Volume 16 Number 18 1988 Nucleic Acids Research

Illegitimate recombination in the histone multigenic family generates circular DNAs in *Drosophila* embryos

Geneviève Pont+, Fabienne Degroote and Georges Picard*

Laboratoire de Génétique, UA 360 CNRS, Université Blaise Pascal-Clermont-Fd. II, 63177 Aubière cedex, France

Received June 10, 1988; Revised and Accepted August 16, 1988

ABSTRACT

From extrachromosomal covalently closed circular DNA molecules purified from Drosophila melanogaster embryos, we have isolated 24 clones homologous to the histone tandemly repeated gene family. Some of the clones harbor one of the two main types of genomic repeated units of 4.8 and 5.0 kb. and probably result from homologous recombination. The remaining clones have a size ranging from 0.2 to 2.5 kb. and most of them carry a single fragment of the repeated unit. Nucleotide sequences of the junction region of six of these clones indicate they are generated by illegitimate recombination between short (8-15 bp.) imperfect direct repeats. The data suggest that most of the histone homologous circular DNA molecules are deleted histone units.

INTRODUCTION

Extrachromosomal covalently closed circular DNA molecules (ccc DNAs) have been found in a large variety of eucaryotes. In cultured cells of primates, rodents, chicken and Drosophila, they constitute a population of molecules highly heterogeneous both in size and sequence (1-5). Ccc DNAs also have been observed in whole organisms including mouse and Drosophila embryos (2,6,7) and plants (8,9). Several lines of evidence indicate that most of the ccc DNAs are homologous to repeated genomic sequences (10-15).

Most of the circular molecules which have been identified to date are homologous to various types of mobile elements and some are potential intermediates of transposition (16-20). In addition to this major class, several non-coding sequences have been detected in the form of ccc DNAs, such as alphoide satellite DNA in primates (21-24), T cells receptor genes sequences resulting from functional somatic rearrangements in mouse thymocytes (25,26) and tandemly repeated sequences of the non-transcribed spacer of rDNA genes in Drosophila embryos (7). Some coding sequences, such as rDNA genes in

Xenopus oocytes (27,28) and in testis of Drosophila males undergoing magnification (29), histone DNA sequences in Drosophila (30) and TS-DHFR genes in methotrexate-resistant lines of Leishmania (31) were also identified in circular DNAs, and shown to be associated with amplification. Recently it was also demonstrated that members of tandemly repeated multigenic families such as rDNA in Pisum sativum (32), 5 S and histone genes in Drosophila melanogaster embryos (7) can be normal components of the ccc DNA populations.

There are about 100 tandemly repeated histone units in the genome of D.melanogaster. Each unit contains one copy of each of the five histone genes, separated by non-transcribed spacers (Figure 1). Two types of units were described which differ mainly by an insertion of about 200 bp. in a spacer region (33-35). In our previous report, we showed that the two major types of histone units are present among the extrachromosomal circular DNA molecules purified from D.melanogaster embryos of various origins.

In this paper, we report on the molecular cloning and analysis of histone circular molecules from embryos laid by dysgenic F1 (SF) females produced in the I-R system of hybrid dysgenesis (36). SF females display reduced fertility resulting from the death of a variable proportion of their eggs during early embryogenesis. This material was nevertheless choosen because it contains more ccc DNAs than embryos from the parental inducer and reactive strains (unpublished data). We show that a large proportion of the circular molecules homologous to the histone multigenic family are shorter than a genomic repeat. Nucleotide sequences of the junction fragments indicate that these ccc DNAs are generated by a recombinational event involving short directly repeated sequences.

MATERIALS AND METHODS

Extrachromosomal circular DNAs from embryos. Embryos were obtained from F1 (SF) dysgenic females that resulted from mass crosses between females of the reactive stock Charolles and males from the inducer stock Canton-S. They were collected approximately every 12 hours, dechorionated in 12% Sodium Hypochlorite, rinsed in 0.7% NaCl, 4% Triton X 100 and thoroughly washed with 0.7% NaCl. After being blotted dry, the embryos were frozen in liquid nitrogen and kept at -70°C. Extrachromosomal circular DNAs were purified as

previously described (7). From 50 g. of embryos it is usually possible to recover 15 to 30 ug. of circular DNAs, of which about half of the molecules are mtDNA.

DNA probes, labeling, hybridizations. pDmHH8 (7) contains a 4.8 kb. genomic histone repeat cloned into the Hind III site of plasmid pUC8 (37). pDmtHC and pDmtHD (38) respectively carry, in pBR322, the 4.7 kb. "C" and the 0.43 kb. "D" fragments generated by a Hind III digest of D. melanogaster mitochondrial DNA (39). Plasmid DNAs were purified as described in (40) and 32P-labeled by nick-translation (41) to a specific activity of 10 dpm/ug. Hybridizations were performed for 36 h. at 40°C in 50% (v/v) formamide, 1 x Denhart's solution, 0.6 M NaCl, 60 mM Sodium Citrate and 50 ug. / ml. of denatured calf thymus DNA.

Cloning of ccc DNAs in pUC18. Two experiments were done. In the first one, 7 ug. of a ccc DNA preparation which contained about 10% of contaminating linear molecules, as estimated by electron microscopy, were heated at 100°C for 5 min. and then treated with Mung-bean nuclease, according to the procedure previously reported (7). Ccc DNAs were digested by Bam HI and ligated by T4 DNA-ligase at 10°C. overnight in 10 ul. to 1 ug. of Bam HI linearized pUC18 plasmid DNA (42). Transformation of the bacterial strain NM522 (43) according to the Hanahan's method (44) yielded 2 x 10^3 recombinant clones which were screened using as a probe the 4.8 kb. Hind III fragment inserted in pDmHH8. Two clones giving a positive signal were recovered and named pDmcHB7 and 8.

