Volume 8 Number 21 1980 Nucleic Acids Research

Structure and organization of the highly repeated and interspersed 1.3 kb EcoRI - BglII sequence family in mice

Rita Heller[1] and Norman Arnheim[2]

Department of Biochemistry and Cellular and Developmental Biology Graduate Program, S.U.N.Y., Stony Brook, NY 11794, USA

Received 18 July 1980

ABSTRACT

EcoRI digestion of total mouse DNA yields a prominant 1.3 kb fragment amounting to between 1 and 2% of the mouse genome. The majority of the 1.3 kb EcoRI fragments have a single BglII site 800 bp from one end. This EcoRI-BglII sequence family shows HindIII and HaeIII sequence heterogeneity. We have cloned representatives of the EcoRI-BglII gene family in Charon 16A and studied their structure and organization within the genome. The cloned 1.3 kb fragments show the expected restriction enzyme patterns as well as additional heterogeneity. Representatives of the EcoRI-BglII sequence family were found to be interspersed throughout the mouse genome as judged by CsCl density gradient centrifugation experiments. Family members were also found to be organized in higher order repeating units. Homologous sequences were also found in other rodent species including rat and Chinese hamster. Cross hybridization between a cloned 1.3 kb mouse fragment and a cloned CHO repeated sequence is of special interest since the latter has been shown to contain sequences homologous to the Human AluI family by nucleotide sequencing.

INTRODUCTION

Different degrees of DNA sequence repetition have been defined in eukaryotic genomes (1-4): highly repetitious tandem arrays of simple sequences or satellites, tandemly arranged moderately repeated sequences coding for structural RNA's or proteins and interspersed repeated sequences. The function of this latter sequence class has been the subject of considerable investigation (see 5,6) but their role in the genome remains unknown.

Interspersed repeated mouse DNA sequences have been recognized as a result of nucleic acid reannealing studies and restriction enzyme analysis (7-9). Reannealing studies showed that a sizable fraction of highly repeated interspersed mouse DNA renatured to give double stranded segments between 1 kb and 1.5 kb (7,8). Restriction enzyme analysis of CsCl purified main band DNA revealed the existence of a similarly sized (1.3 kb) EcoRI fragment class which was also interspersed (9). We have further character-

ized this latter repeated gene family by restriction enzyme mapping and recombinant DNA technology.

MATERIALS AND METHODS

DNA Samples

DNA from Swiss mice (purchased from Taconic Farms) and rat livers (a gift from Dr. G. Brown) and Chinese Hamster ovary cell line E-36 were prepared by the method of Walker and McLaren (10). Livers from other mouse species were gifts from Dr. M. Potter of the NIH.

Restriction enzymes were purchased from New England Biolabs and Boehringer Mannheim. We used the digestion conditions recommended by the supplier.

Electrophoresis, Southern Transfer and Hybridization

Restriction digests were displayed on 1% or 1.5% agarose gels run at 1-3 V/cm. DNA fragments were transferred to nitrocellulose filters according to Southern (11). EcoRI and HindIII digests of λ and HindII and HaeIII digests of ϕx 174 were used as size markers.

Nick translations were carried out using α^{32}P dATP and α^{32}P dCTP (Amersham >350 Ci/mMole) according to the method of Rigby et al (12). Specific activities were 10^5 cpm to 10^7 cpm/μg for gel purified probes and $2x10^8$ cpm/μg for cloned probes.

Filters were hybridized in 50% formamide, 5xSSC for 18-20 hrs at 37°C, washed in the same buffer for 3-5 hours at 37°C, with occasional shaking, and then rinsed briefly with 10 mM or 20 mM Tris at room temperature.

Autoradiography was performed using Kodak XR-1 film at -70°C with Tungstate screens (13).

Purification of Restriction Enzyme Fragments

Restriction fragments were isolated from agarose gels after visualization by EtBr staining using a modification of the freeze squeeze method (14). The excised gel piece was minced with a sharp blade, frozen at -20°C for 1 hr, then thawed and centrifuged at 40K in an SW50.1 rotor for 3 hrs. The supernatant was then either passed over DE-52 and/or ETOH precipitated with carrier E. coli tRNA. When used as a probe the 1.3 kb material was gel purified two times. For cloning only a single gel run was used for purification.

CsCl Gradients

DNA was needle sheared to a 10 kb size average and run on neutral CsCl gradients (initial density 1.700 g/ml , $ND^{25°}$ = 1.3996) for 3 days at 35,000 rpm in a 50.2 Ti rotor.

