Volume 10 Number 23 1982 Nucleic Acids Research

Nucleotide sequence of satellite DNA contained in the eliminated genome of *Ascaris lumbricoides*

Fritz Müller, Philippe Walker*, Pierre Aeby, Heinz Neuhaus, Heinz Felder, Eduard Back and Heinz Tobler+

Zoologisches Institut der Universität Freiburg, Pérolles, CH-1700 Freiburg, Switzerland

Received 22 October 1982; Accepted 5 November 1982

ABSTRACT

Several restriction endonuclease fragments isolated from highly repetitive satellite DNA of the chromatin eliminating nematode Ascaris lumbricoides var. suum have been cloned. Each type of restriction fragment corresponds to a different variant of the same related ancestral sequence. These variants differ by small deletions, insertions and single base substitutions. Restriction and DBM blot analyses show that members of the same variant class are tandemly linked and therefore are physically separated from other variant classes. A comparison of all the determined sequences establishes a 121 bp long and AT rich consensus sequence. There is evidence for an internal short range periodicity of 11 bp length, indicating that the Ascaris satellite initially may have evolved from an ancestral undecamer sequence. The satellite DNA sequences are mostly but not entirely eliminated from the presumptive somatic cells during chromatin diminution. We have no evidence for transcriptional activity of satellite DNA at any stage or tissue analyzed.

INTRODUCTION

During early cleavage divisions of a variety of organisms a process called chromatin diminution can be observed: portions of chromatin are eliminated from the presumptive somatic cells thus leading to somatic cells which contain less genetic material than germ line cells (see 1-3, for review). The phenomenon, which was first described by Theodor Boveri (4) for Parascaris equorum, has since been found to occur in phylogenetically widespread species of nematodes, insects, crustaceans and hypotrichous ciliates (1-3). In Ascaris lumbricoides var. suum (= Ascaris suum), about a fourth of the total amount of DNA is expelled from the presumptive somatic nuclei (5,6).

In order to understand the significance of the elimination process at the molecular level, it is necessary to know the genetic informational content of the eliminated DNA. Earlier work has already shown that the eliminated DNA is enriched for repetitious DNA sequences (5-9). Recently it has been demonstrated by our group (10) and independently by Moritz and his collaborators

0305-1048/82/1023-7493 $2.00/0

(11) that the eliminated material contains highly repetitive satellite sequences. This satellite is composed of 120 bp repeating units, existing in different sequence variations, as demonstrated by restriction enzyme analysis (9-11). In the present work we have isolated, cloned and sequenced different monomer fractions in order to answer questions concerning the extent of sequence heterogeneity in this satellite, its organization within the germ line genome and whether or not it is transcribed in different tissues. Furthermore, we discuss quantitative aspects about the elimination of the satellite DNA sequences from presumptive somatic cells.

MATERIALS AND METHODS

Preparation of Ascaris DNA

Adult Ascaris lumbricoides worms were collected from infected pigs in the slaughterhouse (Micarna Ltd. Courtepin, Switzerland). Intestines from females were dissected in 10 mM Tris (HCl), pH 7.6, 0.14 M NaCl, 10 mM EDTA. Spermatids were isolated by following the method of Tobler et al. (5), oocytes were collected as described by Kuhn and Tobler (12). All these cells can be stored for months at -20° C.

Total high molecular weight DNA from spermatids and intestines was isolated according to the method of Gross-Bellard et al. (13). For DNA extraction, oocytes were suspended in 10 mM Tris (HCl), pH 8.0, 10 mM EDTA, 10 mM NaCl, 0.5% SDS, 100 ug/ml proteinase K (Böhringer) and homogenized in a Tenbroeck homogenizer. After incubation at 37° C for about 16h, the homogenate was extracted twice with phenol/chloroform (1:1; saturated with 0.3 M NaAc, pH 7.6). The aqueous phase was brought to 0.3 M NaAc, pH 7.6 and precipitated with ethanol. The DNA pellet was redissolved in 10 mM Tris (HCl), pH 7.6, 10 mM EDTA, 50 ug/ml of pancreatic RNase (Sigma, DNase free), and incubated for 2h at room temperature. The reaction was stopped by addition of SDS to a final concentration of 0.5% and 100 ug/ml proteinase K. After another phenol extraction followed by an ethanol precipitation, the DNA was further purified as described in Bellamy and Ralph (14).

Restriction Enzyme Digestion, Gel Electrophoresis and Southern Hybridization

Restriction enzymes have been used according to the manufacturer's instruction (New England Biolabs). The reactions were stopped by addition of V4 volume of dye mixture (10 mM Tris (HCl), pH 7.6, 10% glycerol, 20 mM EDTA, 0.3% bromphenol blue, 1.5% Ficoll 400). Samples were loaded on 1.5-2% agarose

slab gels and electrophoresed in 15 g glycine and 15 ml 1 N NaOH per liter. Gels were stained with 1 ug/ml ethidium bromide for 15 minutes and the DNA bands were photographed on Ilford HP5 under 254 nm ultraviolet light.

Transfer of DNA to DBM paper (Schleicher and Schuell, ABM-Form) was performed by the standard technique (15) modified by Wahl et al. (16). Transfer buffer was 0.2 M sodium acetate, pH 4.0. Hybridization of the blots was done in 10x Denhardt solution (17), 2xSSC, 1 mM EDTA, 0.1% sodium pyrophosphate, 1 mg/ml hering sperm DNA. The DBM paper was autoradiographed as described (18).