In the second experiment, another preparation of ccc DNAs was used which was obtained after three successive CsCl-Ethidium Bromide equilibrium density gradient centrifugations instead of two. Because less than 1% of contaminating linear molecules were observed, most of them showing a size lower than 1 kb., the ccc DNAs were heat denatured but no Mung-bean nuclease treatment was done. 4 ug. were digested by Xho I and ligated as previously described to 1.5 ug. of pUC18 cut by Sal I. The ligation mixture was then digested by Sal I to linearize the non recombinant plasmids. Transformation and screening were carried out as in the first experiment. Of the 10^3 recombinant clones obtained, six were recovered which hybridize with the pDmHH8 insert. They were named pDmcHX1, 3.1, 3.2, 4, 6, and 7.

Cloning of ccc DNAs in lambda NM1149. A 1 ug. sample of the ccc DNA preparation used in the second experiment was digested by Hind III after heat denaturation and ligated with 4.5 ug. of Hind III-digested λNM1149 DNA (45) using T4 DNA-ligase, in a total volume of 100 ul., at 10°C. overnight. Approximately 3 x 10^5 recombinant clones were selected by plating on the NM514 bacterial strain (45), and 17 of these clones were found to hybridize to the pDmHH8 probe. They were named λDmcHH.

DNA sequencing. DNA sequences were determined by the dideoxy-chain termination method (46). Nucleotide sequences of junction fragments of pDmcHX3.2, pDmcHX4 and pDmcHX6 were determined after subcloning in M13mp19 while junctions between histone and non-histone DNA of pDmcHX3.1 were determined after subcloning in M13mp18 (42). The DNA sequence of the junction fragment of pDmcHX7 was obtained directly from the pUC18 clone, using an appropriate synthetic oligonucleotide (47). Most of the nucleotide sequence of the 4.8 kb genomic repeat was determined by Goldberg (34) and Lifton (35). We determined the sequence from nucleotides 98 to 220, 324 to 424 and 2427 to 2579 using pDmHH8 as a template and the appropriate oligonucleotides. Nucleotide coordinates used in this paper are from Goldberg (34).

RESULTS

Cloning of ccc DNAs homologous to the histone multigenic family.

Three successive cloning experiments were carried out. In each experiment, the ccc DNA preparation was first heated at 100°C. for 5 min. to denature the linear contaminating DNA molecules. In one experiment a treatment with Mung-bean nuclease was performed to eliminate single stranded DNA. This was done to prevent the linear molecules from being cloned (see Mat. & Met.). The restriction enzymes used for cloning (Bam HI, Xho I and Hind III) were choosen because they cut only once in both types of genomic histone DNA repeats (Figure 1).

The first two experiments were performed using the plasmid vector pUC18 and the restriction enzymes Bam HI and Xho I, respectively. The 19.5 kb. mtDNA molecules are abundant in the circular DNA extracts, however, as D. melanogaster mtDNA does not contain a Bam HI site or only a single Xho I site, there is a very low probability that clones of these molecules will be recovered in

pUC18. In the first experiment, two clones hybridizing to the purified Hind III fragment of pDmHH8, which carries a 4,8 kb. histone unit, were isolated and named pDmcHB. The second experiment yielded six positive clones named pDmcHX. Restriction analysis of the eight clones indicates that they carry inserts ranging from 0.6 to 2.5 kb., smaller than either of the two types of histone DNA repeats previously detected among the ccc DNAs by Southern blot analysis (7).

In order to clone circular molecules corresponding in size to whole histone DNA repeats, we carried out a third experiment in which ccc DNAs digested by Hind III were cloned in the bacteriophage λNM1149. Seventeen clones that hybridized to pDmHH8 were recovered and named λDmcHH. Six of these clones carry a 5 kb. insert, four a 4.8 kb. insert and the seven remaining clones have inserts ranging in size from 0.2 to 1.7 kb. In this experiment, it was possible to estimate the size dependant cloning bias because Hind III cuts the mtDNA molecules into four fragments : HA (8.2 kb.), HB (5 kb., containing the A+T rich region), HC (4.7 kb.) and HD (0.43 kb.) (39). About 10^3 recombinant phages were screened using pDmtHC and pDmtHD as probes (see Mat. & Met.). We found that 286 clones are homologous to the HC fragment while only 69 clones are homologous to the HD fragment. As there is a priori no reason to suppose that the factor responsible for the difference lies in the mtDNA fragments themselves, this observation suggests that in this experiment, fragments approaching the size of the histone DNA repeats were cloned more efficiently than were fragments ten times smaller.

Taken together, these results strongly suggest that most of the ccc DNAs homologous to the histone multigenic family are shorter than the two main types of genomic repeats.

The large ccc DNA clones correspond to whole genomic repeats.

Two λDmcHH clones harboring a 5 kb. insert and two others carrying a 4.8 kb. insert were analysed by digesting the DNA with Hind III, Ava I and Eco RI, simultaneously, and hybridizing the resulting fragments with pDmHH8. The results (not shown) indicate that the two 5 kb. inserts and one of the 4.8 kb. inserts generate the same restriction fragments as those expected for a 5 kb. or a 4.8 kb. histone genomic repeat, respectively (see map of Fig. 1). The other 4.8 kb. insert does not give the expected pattern of a 4.8 kb.