Recombinant DNA Techniques

Gel purified 1.3 kb EcoRI fragments were cloned into the EcoRI site of

Charon 16A (15) by *in vitro* packaging (16). Genomic clones were obtained by screening a Balb/c genome library with gel purified 1.3 kb material. The gene library was constructed by ligating partial EcoRI digests of Balb/c liver DNA with the left and right arms of Charon 4A (15) and was a gift of Dr. Ken Marcu. The cloned ribosomal gene probe has been described previously (17). All phage were grown and handled according to the then appropriate P2-EK2 conditions set forth in the NIH Guidelines. One of the 1.3 kb inserts in a Charon 16A clone was subcloned into the EcoRI site of pBR322. Selection of plasmids carrying the 1.3 kb piece was carried out using a modification of the Grunstein-Hogness procedure (18).

RESULTS

Organization of the Mouse 1.3 kb Fragment Class

Total mouse liver DNA digested with EcoRI reveals several prominent bands visible above the characteristic background smear of ethidium bromide staining (Fig. 1A). We estimate that the 1.3 kb fragment class makes up 1-2% of the mouse genome. Southern transfer and hybridization experiments with EcoRI digests using either gel purified 1.3 kb material or a cloned 1.3 kb EcoRI fragment (see below) as a probe reveal a prominant 1.3 kb fragment as expected (Fig. 1C). In addition, a broad smear of hybridization was detected at higher molecular weights. Since some hybridization appeared to occur in the region of the gel known to contain mouse satellite DNA we examined this possibility further. Using a cloned 1.3 kb fragment as a probe we observed no hybridization to the fragments of mouse satellite DNA generated by the restriction enzyme EcoRII (19). This confirms the same conclusion reached by Horz, *et al.* (9). We next examined whether the 1.3 kb fragments are organized in tandem blocks or are interspersed throughout the genome. Partial EcoRI digests of total mouse liver DNA show no multimers of the 1.3 kb piece detectable by ethidium bromide staining. Southern transfer of these partial digests followed by hybridization with a cloned 1.3 kb piece also gave no indication of multimers of the 1.3 kb segment.

To further examine the distribution of the sequences in the genome, DNA which had been sheared to a 10 kb average size was run on neutral CsCl density gradients. Each gradient fraction was digested with EcoRI and run on an agarose gel. The 1.3 kb band was seen in each fraction by ethidium bromide staining (data not shown). Hybridization using 1.3 kb gel purified probe gave the same result (Fig. 2). In this hybridization experiment our 1.3 kb probe was mixed with a cloned ribosomal gene probe (17). This probe will