Preparation of the satellite DNA hybridization probe

Satellite DNA for use as a hybridization probe was obtained by digestion of total Ascaris spermatid DNA with Mbo I. The digestion products were separated on 2% low melting agarose (Sigma A-4018) tube gels (18 cm long x 2.7 cm diameter). The 120 bp satellite monomer was localized by staining with ethidium bromide and the corresponding part of the gel was cut out. Isolation of the DNA from the gel was performed following the modified method of Wieslander (19). After addition of V10 of the volume of 100 mM Tris (HCl), pH 7.8, 5 M NaCl, 10 mM EDTA, the gel slices were melted at 68° C for 10 minutes followed by an incubation at 37° C for another 10 minutes. The DNA was then extracted twice with phenol saturated with 0.3 M NaCl and 0.1% hydroxyquinoline, pH 7.6. After another extraction with chloroform/isoamyl alcohol (24:1), the DNA was recovered by ethanol precipitation. The isolated satellite Mbo I monomers were radioactively labelled at the 3' ends with α-^{32}P dNTP (3000 C/m mole, Amersham) and polymerase I (New England Biolabs) as described (20). For further purification the hybridization probe was denatured and renatured to Cot 0.1 and passed through a HAP column. Double stranded DNA was eluted, desalted over Sephadex G 75, ethanol precipitated and resuspended in TE buffer (10 mM Tris, pH 7.6, 1 mM EDTA).

Isolation and cloning of satellite restriction fragments

Total germ line DNA from spermatids was completely digested with different restriction enzymes and the products separated on low melting agarose gels. Satellite containing DNA fragments were recovered from the gels as described above. The Bam HI and Taq I fragments were ligated directly into the Bam HI site or the Cla I site, respectively of the pBR 322 vector. Each sample containing 250 ug of satellite monomers was mixed with 300 ug of CIP treated plasmid in a ligase buffer (50 mM Tris (HCl), pH 7.6, 10 mM $MgCl_2$, 50 ug/ul

BSA, 10 mM DTT, 140 uM ATP). The reaction was carried out with 0.5 Weiss units of T4 ligase (Böhringer, Mannheim) at 15° C over night. The mixture was used to transform E. coli HB 101 bacteria, and recombinants were screened by the filter hybridization method of Grunstein and Hogness (21). The Hinf I fragments were reacted first with Klenov polymerase in order to convert the 5' overhangs to flush ends. The reaction was carried out over night at 16° C in 50 mM Tris (HCl), pH 7.6, 10 mM $MgCl_2$, 10 mM DTT, 200 u M dNTP's and 2 units of DNA polymerase I (large fragment).

In order to clone blunt ended restriction fragments, they were joined simultaneously to Hind III and Eco RI-linkers (Collaborative Research; labelled with T4 polynucleotide kinase / Biolabs and γ-^{32}P-ATP / Amersham) in about 30 times excess of linker to DNA ends. The reaction was done over night at 16° C in 50 mM Tris (HCl), pH 7.6, 10 mM $MgCl_2$, 10 mM DTT, 1 mM ATP with 3 units of T4 ligase (Böhringer) and stopped at 65° C for 10 minutes. The DNA was then digested with 30 u of Hind III and Eco RI (Biolabs) for 60 minutes. Excess linkers were removed on a Sephadex G 75 column. An aliquot of the DNA was analyzed on a polyacrylamide gel. Radioactive fragments of the expected length were observed. Another aliquot of the DNA was ligated between the Hind III and Eco RI sites of pBR 322. Such recombinant clones are easy to sequence either from the Hind III or from the Eco RI site, respectively.

DNA sequencing analysis

Restriction fragments to be sequenced were labelled either at their 5' ends by γ-^{32}P-ATP (3000 Ci/m mole) and T4 polynucleotide kinase (Biolabs) as described in Clarkson et al. (22) or at the 3' ends with α-^{32}P-dNTP (3000 Ci/m mole) and DNA polymerase I (Biolabs) (20). They were then redigested, separated on low melting agarose gels and extracted as described above. Chemical sequencing was performed according to Maxam and Gilbert (23).

Saturation hybridization

Total DNA from spermatids, oocytes and intestines, as well as λ-DNA was sheared by ultrasonication to about 3 kb length (Landolt et al. in prep.). The DNA was then denatured and loaded on Millipore HAWG filters (13 mm diameter) by following the procedure of Gillespie and Spiegelman (24). Each filter was charged with 30 ng of Ascaris DNA from spermatids, oocytes or intestines, complemented with 500 ng λ-DNA as carrier. Control filters were loaded with 500 ng λ-DNA. Hybridization to the radioactively labelled DNA probes at different concentrations was done as described for Southern hybridization. The

radioactivity of the filters was determined in Instagel in a Packard Tri-Carb liquid scintillation counter.

Isolation of total nuclear RNA and hybridization to satellite DNA

Nuclei from larvae, intestines, spermatids, oocytes and oogonies have been isolated following the citric acid procedure described by Busch (25). Total RNA from these nuclei was extracted according to Roop (26). Upon glyoxylation (27), the RNA probes were subjected to electrophoresis through a 1.5% agarose gel in 10 mM sodium phosphate buffer at pH 6.8 (100 V, 6V4h). The separated RNA fragments were transferred from the agarose gel to DBM paper (15). After prehybridization for 24h in the buffer described (15), complemented with 25 mM sodium phosphate (pH 6.5) and 0.2% SDS, hybridization was carried out with 10% dextran sulfate (16) for 15h at 40° C. Cloned Bam HI and Taq I monomers were used as hybridization probes after end-labelling with DNA polymerase I as described above.