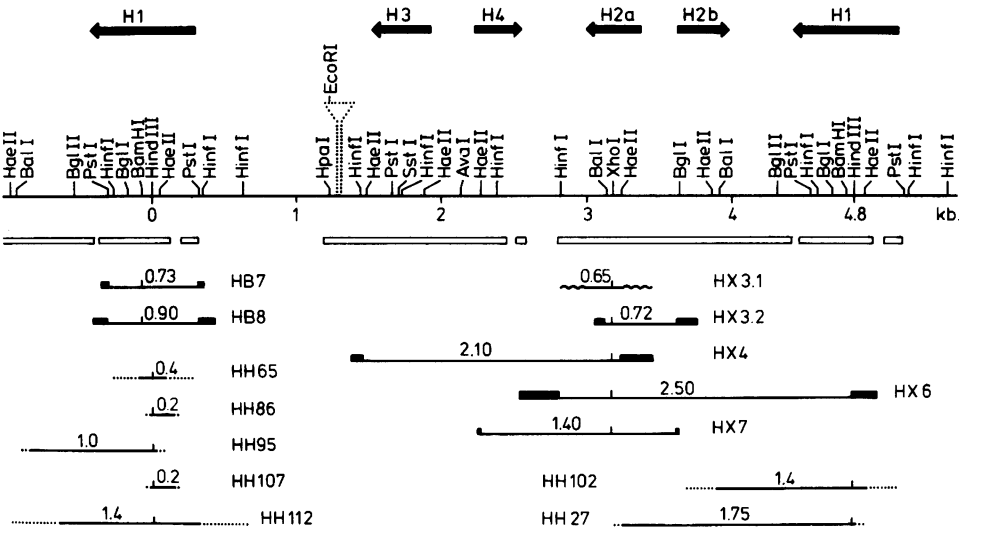


Fig. 1. Restriction maps of a 4.8 kb. histone genomic unit and of the ccc DNA clones shorter than a complete unit. A detailed restriction map of a 4.8 kb. unit is shown in the upper part of the Figure. The 200 bp. insertion found in the 5 kb. units is indicated by dotted lines. Heavy arrows represent the coding regions of each of the five histone genes. Open bars indicate parts of the 4.8 kb. unit for which the nucleotide sequence is known. All these data are from Lifton *et al.* (33), Goldberg (34) and Lifton (35).

In the lower part of the Figure the maps of the ccc DNA clones are given. Maps of the λcHH clones were established from the results shown on Fig. 2 and from Pst I and Bgl II digests for HH27, 102 and 112. The dotted lines indicate that the limits of the clones have not been accurately determined. Data concerning the pDmcHB and pDmcHX clones were obtained using the restriction enzymes shown on the map. New junction fragments are shown by bars. The name of each clone and its size in kb. is indicated. Non-histone DNA of pDmcHX3.1 is represented by wavy lines.

repeat : the 2.12 kb. Hind III-Ava I fragment is missing. However, two new fragments of approximately 0.9 and 1.2 kb. were detected suggesting that the clone carries a 4.8 kb. repeat that contain either an Eco RI site or an additional Ava I site. No more experiments were done with the 10 clones containing either a 4.8 or a 5 kb insert. Nevertheless, the available data lead to the conclusion that most if not all of these inserts correspond to the circular forms of the two main types of genomic histone repeats described in our previous report (7).

The short ccc DNA clones correspond to deleted genomic repeats.

Restriction analysis of the pDmcHX and pDmcHB ccc DNA clones indicates that all but one of them contain a single Xho I or Bam HI fragment. pDmcHX1 contains three Xho I fragments, only one (1.25 Kb.) of which shares homology with the histone genomic repeat. This clone was not investigated further.

Detailed restriction maps of the seven remaining clones are presented on Figure 1. Six of these clones display a map similar to that of a corresponding segment of the genomic repeat. The presence in every pDmcHX clone of the 70 bp. Bal I-Xho I and 49 bp. Xho I-Hae II fragments was confirmed by hybridization of a digest of the clones with a pDmHH8 probe. In the same way, the presence in the pDmcHB clones of the 250 bp. Pst I-Bam HI and 70 bp. Bam HI-Hind III fragments was verified. In each of the six clones, a new junction fragment was identified. The other clone, pDmcHX3.1 carries a single 650 bp. Xho I fragment which contains about 400 bp. of non histone encoding DNA in addition to a histone DNA segment that includes the Xho I site.

No detailed restriction map was made for the λDmcHH clones. However, we investigated whether the regions normally surrounding the Hind III cloning site are present in these ccc DNA clones. DNA of pDmHH8 was digested with various combinations of restriction enzymes, the fragments separated by electrophoresis and blotted onto nitrocellulose. Six identical filters were made and each was hybridized with the labeled DNA of one of the pDmcHX or λDmcHH clones. After removing the probes, the filters were used again for hybridizing the remaining clones.