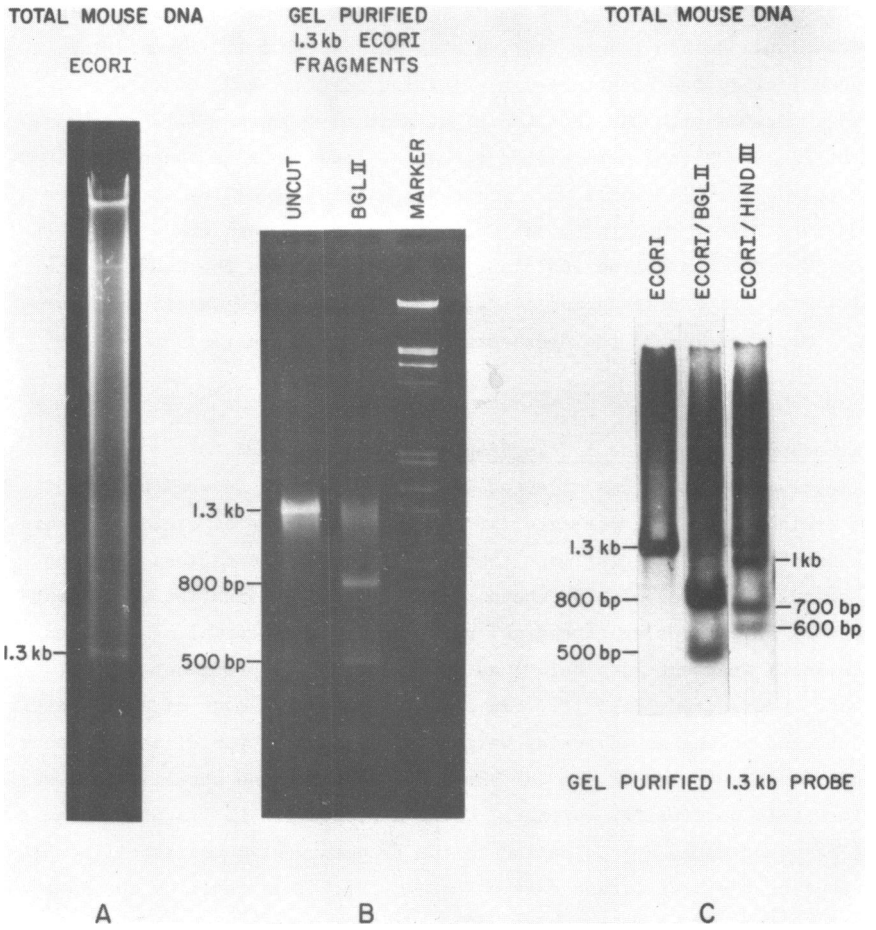

Figure 1. A) Ethidium bromide staining pattern of EcoRI digested mouse DNA. B) Restriction enzyme analysis of gel purified 1.3 kb material after ethidium bromide staining. C) Restriction enzyme analysis of sequences homologous to the 1.3 kb gene family in total mouse DNA. The hybridization probe was nick-translated gel purified 1.3 kb material. The 400 bp EcoRI-HindIII fragment is not visible in this experiment.

detect the 6.6 kb mouse rDNA Eco RI fragment carrying part of the 18S gene, the internal transcribed spacer and most if not all of the 28S gene. If our gradient fractionated the mouse DNA according to its G+C content we would expect hybridization with the ribosomal gene probe to occur at the higher densities (20). As is shown in Figure 2, the 1.3 kb fragment is found throughout the gradient whereas the 6.6 kb rDNA piece is found primarily in

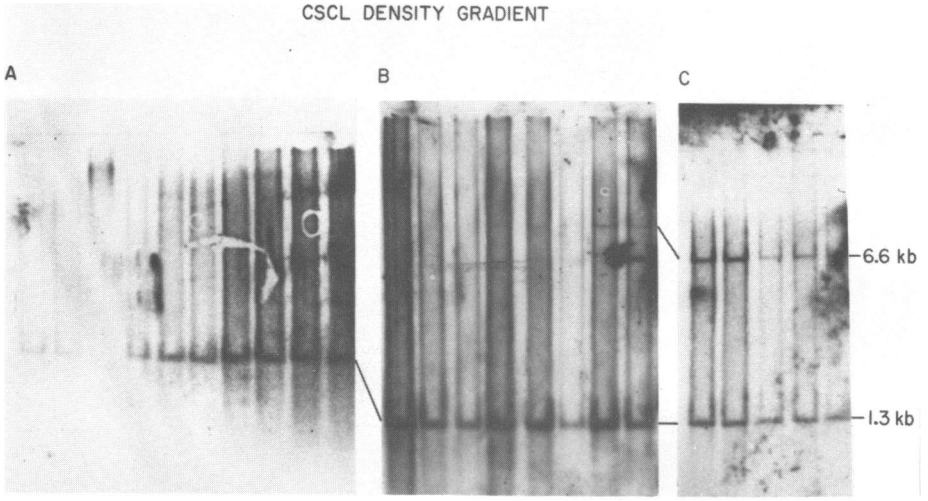

Figure 2. Distribution of the 1.3 kb repeated sequence family in mouse DNA fractionated by CsCl density gradient centrifugation. One-third of each sample was ethanol precipitated, digested with EcoRI, run on a 1.5% agarose gel and hybridized to a nick-translated mixture of gel purified 1.3 kb DNA and cloned 6.6 kb rDNA. The samples were analyzed on three separate gels beginning with the lowest density fractions (A) and progressing to those of higher density (C).

the highest density fractions. This confirms that the 1.3 kb EcoRI sequences are distributed throughout the genome.