Following hybridization, the DBM paper was washed 3 times for 10 min in 2xSSC and 0.1% SDS at room temperature and 2 times for 20 min in 0.1xSSC and 0.1% SDS at 42° C. The DBM paper was then exposed using an intensifying screen (18).

RESULTS

Structure and Organization of the Satellite DNA

Satellite monomers and multimers were isolated from total digests of *Ascaris* germ line DNA with different restriction enzymes and cloned (partially with the help of Hind III and Eco RI linkers) into pBR 322. As expected, the *Ascaris* DNA was found to be inserted in both orientations relative to the vector. The unambiguous sequence of 12 cloned monomers could be established from both strands and are represented in Fig.1.

As expected from previous experiments, the lengths of these sequences were found to be quite homogeneous. They vary between 119 bp for the Hae III and 123 bp for the Alu I fragment. Moreover, each restriction fragment type corresponds to a different sequence variant, all of which can be derived from the same common ancestor sequence. This is in perfect agreement with the hybridization data (10). The examined repeats differ by single base substitutions and small insertions and deletions (cf. Fig.1). However, there is always the possibility that molecular cloning has introduced some of the observed variations. The data presented suggest that this is very unlikely for

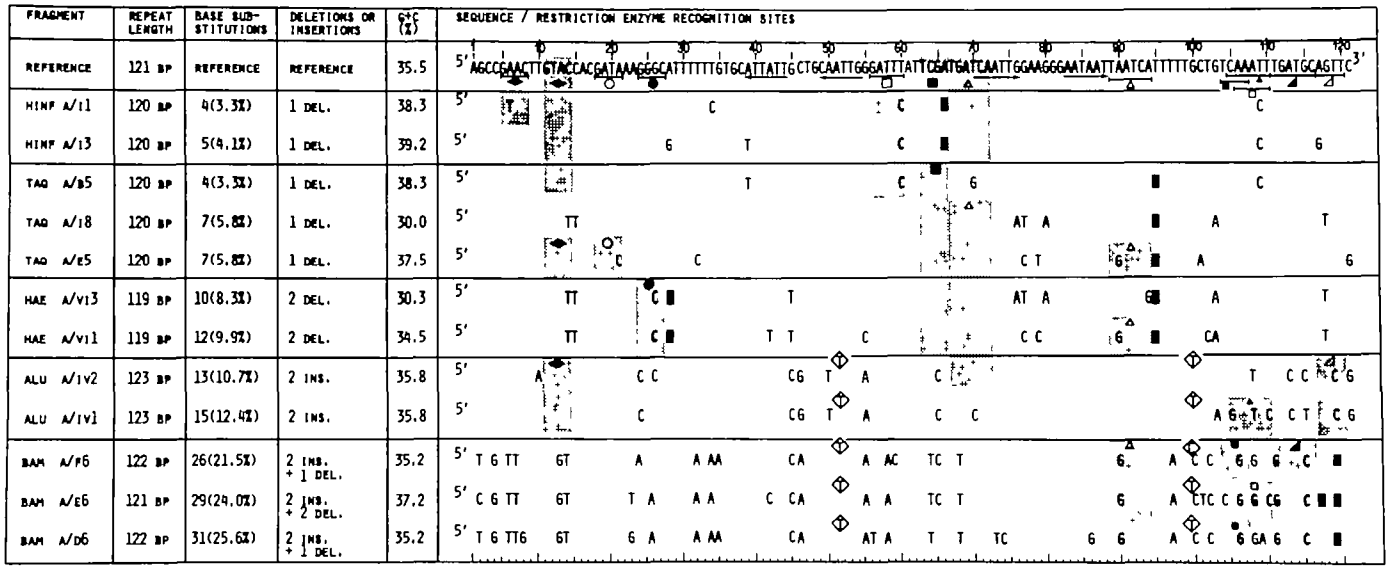

Fig. 1. Consensus sequence (on top) and sequences of 12 individual clones of highly repetitive germ line DNA of Ascaris lumbricoides. The reference sequence represents the consensus sequence deduced from all sequenced DNA fragments. The different sequences are listed in order of decreasing homology. Base substitutions relative to the reference sequence are indicated by the respective bases, deletions by ▬ and insertions by ◇ . Arrows underlining the reference sequence designate small palindromic sequences. The restriction enzyme recognition sites are ◆ Rsa I, ○ Mbo I, ● Hae III, □ Hinf I, ■ Taq I, △ Bcl I, ▲ Eco RI, ◢ Bam HI, ◿ Alu I; they are indicated by stippled areas.

the following two reasons: (i) The length of the cloned repeats corresponds to the length of the uncloned material. (ii) Most of the alterations in sequence among our sequenced repeats are found to be clustered at the same position in several, but not all clones. Fig. 1 also shows some variations in the G+C content of the different monomers, ranging from 30.0% to 38.3%. Nevertheless, all these data indicate that the satellite under investigation is very rich in A+T.

