATS, IN THE CHICKEN GENOME

BIRGITTA OLOFSSON and GIORGIO BERNARDI

Laboratoire de Génétique Moléculaire, Institut Jacques Monod, 2 Place Jussieu, 75005 Paris (France)

(Received February 23rd, 1983)

Key words: Interspersed repeat; Evolution; Gene structure; Hybridization; (Chicken genome)

We have studied the distribution of CR1, a family of short interspersed repeats, in the chicken genome; this family is homologous to the *Alu*I family of man and to the B1-B2 families of mouse. Hybridization with a suitable probe showed that the vast majority of CR1 are located on the heaviest major component (1.708) of the genome which only represents 9% of chicken DNA. Some repeats were also found on the 1.702 and 1.704 components, but none on the 1.699 component (components are denoted by their buoyant densities in CsCl). The GC content of the repeats, 48%, matches that, 47%, of the major component mainly harboring it.

Main-band DNAs (from which satellite and minor components have been removed) from warm-blooded vertebrates can be fractionated by density gradient centrifugation into four families of fragments, derived by preparative breakage of long DNA segments (over 300 kb in size) of fairly homogenous composition [1–6]. These segments, called isochores, may be responsible for chromosomal Giemsa and Reverse banding patterns [4]. Two of the fragment families represent about two-thirds of main-band DNA and have buoyant densities in the 1.697–1.703 g/cm^3 range; the other two have buoyant densities of 1.704 and 1.708 g/cm^3, respectively. In contrast, main-band DNAs from cold-blooded vertebrates do not have the wide compositional heterogeneity of warm-blooded vertebrate DNAs and have, in most cases, buoyant densities in the same range as the light components of the latter [2,7].

The isolation of major components, as we have called the families of fragments discussed above, from mammalian and avian DNAs [4,6] has allowed us to study the genomic distribution of specific nucleotide sequences. Here we report that the CR1 family of short interspersed repeats [8] is essentially present in the 1.708 component, (which only represents 9% of chicken DNA [5,6]), and matches this component in base composition. This match, as well as a strikingly non-uniform genomic distribution, have been found for other interspersed repeats from the mouse and the human genomes [9,10]; in particular the B1-B2 repeats of mouse [11] and the *Alu*I repeats of man [12], which are homologous to the CR1 repeats [8], exhibit a genomic distribution similar to that of the latter, indicating that this distribution is conserved in evolution.

The chicken DNA preparation and the major components used in this work have been described elsewhere [5,6]. The methods used are as described previously [10]. The probe used to detect the repeats was a plasmid containing the CR10va repeat which is located in the neighborhood of the chicken ovalbumin gene and is one of the two repeats of the CR1 family which have been cloned [8].

Hybridization of radioactive CR10va repeat on *Eco*RI digests of total chicken DNA yielded a strong hybridization smear in a region of high molecular weight fragments (6–9 kb). Hybridization on the four major components (Fig. 1) re-

Fig. 1. Autoradiograms of pCR10va hybridized on *Eco*RI digests from unfractionated chicken DNA (T) and its major components, 1.699 (1), 1.702 (2), 1.704 (3) and 1.708 (4). Specific activity of the labeled pCR10va was 10^8 cpm/μg. Stringency of the last washing ws 0.5 standard saline solution, 65°C. Loads were 30 μg for total DNA, and 14 μg, 14μg, 8.4 μg and 5.5 μg, respectively, for components 1–4; loading of components 2–4 was proportional to their percentages in the DNA (see Table I).

vealed that most of the label hybridized on the 1.708 component; much lower amounts of radioactivity were detected in the 1.702 and 1.7.04 components, and none in the 1.699 component. In the case of the 1.708 component, a strong smear covered the 5–9 kb region; in addition, some bands appeared in the 2–5 kb region, indicating a discrete distribution of DNA fragments containing the repeat in this size range.

These results indicate a distinctly non-uniform distribition of the CR1 family over the chicken genome with major abundance in the heaviest DNA component. This distribution is similar to those of the homologous B1-B2 families of mouse and of the *Alu*I family of man [10]. In fact, the preferential distribution in the heaviest component is even sharper in the case of the chicken genome. This finding stresses the evolutionary conservation of the genomic distribution of homologous families of repeats in vertebrates.

As far as the low hybridization found in the 1.702 and 1.704 components is concerned, two points should be noted. First pCR10va contains a repeat located approx. 12 kb downstream of the 3′ end of the ovalbumin gene; three other cross-reacting sequences are 25 kb from the 3′ end of the ovalbumin gene and 18 and 25 kb from the 5′ end of the X gene, respectively [8], as displayed in Fig. 2. Second, a region encompassing at least 100 kb on either side of the ovalbumin gene has been shown to be localized on the 1.702 component [5]. The actual demonstration of CR1 repeats in this component accounts, therefore, for the weak hybridization on it.