The 1.3 kb EcoRI Fragments Contain a Single BglII Site

If total mouse DNA is digested with EcoRI followed by digestion with BglII the 1.3 kb fragment class essentially disappears and two new fragments, 500 bp and 800 bp in size, can be detected by ethidium bromide staining and after Southern transfer and hybridization with gel purified or cloned 1.3 kb probe (Fig. 1C). Virtually all of the 1.3 kb EcoRI fragments have a single BglII site about 800 bp from one end. When gel purified 1.3 kb material is doubly digested with EcoRI and BglII the same two fragments are detected (Fig. 1B). The conservation of the EcoRI and BglII sites in these repeated and interspersed sequences prompts us to define them as the EcoRI-BglII sequence family. Further studies were carried out using cloned material derived from gel purified EcoRI 1.3 kb fragments and inserted into Charon 16A. Two were selected for further investigation. Both recombinants contain a 1.3 kb EcoRI fragment which hybridizes to gel purified 1.3 kb probe and cross-hybridize with each other.

A restriction enzyme map of these two clones is shown in Fig. 3A. One

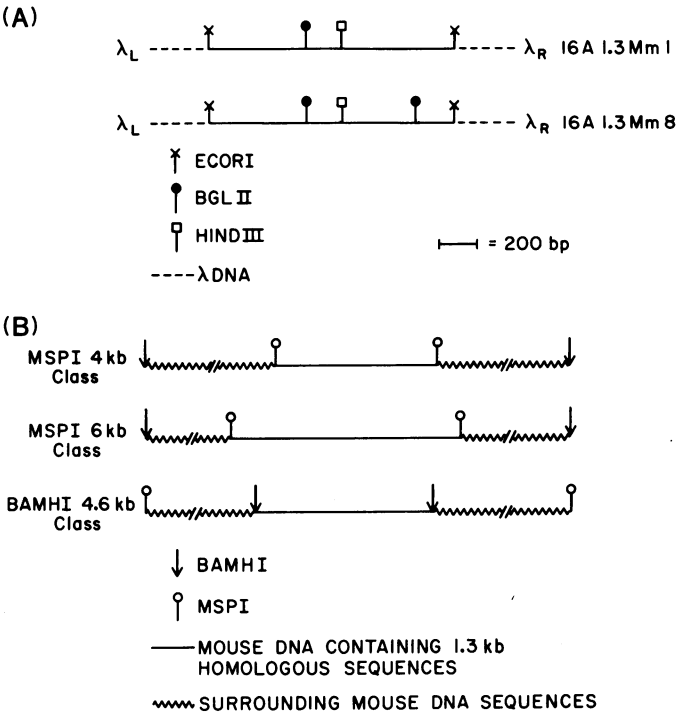

Figure 3. A) Restriction enzyme maps of two cloned 1.3 kb fragments. B) Model for the higher order organization of 1.3 kb homologous sequences in the mouse genome. See text for a detailed discussion.

clone (16A 1.3 Mm1) yields a 800 bp and 500 bp fragment when doubly digested with EcoRI and BglII as expected from our results on total mouse DNA. The other clone (16A 1.3 Mm 8) produces three BglII-EcoRI fragments; an expected 500 bp piece and two new fragments (600 bp and 200 bp) whose sizes sum up to 800 bp. This 1.3 kb fragment may represent a minor class found in the mouse genome since the 600 bp and 200 bp Eco RI-BglII fragments are not readily detectable in Southern transfer and hybridization experiments with total mouse DNA.

Three additional 1.3 kb fragments were studied. These fragments were part of larger genomic clones obtained by screening a partial EcoRI Balb/c mouse gene library with 1.3 kb gel purified probe. The three clones, each between 10 kb and 15 kb, gave different EcoRI restriction enzyme fragment patterns consistent with the idea that the 1.3 kb sequences are found scattered throughout the genome. Two of these genomic clones (4A 1.3 Mm 1 and 4A 1.3 Mm 2) each produced a 1.3 kb EcoRI fragment which was further digested

by BglII into 800 bp and 500 bp pieces. These EcoRI-BglII fragments hybridized to 1.3 kb gel purified or cloned probe. The third genomic clone (4A 1.3 Mm 3) yielded a 1.4 kb EcoRI fragment which hybridized to the 1.3 kb probe. Further digestion with BglII gave an expected 800 bp piece and a 600 bp rather than 500 bp fragment. Both pieces hybridized to the 1.3 kb probe. This clone may also represent a minor component of the 1.3 kb sequence family since a 600 bp EcoRI-BglII fragment which hybridizes to the 1.3 kb probe is not readily detected in experiments using total mouse DNA.

In conclusion, 3 of the 5 cloned 1.3 kb EcoRI fragments give the BglII digestion fragments characteristic of EcoRI-BglII double digests of total mouse DNA. In each of the two other clones one of the two expected fragment classes were detected. These studies on the cloned material further support the existence of a relatively homogeneous repeated and interspersed EcoRI-BglII sequence family.