A consensus sequence for all the analyzed repeats has been established and is listed on top of Fig. 1. This consensus sequence is 121 bp long and has a G+C content of 35.5%. The analyzed monomers, grouped into variant classes (which are defined by their common flanking restriction sites), were compared with this consensus sequence and are listed in order of decreasing homogeneity. The smallest difference to the consensus sequence is found within the Hinf I fragments: In the Hinf A/I monomer only 3 base pair changes and 1 deletion are noted. The Bam variants on the other hand show the largest difference: they have up to 31 base substitutions (25.6%), 2 insertions and 1 deletion. All these changes between the different variant classes seem to be scattered more or less randomly over the whole repeat rather than being localized on preferential sites.

Our data show that variations in sequence also occur between fragments of the same variant class. These changes are rather small compared to those between some variant classes, so that the members of the same class are generally very similar to each other. An important question regarding the structure of the satellite is the extent to which neighbouring repeat units are structurally related. We therefore sequenced a cloned Hinf I dimer and a Hae III pentamer (cf. Fig. 2). The structure of these multimeric fragments resembles very much the sequenced monomers of the respective variant classes. Furthermore, the results show that immediately adjacent repeating units diverge slightly from each other by a few base changes, which are scattered over the whole repeat length. The random distribution of these point mutations indicates that there is no strong selective pressure for any particular region within the sequence of this satellite. The present data furthermore show that such occasional mutations in the Hinf I and Hae III recognition sites within the dimer and the pentamer sequences, rendered the recognition sites inaccessible to the effect of the appropriate restriction enzymes. Such random mutations, rather than DNA modifications, seem to be responsible for the for-

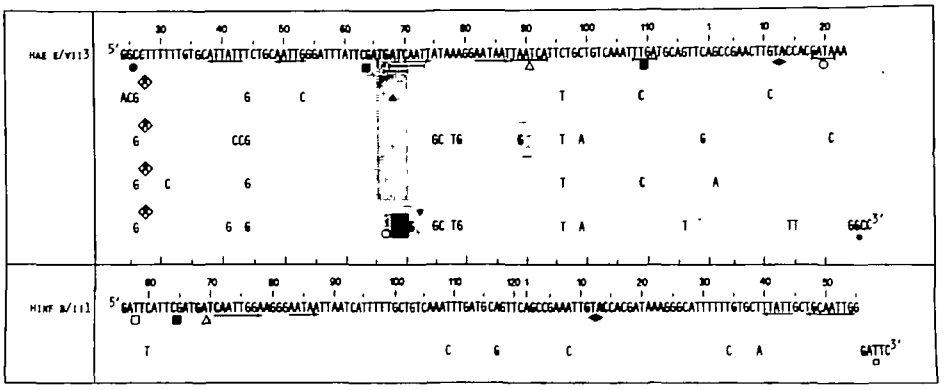

Fig. 2. Sequence of a cloned Hae III pentamer and a Hinf I dimer isolated from highly repetitive germ line DNA of Ascaris lumbricoides. In the pentamer sequence, only the first 119 bases are fully depicted, whereas in the following 4 monomeric sequences merely the base changes relative to the first sequence are indicated. Likewise, only the first 121 bases are completely represented for the Hinf I dimer. Base substitutions relative to the first monomeric sequence are indicated by the respective bases, insertions by ◇ . Arrows underlining the reference sequence designate small palindromic sequences. The restriction enzyme recognition sites are ● Hae III, ■ Taq I, △ Bcl I, ○ Mbo I, ▼ Cla I, ◀▶ Rsa I, □ Hinf I; they are indicated by stippled areas.

mation of most if not all multimeric fragments in a total digest of Ascaris satellite DNA (Fig. 4 and Ref. 10). Further evidence comes from the fact that several cloned multimers remain totally resistent to digestion with the appropriate restriction enzyme (results not shown), even though E. coli would not be expected to modify DNA in the same way as Ascaris.

A computer analysis of the consensus sequence reveals the presence of a dyad symmetry, which is localized between bases 39 and 88. This region is able to form a hairpin structure with two basepaired regions, each of which being 6 bp long (see Fig. 1). Interestingly, the same dyad symmetry is also present in slightly altered form at the same position in all our sequenced monomers.

On the other hand we have found that the consensus sequence can be subdivided into 11 subrepeats, which all seem to be variants of the same sequence motif (Fig. 3). The prototype sequence which can be deduced is an undecanucleotide with the sequence 5'GCA($\frac{T}{A}$)TT($\frac{T}{G}$)TGAT. The same motif is found, with some variations, in all our sequenced fragments. The first 6 nucleotides of this undecanucleotide are significantly conserved throughout all the se-

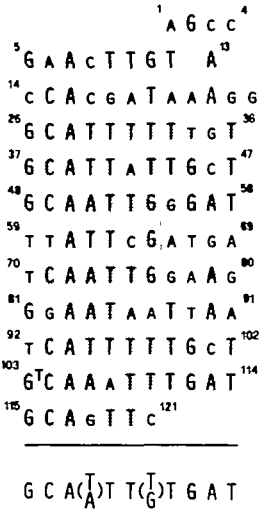

Fig. 3. The 121 bp long consensus sequence subdivided into 11 subrepeats and arranged in order to demonstrate the highest degree of sequence homologies. The undecanucleotide prototype sequence is listed below. Regions of homologous bases are indicated by stippled areas and by larger capital letters.

quenced fragments, while the other 5 seem to be more variable. Based on our findings we postulate that the Ascaris satellite initially has evolved from an ancestral variant of this 11 bp long prototype sequence.