The results given in Fig. 1 are also of interest in connection with the question of the copy number of CR1 repeats. On the basis of reassociation kinetics this number has been estimated as 1300/haploid genome [8]. Correction of this estimate for sequence divergence and sequence length on the basis of data on the *Alu*I repeats would lead to 7000 copies [8]. Taking into account the length of the repeats, 180 bp, and the genome size, 1.25 pg DNA/haploid genome (an average of published values), this copy number corresponds to 0.1% of the genome. This represents only a minute fraction of the intermediate sequences, which account for 10% of chicken DNA (see Table I). Likewise, since the vast majority of repeats is located in the 1.708 component, which represents 9% of the genome, the concentration of repeats in it is approx. 1% out of the total of 17% intermediate sequences in this component (Table I). The extremely low calculated amounts of CR1 repeats in both total DNA and 1.708 component suggest that even the corrected value of the copy number, which is 50–100-times lower than that of *Alu*I repeats, is an underestimate.

A final observation is that the GC content of the CR1 repeats, 48% [8], matches very well the

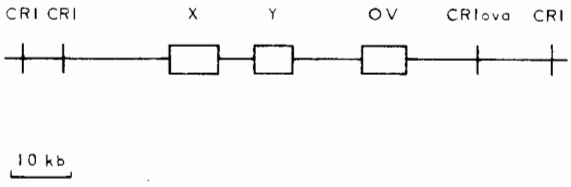

Fig. 2. Scheme of the ovalbumin gene region and of the localization of CR1 repeats. The position of CR10va is indicated as well as those of three other CR1 repeats. X, Y and OV indicate the positions of X, Y and ovalbumin genes. (From results of Refs. 8, 14 and 15).

TABLE I

GC CONTENT RELATIVE AMOUNTS AND % INTERMEDIATE SEQUENCES IN THE MAJOR COMPONENTS OF THE CHICKEN GENOME [a]

Component	GC (%)	Relative amount (%)	Intermediate repeated sequences (%)
1.699	38.3	39	5
1.702	39.7	25	10
1.704	42.7	15	2
1.708	46.7	9	17
Total	39.3	100	10

[a] Modified from Ref. 6.

value of 47% [6], of the major component in which it is predominantly localized. This result is in agreement with similar ones obtained on mouse and human interspersed repeats [9,10].

To sum up, the present results extend those previously obtained on the mouse and human repeats [9,10] in that the genomic distribution of the CR1 *Alu*I-like family of repeats appears to be highly non-uniform and conserved in evolution. The base composition of these short repeated sequences matches that of the long DNA stretches in which they occur. It is clear from these data that the term 'ubiquitous' often used in conjunction with the genomic distribution of *Alu*-like repeats is most inappropriate. Furthermore, if these repeats are formed by mobile, retroposon-like [12] sequences as suggested for both the *Alu*I and the B1-B2 families (see Ref. 13 for a very recent review) sequence movement must be largely limited to certain chromosomal domains which correspond to the heavy isochores.

We wish to express our appreciation to Dr. M.-J. Tsai for the probe used in the present work.

References

1 Filipski, J., Theiry, J.P. and Bernardi, G. (1973) J. Mol. Biol. 80, 177–197
2 Thiery, J.P., Macaya, G. and Bernardi, G. (1977) J. Mol. Biol. 108, 219–235
3 Macaya, G., Thiery, J.P. and Bernardi, G. (1977) J. Mol. Biol. 108, 237–254
4 Cuny, G., Soriano, P., Macaya, G. and Bernardi, G. (1981) Eur. J. Biochem. 115, 227–233
5 Cortadas, J., Olofsson, B., Meunier-Rotival, M., Macaya, G. and Bernardi, G. (1979) Eur. J. Biochem. 99, 179–186
6 Olofsson, B. and Bernardi, G. (1983) Eur. J. Biochem. 130, 241–245
7 Hudson, A.P., Cuny, G., Cortadas, J., Haschemeyer, A.E.V. and Bernardi, G. (1981) Eur. J. Biochem. 112, 203–210
8 Stumph, W.E., Kristo, P., Tsai, M.-J. and O'Malley, B.W. (1981) Nucleic Acids Res. 9, 5383–5397
9 Meunier-Rotival, M., Soriano, P., Cuny, G., Strauss, F. and Bernardi, G. (1982) Proc. Natl. Acad. Sci. U.S.A. 79, 355–359
10 Soriano, P., Meunier-Rotivl, M. and Bernardi, G. (1983) Proc. Natl. Acad. Sci. U.S.A., in the press
11 Krayev, A.S., Kramerov, D.A., Skryabin, K.G., Ryskov, A.P., Bayev, A.A. and Georgiev, G.P. (1980) Nucleic Acids Res. 8, 1201–1215
12 Rogers, J. (1983) Nature 301, 460
13 Sharp, P. (1983) Nature 301, 471–472
14 Royal, A., Garapin, A., Cami, B., Perrin, F., Mandel, J.-L., Le Meur, M., Brégégère, F., Gannon, F., Le Pennec, J.P., Chambon, P. and Kourilsky, P. (1979) Nature 279, 125–132
15 Heilig, R., Perrin, F., Gannon, F., Mandel, J.-L. and Chambon, P. (1980) Cell 20, 625–637