Studies with Other Restriction Enzymes

Double digestion of total mouse DNA with EcoRI and HindIII reveals heterogeneity in the structure of the 1.3 kb EcoRI-BgII family. The major portion of the 1.3 kb material is cleaved by HindIII into 4 bands of about 1 kb, 700 bp, 600 bp and 400 bp as detected on Southern transfers (Fig. 1C). In addition, some of 1.3 kb material is not digested by HindIII. These results suggest the existence of at least 3 HindIII sub-classes of the 1.3 kb EcoRI size class.

HindIII studies on the two 1.3 kb fragments cloned in Charon 16A were also carried out. HindIII digests of the 1.3 kb EcoRI fragment in both of these clones yields a 700 bp and 600 bp piece. The position of the HindIII sites relative to the BglII sites are shown in Fig. 3A. These two clones generate two of the EcoRI-HindIII fragments detected in total mouse DNA.

HaeIII-EcoRI double digest of total mouse DNA produces a pattern of bands suggesting the structure of the EcoRI-BglII family is still more complex although fragments about 750 bp and 600 bp are the major restriction digest products (data not shown).

Higher Order Organization of the 1.3 kb Sequences in the Mouse Genome

Using total mouse DNA and Southern transfer and hybridization experiments with the 1.3 kb probe we found a number of enzymes which when used in double digest experiments with EcoRI do not appear to cut the 1.3 kb fragment class. However in some of these double digest experiments fragments larger than 1.3 kb and which were not detected when EcoRI was used alone were found to hybridize to the 1.3 kb probe.

This suggested that not all of the sequences which hybridize to the 1.3 kb probe are bounded in the genome by EcoRI sites 1.3 kb apart. The inability to detect these homologous sequences when DNA is digested with EcoRI alone might be due to the fact that they are bounded by randomly distributed EcoRI sites and therefore form a continuous distribution of EcoRI fragment sizes. Evidence supporting this also comes from the observation that a smear of hybridization detected with the 1.3 kb probe occurs at molecular weights higher than 1.3 kb when mouse DNA is digested with EcoRI (Fig. 1C). The fact that digestion with a second enzyme can often produce discrete high molecular weight fragments suggests that the EcoRI-BglII family may be present in some other higher order structure in the genome.

From double digest experiments of total DNA with EcoRI and either MspI or BamHI we know that the majority of the 1.3 kb EcoRI class lacks internal BamHI and MspI sites. When total mouse DNA is digested with MspI, BamHI or both enzymes and hybridized with a cloned 1.3 kb fragment after Southern transfer (Fig. 4), two mspI fragments, 4 kb and 6 kb, are detected. BamHI digests show only a single band 4.6 kb in size. When double digests using both enzymes were carried out only those fragments observed with either enzyme alone were seen. No new bands were detected. This data is consistent with the model shown in Figure 3B. The 4 kb and 6 kb MspI fragment classes each have no internal BamHI sites. The DNA which surrounds these MspI fragments contain BamHI sites but the location of these surrounding BamHI sites varies randomly among the members of the MspI fragment class. Likewise the 4.6 kb BamHI fragment class contains no internal MspI sites. The surrounding DNA however does have MspI recognition sequences but again the positions of these sites varies randomly around the members of the BamHI fragment class. The consequence of the arrangement of these sites is that upon BamHI digestion the sizes of the BamHI fragments carrying the 4 kb and 6 kb MspI classes form a continuous size distribution. Only the 4.6 kb BamHI fragment class carrying the EcoRI-BglII family is observed to hybridize at a discrete position. Similarly MspI digestion reveals only the 4 kb and 6 kb fragments classes carrying the EcoRI-BglII family sequences. The sizes of the MspI fragments carrying the BamHI class also forms a continuous size distribution and are not detected as a discrete class with the 1.3 kb probe.

Evolutionary Conservation of the EcoRI-BglII Family

DNA samples from two Asian (*Mus caroli* and *Mus cervicolor*) were found to contain sequences homologous to the EcoRI-BglII family (data not shown).