We have also analyzed our sequences with respect to the distribution of several restriction enzyme sites which are indicated in Fig. 1. None of them is present in all the fragments. Random point mutations occurred at several sites, rendering the respective repeat fragment insensitive to the action of the corresponding restriction enzyme. Nevertheless, some of these enzymes cut quite often within every repeat: Mbo I and Bcl I have one or even several restriction sites in nearly all of the fragments, Taq I in about 50% of them. As an example, the restriction enzyme Mbo I cleaves the Taq A/E5 fragment in not less than 3 positions, namely 18, 68 and 90.

If the sequences of our fragments are representative for the structure of the whole Ascaris satellite, we would expect to find a similar distribution of recognition sites for the different restriction endonucleases within the total chromosomal satellite. That this is the case has been verified by digesting aliquots of total germ line DNA with these enzymes. The digestion products were separated by agarose gel electrophoresis and transferred to DBM paper, which was then hybridized to a radioactively labelled satellite probe. The autoradiograph is presented in Fig. 4. Slot b, c and d show total digests of satellite DNA with the enzymes Bcl I, Mbo I and Taq I, respectively. In agreement with our expections, all 3 enzymes cut the satellite to a high de-

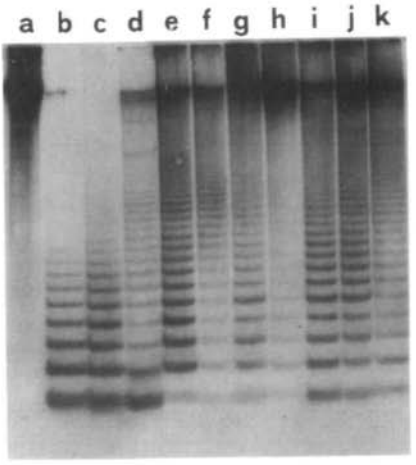

Fig. 4. Hybridization pattern of total satellite DNA from Ascaris oocytes, digested with different restriction enzymes. 3 ug aliquots of total oocyte DNA were digested with Bcl I (b), Mbo I (c), Taq I (d), Hinf I (e), Alu I (f), Hae III (g), Bam HI (h), Bam HI and Hae III (i), Hae III and Alu I (j), Alu I and Bam HI (k). The restriction fragments were separated on a 2% agarose slab gel, transferred to DBM paper and hybridized with a labelled satellite DNA probe. For comparison, slot (a) shows an aliquot of total undigested oocyte DNA.

gree. Some other restriction enzymes have only a few recognition sites within the sequenced monomers, among them Alu I, Hae III and Bam HI. These three enzymes are of special interest because each of them cuts only one distinct variant class (cf. Fig. 1). Again, we have tested whether our findings made on the 12 sequenced monomers are also valid for the whole satellite, by digesting total chromosomal satellite DNA with these 3 restriction endonucleases. The results of this experiment are presented in Fig. 4, lanes f-k. As expected, each of these enzymes degrades total satellite DNA only partially, leading to a whole ladder of regularly spaced fragments corresponding to the 120 bp monomer and its multimers (Fig. 4, slots f, g and h). This indicates that the variant copies, cleaved by each of these enzymes, are clustered in tandem array within the Ascaris genome. The resistent, nondigested part of the satellite DNA has no or only a few cuts for the respective restriction enzyme, since it remains as high molecular weight material on top of the gels. Moreover, cross digestions with Bam HI/Hae III (slot i), Hae III/Alu I (slot j) and Alu I/Bam HI (slot k) do not lead to the appearance of new fragments others than those occurring in single digests of the respective restriction enzymes. Each of these enzymes must therefore cleave in different areas or domains of the satellite. Therefore, the conclusion can be drawn that the Ascaris satellite DNA is composed of an unknown number of sets of variant classes (of which we have analyzed five by random selection). Members of the same variant class are arranged as tandemly clustered repeating units in long

Fig. 5. Hybridization pattern of Hae III digested total DNA from oocytes (a) and intestines (b) of Ascaris lumbricoides. The restriction fragments were separated on a 2% agarose slab gel, transferred to DBM paper and hybridized with a labelled Taq I monomer fragment.

blocks (domains) within the genome, physically separated from other variant classes.

Amount of Satellite DNA in Germ Line and Somatic Cells

The bulk of satellite DNA sequences is clearly eliminated from the somatic genome of Ascaris lumbricoides. This has been demonstrated by hybridizing labelled Taq I monomer fragments to DBM-blots of Hae III digested germ line and somatic DNA (Fig. 5). However, weak signals are still detectable in the track of somatic DNA. Since this material has been isolated from carefully prepared intestines, it must be of pure somatic origin without any contamination of germ line DNA. The above hybridization data therefore demonstrate that some copies of satellite DNA are retained in the somatic genome during the process of chromatin elimination.

