J Mol Evol (1987) 26:311–319

Journal of
Molecular Evolution

Cotransposition of a Highly Repetitive DNA Element with Flanking Sequences in the Genome of the Midge *Chironomus thummi*

Th. Hankeln and E.R. Schmidt

Institute of Genetics, Ruhr-Universität, Postfach 10 21 48, D-4630 Bochum 1, West Germany

Summary. A family of highly repetitive DNA elements, the Cla-elements, is present in the genomes of the two sibling species *Chironomus th. thummi* and *Ch. th. piger*. These Cla-elements are organized in large tandem repetitive clusters as well as occurring as interspersed monomeric elements, in both subspecies. The analysis of a monomeric Cla-element and several kilobases of its flanking sequences from *Ch. th. piger* revealed that the short Cla-elements are cotransposed together with adjacent DNA. We found the same association of Cla-elements with specific flanking DNA in clones obtained from the rDNA of *Ch. th. thummi* and from nonribosomal Cla-DNA of *Ch. th. piger*. The Cla-element-flanking DNA is clearly also repetitive, but mainly of interspersed organization.

Key words: *Chironomus thummi* – Repetitive DNA – Cotransposition – NTS of rDNA – Cla-elements – Satellite DNA – DNA amplification

Introduction

In the genome of many *Chironomus* species a family of repetitive DNA elements is present, which has been termed "Cla-elements" (Schmidt 1981, 1984) because most of these elements are characterized by a recognition site for the restriction endonuclease ClaI. The most intriguing feature of the Cla-elements is the fact that they have probably undergone an evolutionary amplification process during the divergent evolution of two subspecies of *Chironomus thummi* (Schaefer and Schmidt 1981).

Offprint requests to: Th. Hankeln

Chironomus thummi thummi and *Ch. th. piger* can be crossed to produce fully fertile offspring (Keyl and Strenzke 1956). In spite of the close relatedness between the two species, there is a significant difference in genome size: *Ch. th. thummi* contains about 30% more DNA per genome than *Ch. th. piger* (Keyl 1965). About one-fifth of this additional DNA in *Ch. th. thummi* is represented by the higher copy number of Cla-elements (Schaefer and Schmidt 1981; Schmidt 1984). From a cytological point of view, the larger genome of *Ch. th. thummi* represents the phylogenetically younger one (Keyl 1965) and thus, one can conclude that the higher copy number of Cla-elements is the result of an amplification during the evolution of the *thummi* genotype rather than a loss of Cla-elements during the evolution of the *piger* genotype.

In addition to the gross difference in the copy number of Cla-elements in the two genomes there is a remarkable difference in the distribution of Cla-elements in the chromosomes of *Ch. th. thummi* and *Ch. th. piger* (Schmidt 1981). In situ hybridization suggests that the Cla-elements are confined to the centromeric heterochromatin in the chromosomes of *Ch. th. piger* but, in addition to the centromeres, are distributed over more than 200 different sites in the chromosomes of *Ch. th. thummi* (Schmidt 1984). Probably all these sites contain clusters of Cla-elements, because single Cla-element copies are presumably too short to provide sufficient homology in situ. One of these 200 Cla-element sites is the nontranscribed spacer (NTS) of the rDNA of *Ch. th. thummi*, which contains a large Cla-element cluster of varying length, generating extreme length heterogeneity of the individual *thummi*-rDNA repeating units (Israelewski and Schmidt 1982; Schmidt et al. 1982). The NTS of *Ch. th. piger* does

not contain any Cla-elements (Schmidt and Godwin 1983). The comparison of the NTS of the two subspecies has led to the hypothesis that Cla-elements might have been transposable, either as isolated repetitive elements or in association with adjacent DNA sequences found flanking the Cla-element cluster in the NTS of *thummi*.

Although cloned Cla-DNA from *Ch. th. thummi* has already been analyzed (Schmidt 1984), up to now there was little information on the organization of Cla-elements in the phylogenetically older genome of *Ch. th. piger*. Genomic Southern analysis indicated that there were clustered as well as interspersed copies of Cla-elements throughout the DNA of *Ch. th. piger*. However, localization of Cla-elements by in situ hybridization yielded only a significant signal over the centromeric region, showing that this method is not sensitive enough to detect single copies of monomeric Cla-elements. It was, therefore, necessary to analyze Cla-elements and their flanking DNA cloned from the genome of *Ch. th. piger*. This would enable us to compare the organization of Cla-elements of the phylogenetically younger with the phylogenetically older genome.

In the present study we were able to isolate for the first time a monomeric Cla-element, which is surrounded by repetitive DNA sequences that possibly originate from a noncentromeric chromosomal region. The detailed molecular analysis in this study supports the idea that the Cla-elements undergo transposition together with adjacent DNA sequences and that the transpositional events have also taken place in the genome of *Ch. th. piger*.

Materials and Methods

Cloning of Cla-elements from Ch. th. piger DNA. Genomic DNA of *Ch. th. piger* was prepared from larvae raised in laboratory culture as described previously (Schmidt et al. 1980). A library of *Ch. th. piger* DNA was constructed using the lambda replacement vector EMBL 3 (Frischauf et al. 1983). DNA from larvae was cut partially with Sau3A and fractionated by agarose gel electrophoresis. Fragments of 12–23 kilobases (kb) were isolated, ligated to the BamHI ends of the EMBL 3 vector arms, and packaged in vitro. The selection of recombinant phages was carried out biochemically as described by Frischauf et al. (1983).

The library was screened for Cla-element-containing clones by the Benton and Davis (1977) method. Cla-elements, isolated from genomic DNA of *Ch. th. thummi* and labeled radioactively by nick-translation (Rigby et al. 1977), served as hybridization probes. Restriction fragments of the EMBL 3 clones were subcloned into the plasmid vector pBR322. After transformation of *Escherichia coli* strain SF8 by the standard calcium procedure, recombinant bacterial clones were enriched by cycloserine selection (Bolivar and Backman 1979).