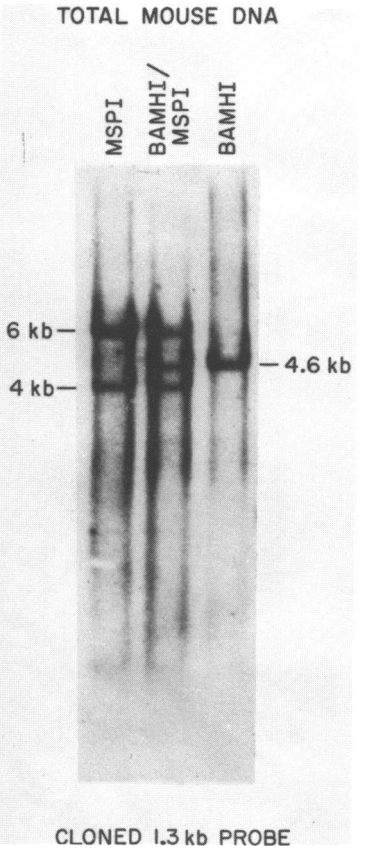

Figure 4. Southern transfer and hybridization experiments analyzing the organization of the 1.3 kb gene family in the mouse genome. Total mouse DNA was digested with either MspI, BamHI or both enzymes and analyzed with a cloned 1.3 kb probe.

Rat and Chinese hamster cell line DNA's also exhibited homologous sequences. Southern transfer and hybridization experiments with EcoRI digests of rat DNA revealed a 1.3 kb fragment that hybridizes to the cloned mouse probe (Fig. 5). EcoRI digests of Chinese hamster cell line DNA give rise to a 2 kb fragment which is also homologous to the cloned mouse probe (Fig. 5). Double digestion of rat DNA with EcoRI and BglII show that the 1.3 kb EcoRI fragments have a single BglII site which is characteristic of the mouse EcoRI-BglII family (data not shown). A more complex pattern is observed in the case of the homologous sequences found in the Chinese hamster DNA sample.

A cloned fragment of Chinese hamster cell line DNA (clone 63) homologous

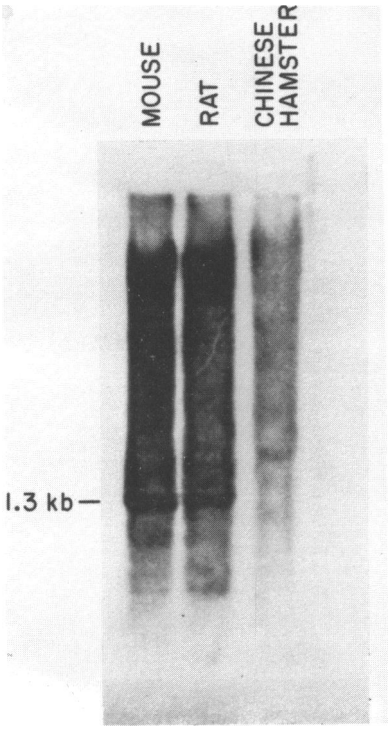

Figure 5. Southern transfer and hybridization analysis of sequences homologous to the mouse EcoRI-BglII family in rat and Chinese hamster. Total DNA of each species was digested with EcoRI. In this experiment gel purified 1.3 kb probe was used. Other experiments in which cloned 1.3 kb probe was employed gave identical results.

to an hnRNA repeated sequence (6) was found to cross hybridize with the cloned mouse 1.3 kb EcoRI piece (L. Leinwand, personal communication).

DISCUSSION

We have investigated some of the properties of the highly repeated 1.3 kb EcoRI fragment class found in the mouse genome. Partial EcoRI digestion and CsCl gradient centrifugation experiments show that these 1.3 kb pieces are not tandemly arranged but interspersed throughout the genome. Restriction enzyme analysis revealed that the majority of the 1.3 kb EcoRI fragments had a single BglII site 800 bp from one end. Thus, in spite of their interspersion throughout the mouse genome, there is conservation of the BglII site which defines these sequences as the EcoRI-BglII family. HindIII

and HaeIII analysis showed that the EcoRI-BglII sequence family also exhibits a sequence heterogeneity expected of a highly reiterated interspersed gene family. Additional variation was found when cloned 1.3 kb fragments were also examined.

It has been proposed that interspersed repeated DNA sequences play a role in DNA replication, transcription or RNA processing (5,6,21). We do not know if the EcoRI-BglII sequence family is transcribed. Clones carrying DNA sequences which are highly reiterated in the mouse genome have been selected using hnRNA probes (21) but the organization of these DNA segments in the mouse genome have not been mapped with restriction enzymes.