In order to quantify the amount of the eliminated satellite DNA, saturation hybridization experiments with radioactively labelled Mbo I monomers were performed. Fig. 6 shows the results. From the difference in the saturation values we have to conclude that chromatin diminution expels over 99.5% of the satellite DNA sequences from the presumptive somatic cells. For this

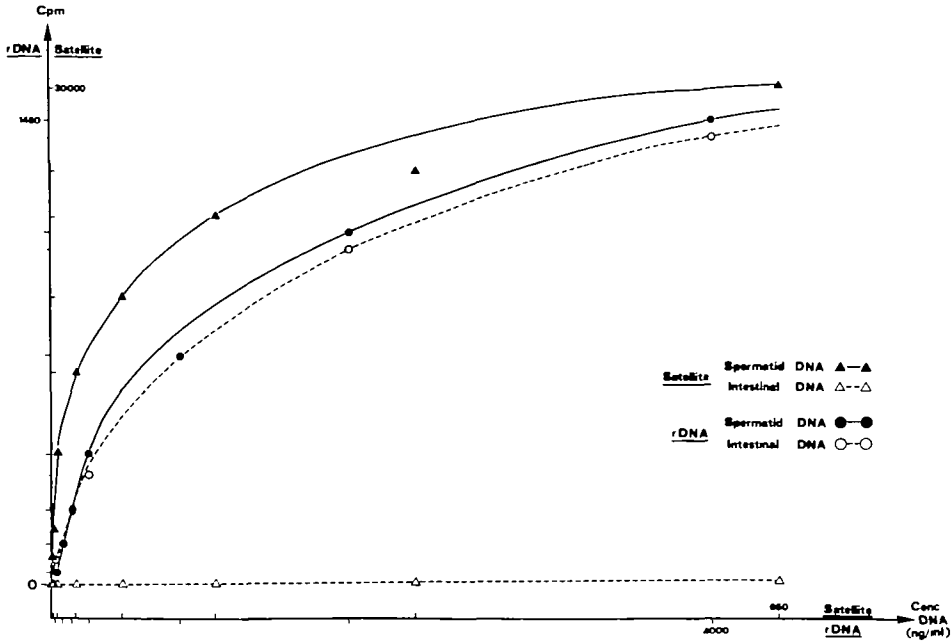

Fig. 6. Saturation hybridization of the Mbo I monomeric fragment to spermatid (▲—▲) and intestinal (△--△) DNA of Ascaris lumbricoides. As an internal control, ^{32}P-labelled cloned 18S and 28S rDNA from Ascaris lumbricoides were hybridized to the same filters loaded either with 30 ng of spermatid (●—●) or with 30 ng of intestinal (○--○) DNA.

experiment an internal hybridization control was used: filters loaded with germ line and somatic DNA were also hybridized with radioactively labelled cloned Ascaris 18S and 28S rDNA. As Fig. 6 demonstrates, the saturation plateaus are identical for spermatid and intestinal DNAs.

Satellite DNA is not transcribed

The question whether or not satellite DNA is transcribed in different tissues has been investigated by hybridizing radioactively labelled Taq I and Bam HI cloned monomers of satellite DNA to total nuclear RNAs isolated from larvae, intestines, spermatids, oogonies and oocytes. These different nuclear RNAs, as well as 18S and 28S rRNA, pBR 322/Hinf I and a cloned monomer Bam F 6 as an internal standard were denatured, electrophoresed through a 2% agarose gel, transferred to DBM paper and hybridized with the satellite DNA probe. As Fig. 7 clearly demonstrates, neither total nuclear RNA nor 18S and 28S rRNA from Ascaris show any hybridization with the labelled monomer probe, al-

Fig. 7. Total nuclear RNAs were isolated from larvae (L), intestines (Int), spermatids (Sp), oocytes (Ooc) and oogonies (Oog) of Ascaris lumbricoides. 5 ug aliquots of these different RNAs, as well as 0.5 ug of 18S and 28S rRNA, 50 ng of pBR 322/Hinf I together with 50 ng of the clone Bam F 6 (corresponding roughly to 1.25 ng of the cloned Bam HI monomer) as an internal standard were run on a 2% agarose slab gel, transferred to DBM paper and hybridized with labelled Taq I 8 and Bam F 6 cloned monomers. Hybridization of the labelled probe is only detectable with the internal standard.

though hybridization with the internal standard (Bam F 6 monomer), which corresponds to 0.025% of the nuclear RNA in each slot, gives a clear signal. In a reciprocal experiment, satellite DNA was bound to a filter and hybridized with labelled nuclear RNAs isolated from different tissues and stages. Again, no hybridization could be detected (results not shown). We therefore conclude that satellite DNA of Ascaris lumbricoides is either not transcribed in any of the tested cells, or that the transcripts are so few that they go undetected by these experimental methods.

DISCUSSION

We have cloned satellite monomers resulting from digests of Ascaris germ line DNA with the restriction enzymes Alu I, Bam HI, Hae III, Hinf I and Taq I, respectively. A few recombinants of each fragment type were sequenced. Due to their random selection, they are expected to be a representative sample for the most abundant sequence of the respective fragment type. As our results show (Fig. 1), each fragment type corresponds to a different variant of the same related ancestral sequence. Sequence diversity between these va-

riant classes can amount up to 30% base substitutions due to single base substitutions, insertions or deletions. In contrast, members of the same variant class seem to be quite homogeneous to each other, apart from some microheterogeneities which result from point mutations. However, because of the limited data it is impossible to make more precise statements about the degree of possible sequence variations within a given variant class.

It should be pointed out that the DNA was extracted from a great number of animals, so that the total variation within an individual cannot be accurately determined. However, repeating units which lie immediately adjacent to each other diverge slightly by randomly distributed point mutations. This is exemplified by the cloned Hinf I dimer and the Hae III pentamer (cf. Fig. 2) as well as by the existence of multimeric fragments within total digests of satellite DNA with any restriction enzyme, which are thought to be due to occasional mutations of the internal recognition sites (see results). Thus, some sequence variation must already be present on a single chromosome.