The results obtained with the pDmcHX probes (not shown) are in good agreement with their restriction maps. Those obtained using the λDmcHH probes are shown in Figure 2. Every clone hybridized to the fragments immediately adjacent to the Hind III cloning site, i.e. the 1.23 kb. Hind III-Hpa I (lane c) and the 0.54 kb. Hind III-Bgl II (lane d) fragments. The intensity of the hybridization signals is highly variable and with the λDmcHH86 and 107 probes, the 0.54 kb. band is visible only on the original film. Moreover, there is a colinear relation between the DNA sequences of the clones and that of the genomic repeat. In cases where a fragment gave a relatively weak

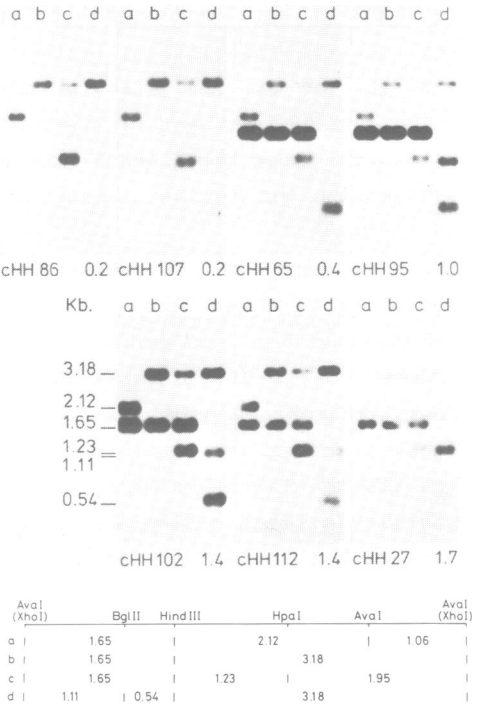


Fig. 2. Hybridization of λDmcHH probes with restriction digests of pDmHH8. a: Ava I + Hind III; b: Xho I + Hind III; c: Xho I + Hind III + Hpa I; d: Xho I + Bgl II + Hind III. The map indicates the sizes of the expected fragments generated by a 4.8 kb. unit digested by each set of restriction enzymes. Note that Ava I cuts at the Xho I site. The weak 3.18 kb band which sometimes appears in lane c results from a partial digestion

signal, suggesting that only a small region of homology is present in the probe (for example the 2.12 kb. Hind III-Ava I fragment with the cHH95 probe, lane a), a subsequent cut in the fragment (lane c) always indicated that the region of homology is located in the more proximal part relative to the cloning site. In addition, the more distal fragments are either never detected (1.06 kb. Ava I-Xho I, lane a) or are only detected (1.11 kb. Xho I-Bgl II, lane d) by the largest probes. Thus, it seems likely that as is the case with most of the pDmcHX and cHB clones, the λDmcHH clones include a part of the histone repeat that contains the Hind III site. Taking into account the relative variations in intensity of the hybridization signals, and

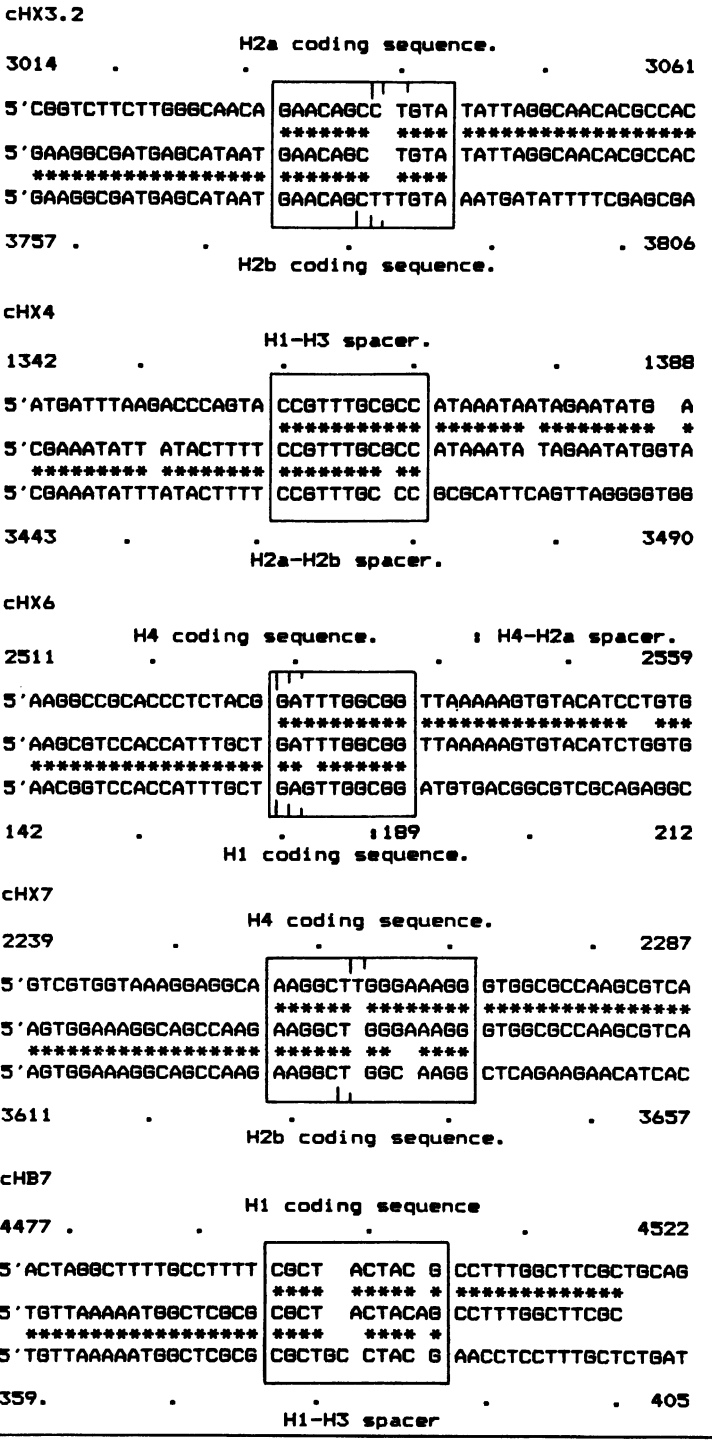
the Pst I and Bgl II restriction patterns of the largest inserts, we have estimated the position on the map of the ccc DNA molecules corresponding to the λDmcHH clones (Figure 1).