In humans the major interspersed repeated sequence family isolated by low Cot reannealing is cut into uniform size fragments by AluI and has been defined as the AluI family (22). Nucleotide sequence comparisons of the human AluI family with a cloned Chinese hamster cell line DNA fragment (Clone 63) which hybridizes with CHO hnRNA reveals a significant 30 bp homology (6). Although one of our EcoRI-BglII family clones cross hybridizes with this CHO clone, we do not know whether the same 30 bp sequence or other evolutionarilly conserved repeated sequences are responsible.

We have found that the EcoRI-BglII sequence family is organized into larger defined units within the mouse genome. MspI and BamHI reveal two higher order organizations and other enzymes hint of additional arrangements. It is possible that the EcoRI-BglII family is arranged with other repeated sequences in clusters (23,24).

Some of the 1.3 kb moderately repeated mouse DNA sequences cloned by Cheng and Schildkraut may be members of the EcoRI-BglII family.

ACKNOWLEDGEMENTS

This work was supported in part by an NIH grant and a PHS Biomedical Research Support Grant.

[1]Present address: Department of Biology, University of Pennsylvania, Philadelphia, PA 19143, USA

[2]From whom reprints should be requested

REFERENCES

1. Britten, R.J. and Kohne, D.E. (1968). Science 161, 529.
2. Brutlag, D.L. (1980) Annual Review of Genetics 14, In press.
3. Davidson, E.H., Galaw, G.A., Angerer, R.C. and Britten, R.J. (1975) Chromosoma 51, 253.
4. Long, E.O. and Dawid, I.B. (1980) Ann. Rev. Biochem. 49, 727.

5. Davidson, E.H. and Britten, R.J. (1979) Science 204, 1052.
6. Jelinek, W.R., Toomey, T.P., Leinwand, L., Duncan, C.H., Biro, P.A., Choudary, P.V., Weissman, S.M., Rubin, C.M., Houck, C.M., Deininger, P.L. and Schmid, C.W. (1980) Proc. Natl. Acad. Sci. USA 77, 1398.
7. Cech, T.R., Rosenfeld, A. and Hearst, J.E. (1973) J. Mol. Biol. 81, 299.
8. Cech, T.R. and Hearst, J.E. (1976) J. Mol. Biol. 100, 227.
9. Horz, W., Hess, I. and Zachau, H.G. (1974) Eur. J. Biochem. 45, 501.
10. Walker, P.M.B. and McLaren, A (1965) J. Mol. Biol. 12, 394.
11. Southern, E.M. (1975) J. Mol. Biol. 98, 503.
12. Rigby, P.W.J., Dieckmann, M., Rhodes, C. and Berg, P. (1977) J. Mol. Biol. 113, 237.
13. Lasky, R.A. and Mills, A.D. (1977) Febs Lett. 82, 316.
14. Thuring, R.J.W., Sanders, J.P.M. and Borst, P. (1975) Anal. Biochem. 66, 213.
15. Blattner, F.R., Williams, B.G., Blechl, A.E., Denniston-Thompson, K., Faber, H.E., Furlong, L.-A., Grunwald, D.J., Kiefer, D.O., Moore, D.D., Schumm, J.W., Sheldon, E.L. and Smithies, O. (1977) Science 196, 161.
16. Becker, A. and Gold, M. (1975) Proc. Nat. Acad. Sci. (U.S.A.) 72, 581.
17. Tiemeier, D.C., Tilghman, S.M. and Leder, P. (1977) Gene 2, 173.
18. Grunstein, M. and Hogness, D.S. (1975) Proc. Nat. Acad. Sci. (U.S.A.) 72, 3961.
19. Southern, E.M. (1975) J. Mol. Biol. 94, 51.
20. Sinclair, J.H. and Brown, D.D. (1971) Biochemistry 10, 2761.
21. Krayev, A.S., Kramerov, D.A., Skryabin, K.B., Ryskov, A.P., Bayer, A.A. and Georgiev, G.P. (1980) Nucleic Acids Research 8, 2101.
22. Houck, C.M., Rinehart, F.P. and Schmid, C.W. (1979) J. Mol. Biol. 132, 289.
23. Wensink, P.C., Tabata, S. and Pachi, C. (1979) Cell 18, 1231.
24. Shen, C-K, J. and Maniatis, T. (1980) Cell 19, 379.
25. Cheng, S-M. and Schildkraut, C.L. (1980) Nucleic Acids Research, In press.