All the differences between variant classes and between individual repeats of the same class are the result of single base substitutions as well as small deletions and insertions, which are scattered more or less randomly over the whole repeating unit. This indicates that there is no strong selective pressure for any particular region within the satellite monomer sequence.

The total amount of the five analyzed variant classes within the whole Ascaris satellite can be estimated from the frequency of the recognition sites of the appropriate restriction enzymes. Each of the enzymes Hinf I, Alu I, Hae III and Bam HI have a recognition site in about 5 to 10% of the total satellite repeats, as evaluated from Southern blots of total germ line DNA (see Fig. 4 and Ref. 10). Thus, these four variant classes together may already be representative for at least 20% of the whole satellite DNA. Taq I is an even more abundant cleaving enzyme, which is able to degrade about 50-70% of the total satellite DNA. However, our sequencing data clearly indicate that Taq I cleaves within several variant classes simultaneously, since Taq restriction sites are not only found in Taq I but also in Hae III and Bam HI monomers (cf. Fig. 1). The monomer fraction of the Taq I digest must therefore contain a mixture of several sequence variant classes. Nevertheless, all three of our randomly selected and sequenced Taq I monomers belong to the same sequence variant, namely the Taq I variant class. This shows that these Taq I variants are at least an important, if not the predominant fraction of

the Taq I monomers, and consequently also of the whole satellite. However, it is impossible to make more precise statements due to the relatively small number of individual clones we have analyzed.

Altogether, the digestion products of the five enzymes used represent an important fraction of the whole satellite DNA. The presented sequences and especially the thereof derived consensus sequence (Fig. 1) should therefore be representative for most of the entire satellite. This view is further supported by the fact that those enzymes which have a recognition site in nearly all our sequenced fragments (Bcl I, Mbo I, Taq I) do also cleave the whole native satellite abundantly (cf. Fig. 4b-d). Additional evidence comes from a recent work of Streeck et al. (11). Independently this group has determined the prototype sequence of two major variants of Ascaris satellite, starting out from uncloned material. Since they used restriction enzymes which have many recognition sites within the satellite for the preparation of their sequencing fragments, their sequences are representative for an important portion of the satellite. However, this method is not suitable to demonstrate the diversity between and within different variant classes as it is possible by molecular cloning and analysis of individual repeats. Nevertheless, the sequences of Moritz's group (11) are in good agreement with our data: one of their published variants is very similar but not identical to our consensus sequence, whereas the second variant can be arranged among our Bam fragments.

A comparative analysis of our consensus sequence has provided evidence for an internal short range periodicity of 11 bp length (Fig. 3). The prototype sequence of this undecamer, as deduced from the consensus sequence is 5'GCA($\frac{T}{A}$)TT($\frac{T}{G}$)TGAT. This motif, with variations, is present in tandem array in all our analyzed monomers. Moreover, each variant class is tandemly clustered on the chromosome, physically separated from other variants. Such an arrangement as well as the existence of an undecamer within the basic units argues strongly for the evolution of the Ascaris satellite from a common ancestral sequence by the mechanism of sequence divergence followed by an amplification mechanism, as has been proposed already for other satellites (28-30, 32). In a first step, the 120 bp segments evolved from the ancestral undecamer sequence. Later on during evolution, the 120 bp fragment diverged into the different variants, followed by amplification to form the large tandem arrays that are observed in the present Ascaris genome.

Preliminary results from in situ hybridization experiments indicate that

the satellite DNA sequences are localized on most but not all chromosomes. Moreover, we have evidence that the individual chromosomes contain different amounts of satellite DNA. There is no correlation between the size of the chromosomes and their amount of satellite DNA sequences (Felder et al., in preparation).

The most important feature of Ascaris satellite DNA is the fact that its quantity differs to a great extent between germ line and soma cells due to the chromatin elimination process. We estimate this satellite to represent about 20% of the germ line genome, corresponding to about 10^6 copies of the 120 bp basic repeat. This value is based on measurements of the DNA quantity within the monomer band of a total germ line digest by the restriction enzyme Mbo I (unpublished results). Somatic cells on the other hand contain only a few satellite DNA sequences. At least 99.5%, but not all of the satellite copies are expelled from the presumptive somatic cells, so that somatic DNA contains at most 0.14% of satellitic sequences (roughly 5000 copies). We think that the satellite must account for most, if not all of the eliminated, highly repetitive material.

Turning now to the possible functions of this Ascaris satellite DNA, there is no experimental evidence that it is transcribed, either in germ line cells (oogonies, oocytes and spermatids) or in somatic tissues (intestines and larvae) (Fig. 7). This is not surprising because in general, highly reiterated sequences in eucaryotes are not thought to be transcribed. However, there are a few exceptions: transcription of satellite DNA sequences has been reported from HTC rat tissue cells (33) and from specific loops of lampbrush chromosomes in oocytes of newts (34-37). In Notophtalmus viridescens, transcription may initiate at promotor sites for histone genes which are embedded within variable stretches of satellite DNA sequences, and then proceed into the downstream satellite DNA sequences (35, 36). Other functions that have been suggested for satellite DNA, not involving its transcription, are mainly based on its general association with heterochromatin (for review see 28-32). The available data are consistent with the notion that satellite DNA exerts its biological effects in processes that are specific for germ line cells. Our present results support this idea, since chromatin elimination in Ascaris lumbricoides demonstrates clearly that the bulk of the satellite DNA is not needed in somatic cells. Therefore, if this satellite DNA has any function at all, it must be germ line limited. Otherwise one would

have to assume that the same function could be exerted in somatic tissues by the few remaining copies, which seems to be rather unlikely.