Therefore, these data indicate that clones shorter than a whole repeated genomic histone unit do not originate from whole circular repeats that have a restriction polymorphism or that carry a non histone insertion. Rather, they effectively represent circular deleted repeats. The maps given in Figure 1 show that most of the deletion breakpoints occur at different locations.

Nucleotide sequences of the junctions.

In order to obtain informations concerning the mechanisms by which circular deleted histone repeats are generated, we determined the nucleotide sequences of the junction regions of pDmcHX3.2, 4, 6, 7 and pDmcHB7. These sequences were compared to the nucleotide sequence established by Goldberg (34) and Lifton (35) which covers a large part of a 4.8 kb. genomic unit. Because the breakpoints of cHX6 and cHB7 lie outside, or very close to the limits of the sequenced regions (see Figure 1) we also determined the nucleotide sequence from positions 98 to 220, 324 to 424 and 2427 to 2579 of the genomic repeat. Comparisons of sequences of the junction regions of ccc DNAs with the corresponding genomic repeat sequences is shown on Figure 3.

The data clearly show that the breakpoints of each circular molecule correspond to short direct repeats in the genomic histone unit. These repeats, which are not related between the different clones, range in size from 10 to 15 bp. and always display 1 to 3 mismatches. In each clone only a single copy remains at the junction indicating that the ccc DNA molecule results from a recombinational event involving the short direct repeats. The position of the mismatches indicates the few possible alternative locations at which this event occurred, if it is assumed that the sequences of the direct repeats are strictly the same in the histone units which effectively gave rise to the cloned circular DNAs and in the histone unit used to establish the available genomic sequence. Indeed, the data obtained with pDmcHB7 indicate that discrepancies can exist. The copy of the repeat remaining in the circular molecule, carries an extra nucleotide pair (A) which is not found at this position in any of the corresponding repeats of the standard genomic sequence.


cHX3.2
H2a coding sequence.
3014
3061
5'CGGTCTTCTTGGGCAACA GAACAGCC TGTA TATTAGGCAACACGCCAC
5'GAAGGCGATGAGCATAAT GAACAGC TGTA TATTAGGCAACACGCCAC
5'GAAGGCGATGAGCATAAT GAACAGCTTTGTA AATGATATTTTCGAGCGA
3757
3806
H2b coding sequence.
cHX4
H1-H3 spacer.
1342
1388
5'ATGATTTAAGACCCAGTA CCGTTTGCGCC ATAAATAATAGAATATG A
5'CGAAATATT ATACTTTT CCGTTTGCGCC ATAAATA TAGAATATGGTA
5'CGAAATATTTATACTTTT CCGTTTGC CC GCGCATTCAGTTAGGGGTGG
3443
3490
H2a-H2b spacer.
cHX6
H4 coding sequence.
H4-H2a spacer.
2511
2559
5'AAGGCCGCACCCTCTACG GATTTGGCGG TTAAAAAGTGTACATCCTGTG
5'AAGCGTCCACCATTTGCT GATTTGGCGG TTAAAAAGTGTACATCTGGTG
5'AACGGTCCACCATTTGCT GAGTTGGCGG ATGTGACGGCGTCGCAGAGGC
142
189
212
H1 coding sequence.
cHX7
H4 coding sequence.
2239
2287
5'GTCGTGGTAAAGGAGGCA AAGGCTTGGGAAAGG GTGGCGCCAAGCGTCA
5'AGTGGAAAGGCAGCCAAG AAGGCT GGGAAAGG GTGGCGCCAAGCGTCA
5'AGTGGAAAGGCAGCCAAG AAGGCT GGC AAGG CTCAGAAGAACATCAC
3611
3657
H2b coding sequence.
cHB7
H1 coding sequence
4477
4522
5'ACTAGGCTTTTGCCTTTT CGCT ACTAC G CCTTTGGCTTCGCTGCAG
5'TGTTAAAAATGGCTCGCG CGCT ACTACAG CCTTTGGCTTCGC
5'TGTTAAAAATGGCTCGCG CGCTGC CTAC G AACCTCCTTTGCTCTGAT
359.
405
H1-H3 spacer

We also determined the nucleotide sequences of the junctions between histone and non-histone DNA of the clone pDmcHX3.1 (Figure 4). As previously, we found that in the genomic unit, the two breakpoints correspond to imperfect direct repeats of 7 and 8 bp. One complete copy of this repeat is present at one histone-non histone DNA junction of the circular molecule while only 5 nucleotide pairs of the repeat are found at the other junction. The nature of the non-histone DNA sequence was not investigated further. It seems possible that the imperfect 5 bp. direct repeats flanking the non-histone DNA sequence in the circular molecule could be the remnant of a target site duplication resulting from the insertion of a transposable element. However, homology was not detected between 100 bp. of the non-histone sequence of pDmcHX3.1 and transposable elements in the EMBL bank.

DISCUSSION

Twenty four clones homologous to a 4.8 kb. histone genomic unit probe were isolated from several libraries of ccc DNAs purified from embryos laid by SF dysgenic females. These clones were found to harbor complete 4.8 or 5.0 kb. repeated histone genes containing units, deleted units of various sizes or, in one case, a deleted unit carrying a non-histone DNA insertion.