ACKNOWLEDGEMENTS

We would like to thank Sonja von Niederhäusern, Marianne Lauper, Hubert Gachoud and Zbigniew Kasprzycki for excellent technical assistance, Dr. S.G. Clarkson for critical comments on this manuscript and Magdalena Tobler for typing. This work was supported by grants No. 3.443-0.79 and 3.233-0.82 from the Swiss National Science Foundation.

*Present address: Institut de Biologie animale, Université de Lausanne, Place de la Riponne, CH-1000 Lausanne, Switzerland.

+To whom correspondence and reprint requests should be addressed.

REFERENCES

1. Beams, H.W. and Kessel, R.G. (1974) Internat. Review Cytol. 39, 413-479.
2. Pavan, C., da Cunha, A.B. and Sanders, P. (1975) in Handbook of Genetics, King, R.C. Ed., Vol. 3, pp. 207-256, Plenum Press, New York.
3. Ammermann, D. (1982) in DNA and Evolution: Natural Selection and Genome Size, Cavalier-Smith, T. Ed., John Wiley and Sons, New York (in press).
4. Boveri, T. (1887) Anat. Anz. 2, 688-693.
5. Tobler, H., Smith, K.D. and Ursprung, H. (1972) Develop. Biol. 27, 190-203.
6. Moritz, K.B. and Roth, G.E. (1976) Nature (London) 259, 55-57.
7. Goldstein, P. and Straus, N.E. (1978) Exp. Cell Res. 116, 462-466.
8. Roth, G.E. (1979) Chromosoma 74, 355-371.
9. Roth, G.E. and Moritz, K.B. (1981) Chromosoma 83, 169-190.
10. Müller, F., Walker, P., Aeby, P., Neuhaus, H., Back E. and Tobler, H. (1982) in Embryonic Development, Part A: Genetic Aspects, Burger, M.M. and Weber, R. Eds. pp. 127-138, Alan R. Liss, New York.
11. Streeck, R.E., Moritz, K.B. and Beer, K. (1982) Nucl. Acid Res. 10, 3495-3502.
12. Kuhn, O. and Tobler, H. (1978) Biochim. Biophys. Acta 521, 251-266.
13. Gross-Bellard, M., Oudet, P. and Chambon, P. (1973) Eur. J. Biochem. 36, 32-38.
14. Bellamy, A.R. and Ralph, R.K. (1968) in Methods in Enzymology, Grossman, L. and Moldave, K. Eds. Vol. XII, B, pp. 156-160.
15. Alwine, J.C., Kemp, D.J. and Stark, G.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5350-5354.
16. Wahl, G.M., Stern, M. and Stark, G.R. (1979) Proc. Natl. Acad. Sci. USA 76, 3683-3687.
17. Denhardt, D.T. (1966) Biochem. Biophys. Res. Commun. 23, 641-646.
18. Laskey, R.A. and Mills, A.D. (1977) FEBS letters 82, 314-316.
19. Wieslander, L. (1979) Anal. Biochem. 98, 305-309.
20. Müller, F. and Clarkson, S.G. (1980) Cell 19, 345-353.

21. Grunstein, M. and Hogness, D. (1975) Proc. Natl. Acad. Sci. USA 72, 3961-3965.
22. Clarkson, S.G., Kurer, V. and Smith, H.O. (1978) Cell 14, 713-724.
23. Maxam, A.M. and Gilbert, W. (1980) in Methods in Enzymology, Grossman, L. and Moldave, K. Eds. Vol. 65, pp. 499-560.
24. Gillespie, D. and Spiegelman, S. (1965) J. Mol. Biol. 12, 829-842.
25. Busch, H. (1967) in Methods in Enzymology, Grossman, L. and Moldave, K. Eds. Vol. XII A, pp. 421-448.
26. Roop, D.R. (1978) Cell 15, 671-685.
27. McMaster, G.K. and Carmichael, G.G. (1977) Proc. Natl. Acad. Sci. USA 74, 4835-4838.
28. Brutlag, D.L. (1980) Ann. Rev. Genet. 14, 121-144.
29. John, B. and Miklos, G.L.G. (1979) Int. Rev. Cytol. 58, 1-114.
30. Singer, M.F. (1982) Int. Rev. Cytol. 76, 67-112.
31. Hilliker, A.J., Appels, R. and Schalet, A. (1980) Cell 21, 607-619.
32. Miklos, G.L.G. and Gill, A.C. (1982) Genet. Res. Camb. 39, 1-30.
33. Sealy, L., Hartley, J., Donelson, J., Chalkley, R., Hutchison, N. and Hamkalo, B. (1981) J. Mol. Biol. 145, 291-318.
34. Varley, J.M., MacGregor, H.C. and Erba, H.P. (1980) Nature 283, 686-688.
35. Varley, J.M., MacGregor, H.C., Nardi, I., Andrews, C. and Erba, H.P. (1980) Chromosoma 80, 289-307.
36. Diaz, M.O., Barsacchi-Pilone, G., Mahon, K.A. and Gall, J.G. (1981) Cell 24, 649-659.
37. Gall, J.G., Stephenson, E.C., Erba, H.P., Diaz, M.O. and Barsacchi-Pilone, G. (1981) Chromosoma 84, 159-171.