In the Hind III cloning experiment, we obtained 10 clones that appear to contain a complete repeat and 7 clones that contain a deleted repeat. However, it seems unlikely that these figures reflect the relative proportions of the two classes of molecules in the ccc DNAs. The results of cloning mtDNA fragments of different sizes

Fig.3. Nucleotide sequence of the junction region of five ccc DNA clones. The upper and lower lines show the sequence of the genomic histone unit around the two breakpoints of each clone. The middle line shows the nucleotide sequence of the ccc DNA in the junction region. Homologies between corresponding sequences are indicated by asterisks. Boxes show the limits of the short directed repeats found at each junction while vertical bars give the different possible locations of the cuts in the genomic repeat which can lead to the observed recombination. Nothing was indicated for cHX4 and cHB7 because there are too many possibilities.

We established the sequence of pDmHH8 from positions 98 to 220 to fill a gap in the available sequence from positions 121 to 189 in which lies one of the cHX6 breakpoint. We found the gap smaller than expected, which leads to a discrepancy in the coordinates in the lower line of cHX6.

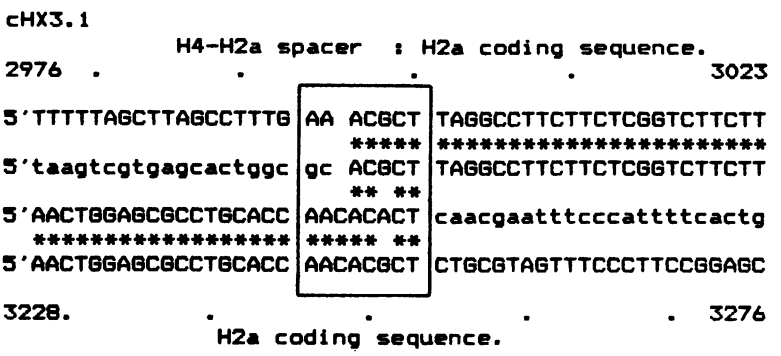


Fig. 4. Nucleotide sequence of the junction regions of cHX3.1. Same legend as Fig.3, except that the two middle lines give the nucleotide sequence of the ccc DNA around the junctions between histone and non-histone DNA. The sequence of the non-histone DNA is in small letters.

indicated that larger fragments (4.7 kb.) are cloned about four times more efficiently than smaller fragments (0.43 kb.). Also, the experimental design is expected to permit cloning of all the circular molecules corresponding to a whole histone repeat, but only of a fraction of those corresponding to a deleted repeat since such molecules have to include the Hind III restriction site. These observations strongly suggest that most of the ccc DNA molecules homologous to the histone DNA probe are deleted histone repeats.

It seems likely that the circular whole repeated units are generated by homologous recombination between two adjacent genomic units. Three of the studied λcHH clones do not show any major rearrangement and the fourth probably only differs from a standard 4.8 kb. unit by a restriction site polymorphism. Moreover, in our previous report (7), we showed that the histone homologous ccc DNAs detected by Southern blotting have a restriction pattern identical to that of the 4.8 and the 5,0 kb. unit. The relative proportions of the two types of complete units were roughly similar in the ccc DNAs and in the genome where the 5.0 kb. unit is known to occur in a higher frequency (35). The production of ccc DNAs from the Sau 3A family of human alphoid-like tandemly repeated sequences also involved homologous recombination even in case where the homology between the interacting repeats is less than 80% (48,49).

A different mechanism appears to be responsible for the generation of the 14 isolated ccc DNA molecules that each consist of

a deleted histone unit. The nucleotide sequences of the junction region of six of these units, including pDm cHX3.1 which carries a non-histone DNA insertion, show that in every case the deletions breakpoints are located within short imperfect directly repeated sequences. No nucleotide sequencing was done for the 8 remaining clones. However, we made a computer analysis to search for short direct repeats which could account for the formation of the ccc DNAs cloned in λDmcHH65, 86 and 107. The genomic sequence was screened from coordinates 4200 to 300. Ten pairs of 9 to 15 bp. perfect or imperfect direct repeats were detected which could give rise to putative circular DNAs including the Hind III site and ranging in size from 215 to 725 bp. Two of these pairs of direct repeats could account for λDmcHH65 and three others for λDmcHH86 or 107. Therefore, it seems likely that most of the circular molecules carrying a deleted histone unit result from a recombination between short directly repeated sequences. These sequences appear to be randomly scattered throughout the unit and as they are found in both coding sequences and spacer regions, their location must be independant of the structural organization of the histone genes.

However, it is not possible to determine whether those ccc DNAs which are generated by some kind of illegitimate recombination, result directly from genomic excisions or from complete circular units that were excised from the chromosomes by homologous recombination. In the first case, the production of ccc DNAs would generate deletions in the genomic histone cluster while in the second case it would only affect the number of repeated histone units. Two ccc DNAs produced by recombination between imperfect or perfect 9 bp. direct repeats have been cloned previously. One of them which consists of part of an unidentified repeated element, was isolated from hamster CHO cultured cells (50). The other, which contains part of a L1 transposable element and a fragment of the adjacent single copy genomic sequence, was isolated from human HeLa cells (51). In this latter case, the ccc DNA is not part of a tandemly repeated element and its direct excision from the chromosome seems very likely. In only a few cases have circular DNAs been described in relation to genomic deletions produced by illegitimate recombination. In the cyanobacterium Anabaena, the intervening DNA which separates the nif-K and nif-D genes was recovered as a circular molecule after the genes were joined by

recombination between 11 bp. direct repeats (52). Also, excision products resulting from the joining of V-J and V-D-J regions of the α and β chains genes of T cell receptor were found among the ccc DNAs of mouse thymocytes (25,26).

In both procaryotes and eucaryotes, various genetic rearrangements resulting from recombination between short direct repeats have been reported. In E.Coli, Whoriskey et al. (53) described tandem duplications of DNA fragments which are amplified 50 to 100 fold during selected growth. In most cases, the endpoints of the duplicated fragments are located within short (less than 15 bp.) imperfect directly repeated sequences. A single copy of the repeat was found at the junction of the tandemly amplified fragments. Genomic deletions involving recombination between short repeated sequences also have been reported in E.Coli (54 and references therehein, 55) and in mammalian cells (56-58). These observations suggest that the mechanisms which generate such genetic rearrangements are related to those which give rise to ccc DNAs. One of the proposed models to explain this kind of illegitimate recombination involves slipped strand mispairing (54). However, this model does not seem to apply to the recombination which gives rise to the histone ccc DNAs. This is mainly because it predicts the excision of a single stranded DNA fragment which would have to be converted into a double stranded covalently closed circular molecule. Moreover, in the present case, no palindromic sequences long enough to help form the necessary mispaired structure were detected close to the breakpoints. Therefore, it seems to us more likely that recombination is mediated by a specific enzyme(s) which can recognize DNA structures or prefered sequences, as does topoisomerase I (59). However, no consensus sequence was found at the junction of the ccc DNAs.

It appears from our data that at least in embryos laid by SF females, the production of ccc DNAs results more frequently from illegitimate recombination than from homologous recombination. Since SF females allow the transposition of the I element with a high efficiency (60) it will be interesting to establish whether the higher frequency of illegitimate recombination is common to all D. melanogaster strains or in some way related to the I-R hybrid dysgenesis. In any case, the production of circular DNA molecules,

either by homologous or illegitimate recombination, or both, appears to be a powerful mutagenic force. If such an event took place in germ cells, then, it may be of evolutionary importance.

ACKNOWLEDGEMENTS

We thank Dr. J.C.Bregliano for helpful discussions throughout this work, Dr. D.R. Wolstenholme for critical reading of the manuscript, Drs. D.S. Hogness, M.L. Goldberg and R.P. Lifton for providing nucleotide sequences of the histone repeated unit, Dr. M. Renaud for plasmids pDmtHC and HD and Dr. C. Vaury for help in sequencing. This work was supported by grants from CNRS (UA 360), Université Clermont-FdII, ARC, INSERM (CRE 831020) and Fondation Médicale. G. Pont was supported by an ARC award.

+Present address: Department of Biology, University of Utah, Salt-Lake City, UT 84112, USA

*To whom correspondence should be addressed

REFERENCES

1. Smith, C.A. & Vinograd, J. (1972). J. Mol. Biol. 69, 163-178
2. Stanfield, S.W. & Helinski, D.R. (1976). Cell 9, 333-345.
3. Stanfield, S.W. & Helinski, D.R. (1984). mol. Cell. Biol. 4, 1, 173-180.
4. Delap, R.J., Rush, M.G., Zouzias, D. & Khan, S. (1978). Plasmid 1, 508-521.
5. Tsuda, T., Yamagishi, H., Ohnishi, N., Yamada, Y., Izumi, H. & Mori, K.J. (1983). Plasmid 10, 235-241.
6. Yamagishi, H., Kunisada, T., Iwakura, Y., Nishimune, Y., Ogiso, Y., & Matsushiro, A. (1983). Develop. Growth and Differ. 25, 6, 563-569.
7. Pont, G., Degroote, F. & Picard, G. (1987). J. Mol. Biol. 195, 447-451.
8. Wong, F.Y. & Wildman, S.G. (1972). Biochim. Biophys. Acta 259, 5-12.
9. Kinoshita, Y., Ohnishi, N., Yamada, Y., Kunisada, T. & Yamagishi, H. (1985). Plant Cell Physiol. 26, 7, 1401-1409.
10. Delap, R.J. & Rush, M.G. (1978). Proc. Natl. Acad. Sci. USA. 75, 12, 5855-5859.
11. Stanfield, S.W. & Lengyel, J.A. (1979). Proc. Natl. Acad. Sci. USA. 76, 12, 6142-6146.
12. Bertelsen, A.H., Zafri-Humayun, M., Karfopoulos, S.G. & Rush, M.G.(1982). Biochemistry 21, 2076-2085.
13. Kunisada, T. & Yamagishi, H. (1984). Gene 31, 213-223.4.
14. Riabowol, K., Shmookler-Reis, R.J. & Goldstein, S. (1985). Nucl.Acids Res. 13, 15, 5563-5584.
15. Sunnerhagen, P., Sjorberg, R.M., Karlsson, A.L., Lundh, L. & Bjursell, G. (1986). Nucl. Acids Res. 14, 20, 7823-7838.
16. Rush, M.G. & Misra, R. (1985). Plasmid 14, 177-191.
17. Arkhipova, I.R., Gorelova, T.V., Ilyin, Y.V. & Schuppe, N.G. (1984). Nucl. Acids Res. 12, 19, 7533-7548.

18. Ilyin, Y.V., Schuppe, N.G., Lyubomirskaya, N.V., Gorelova, T.V. & Arkhipova, I.R. (1984). Nucl. Acids Res. 12, 19, 7517-7531.
19. Sundaresan, V., & Freeling, M. (1987). Proc. Natl. Acad. Sci. USA. 84, 4924-4928.
20. Misra, R., Shih, A., Rush, M., Wong, E. & Schmid, C.W. (1987). J. Mol. Biol. 196, 233-243.
21. Jones, R.S. & Potter, S.S. (1985). Nucl. Acids Res. 13, 3, 1027-1042.
22. Hollis, M. & Hindley, J. (1986). Gene 46, 153-160.
23. Kiyama, R., Matsui, H. & Oishi, M. (1986). Proc. Natl. Acad. Sci. USA. 83, 4665-4669.
24. Kiyama, R., Matsui, H., Okumura, K. & Oishi, M. (1987). J. Mol. Biol. 193, 591-597.
25. Fujimoto, S. & Yamagishi, H. (1987). Nature 327, 242-243.
26. Okazaki,K., Davis,D.D. & Sakano,H. (1987). Cell 49, 477-485.
27. Buongiorno-Nardelli, M., Amaldi, F. & Lava-Sanchez, P.A. (1976). Exp. Cell Res. 98, 95-103.
28. Bird, A.P. (1978). Cold Spring Harbor Symp. Quant.Biol. 42, 1179-1183.
29. Graziani, F., Caizzi, R. & Gargano, S. (1977). J. Mol. Biol. 112, 49-63.
30. Chernyshev, A.I. (1982) Mol. Gen. Genet. 185, 176-180.
31. Beverley, S.M., Coderre, J.A., Santi, D.V. & Schimke, R.T. (1984). Cell 38, 431-439.
32. Kraszewska, E.K., Bjerknes, C.A., Lamm, S.S. & Van't Hof, J. (1985). Plant Mol. Biol. 5, 353-361.
33. Lifton, R.P., Goldberg, M.L., Karp, R.W. & Hogness, D.S. (1977). Cold Spring Harbor Symp. Quant. Biol. 42, 1047-1051.
34. Goldberg, M.L. (1979). PhD. Thesis, Stanford University, Palo-Alto, CA.
35. Lifton, R.P. (1985). PhD. Thesis, Stanford University, Palo-Alto, CA.
36. Bregliano, J.C. & Kidwell, M.G. (1983). In Mobile genetic elements, (Shapiro, J.A. ed.) pp. 363-410, Acad.Press, NY.
37. Viera J. & Messing, J. (1982). Gene 19, 259-268.
38. Berthier, F., Renaud, M., Alziari, S. & Durand, R. (1986). Nucl.Acids Res. 14, 11, 4519-4533.
39. Merten, S.H. & Pardue, M.L. (1981). J. Mol. Biol. 153, 1-21.
40. Micard, D., Sobrier, M.L., Couderc, J.L., & Dastugue, B. (1985). Anal.Bio-chem. 148, 121-126.
41. Rigby, P.W.J., Dieckmann, M., Rhodes, C. & Berg, P. (1977). J. Mol. Biol. 113, 237-251.
42. Norrander, J., Kempe, T. & Messing, J. (1983). Gene 26, 101-106.
43. Gough, J.A. & Murray, N.E. (1983). J. Mol. Biol. 166, 1-19.
44. Hanahan, D. (1983). J. Mol. Biol. 166, 557-580.
45. Murray, N.E. (1983). In Lambda II, (Hendrix, R.W., Roberts, J.W. Stahl, F.W.& Weisberg, R.A. eds.). pp 395-432, Cold Spring Harbor Laboratory Press, New York.
46. Sanger, F., Coulson, A.R., Smith, A.J.M. & Roc, B.A. (1980). J. Mol. Biol. 143, 161-178.
47. Guo, L.H., Yang, R.C.A. & Wu R. (1983). Nucl.Acids Res. 11, 5521-5539.
48. Okumura, K.,Kiyama, R. & Oishi,M. (1987). Nucl.Acids Res. 15, 18, 7477-7489.

49. Kiyama, R., Okumura, K., Matsui, H., Bruns, G.A.P., Kanda, N. & Oishi, M. (1987). J. Mol. Biol. 198, 589-598.
50. Stanfield, S.W. & Helinski, D.R. (1986). Nucl.Acids Res. 14, 8, 3527-3538.
51. Jones, R.S. & Potter, S.S. (1985). Proc.Nat.Acad.Sci.USA. 82, 1989-1993.
52. Golden, J.W., Robinson, S.J. & Haselkorn,R. (1985). Nature 314, 419-423.
53. Whoriskey, S.K., Nghiem, V.H., Leong, P.M., Masson, J.M. & Miller, J.H. (1987). Genes & Development 1, 227-237.
54. Albertini,A.M., Hofer,M., Calos,M.P. & Miller,J.H. (1982). Cell 29, 319-328.
55. Glickman, B.W. & Ripley,L.S. (1984). Proc.Natl.Acad.Sci.USA 81, 512-516.
56. Bullock, P., Champoux, J.J. & Botchan, M. (1985). Science 230, 954-957.
57. Henthorn, P.S.,Mager, D.L.,Huisman, T.H.J. & Smithies, O. (1986). Proc.Natl. Acad.Sci.USA 83, 5194-5198.
58. Nalbantoglu, J., Hartley, D., Phear, G. Tear, G. & Meuth, M. (1986). Embo J. 5, 6, 1199-1204.
59. Been, M.D., Burgess, R.R. & Champoux, J.J. (1984). Nucl. Acids.Res.., 12, 7, 3097-3113.
60. Picard, G. (1976). Genetics 83, 107-123.