Analysis of Cla-Element-Containing Clones. Preparation of plasmid DNA and agarose gel electrophoresis were performed as described previously (Schmidt 1981; Schmidt and Godwin, 1983). Blot-filter hybridizations were carried out according to Southern (1975) and DNA sequencing was done according to Maxam and Gilbert (1980). Restriction mapping was performed according to the method of Boseley et al. (1980).

In Situ Hybridization to Polytene Chromosomes. Polytene chromosome preparations of *Ch. th. thummi* and *Ch. th. piger*, from salivary glands of fourth-instar larvae, were heat-treated at 70°C for 30 min in 2× standard saline citrate (0.15 M NaCl, 0.015 M sodium citrate; SSC) (Bonner and Pardue 1976). The spreads were digested with RNase and the DNA in the chromosomes was then denatured in 0.07 N NaOH for 1 min. Hybridization was carried out at 55°C in 5× SSC for 4–6 h. The DNA probe was labeled with biotin–dCTP (ENZO, Biochem. Inc.) by "oligo-labeling" (Feinberg and Volgelstein 1983). Sites of hybridization were detected by a two-step antibody reaction: The slides were washed in phosphate-buffered saline (PBS) for 5 min. Subsequently, 10 µl of rabbit-anti-biotin-IgG (ENZO, diluted 1:50 in PBS) were added per slide. Slides were incubated at 37°C for 3 h in a moist chamber, then washed in PBS for a total of 5 min. Goat-anti-rabbit-IgG conjugated to FITC (10 µl, diluted 1:100 in PBS) was added and the slides incubated for 1 h at 37°C. After a washing step of 10 min in PBS, chromosome preparations were embedded in buffered *p*-phenylenediamine/glycerol-antifading mixture.

Preparations were inspected using a Zeiss photomicroscope equipped with a fluorescent device (excitation 455 nm).

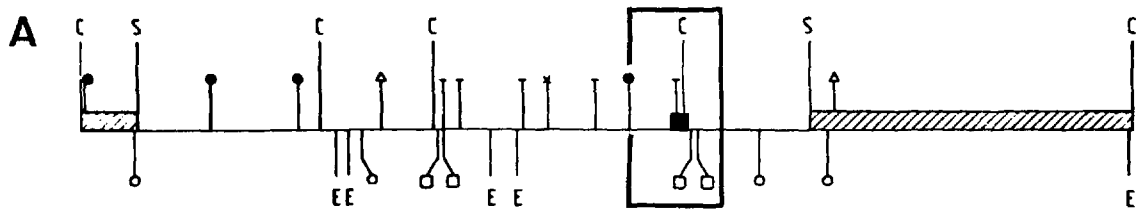
Results

Isolation and Characterization of Cla-Element-Containing Clones from Ch. th. piger

A genomic library of *Ch. th. piger* containing 1.5×10^4 recombinant EMBL 3 clones was screened with radioactively labeled Cla-elements as hybridization probes. Approximately 1% of the screened λ clones showed hybridization with Cla-elements. With an average size of cloned fragments of 18 kb and the size of the *Ch. th. piger* genome (1.25×10^5 kb; Keyl, personal communication), this is equivalent to an approximate number of 70 Cla-DNA-containing clones per *Ch. th. piger* genome. In order to analyze a Cla-element-containing region in detail, we have chosen arbitrarily one clone (λ22A) out of 50 isolated clones. λ22A was characterized by restriction analysis and blot-filter hybridization (not shown). Restriction of the λ22A DNA by SalI generated a SalI-fragment of about 8 kb, which contained the Cla-element-homologous sequences. This fragment was subcloned into pBR322 and termed pPC 170. Figure 1A shows the restriction map of the Cla-containing subclone pPC 170. The Cla-element-homologous DNA was localized by blot-filter hybridization (not shown). About 900 bp, spanning the Cla-homologous region, were analyzed in detail by DNA sequencing (Fig. 1B).

DNA Sequences of the Ch. th. piger Cla-element and Its Flanking DNA

With the clone pPC 170, for the first time a monomeric member of the Cla-DNA family, surround-



B

1 Kb

```
5'
                    20                  40                  60                  80                 100
GAGGATGAAGATGATAAGCTGGACAACAATGGAAGCACCAATGAAATTGCCATAGTGGAGATAGAAGTGAGTGGAGGATTTTCATAATCCAAAATTAAAA

                   120                 140                 160                 180                 200
GGACTCAGCCAATATCGTGTGGTTTGTATGATTTACATTCAAGAAGCATCATGTCAAGTTGGTACTTCCAACTGGATTGAAACTGACAAAAATGTTTTTA

                   220                 240                 260                 280                 300
CAACATTTCTTGACTTTTTTAAAGATTTTTGAGAGGGGTCCAGATTTAAGGTAACTAAACACTTCTTTCATTATCCAATTGAAATTTTTCGGTTCTAAGT

                   320                 340                 360                 380                 400
AGTCATACGATGCGTACTGGCTGCATTGGAACCTACGGAACTGCTTAGTTACCTCAGAACTAGACCCATCTTTAAAATCGTTTGAAAAGTCAAGAAATTG

                   420                 440                 460                 480                 500
AAGTACAATTTTTTTGTCAGTTCCACCCAGTTGGAAGTGCCAACTAGACATGATGCTTCTTGCATGGAAATAATGCTAATCATGCAGGCAAGCAGCTGAT

                   520                 540                 560                 580                 600
ATTCATCATCCTTTTTGTTGGTGAATTTTTATTTTTAGCCTACTATGGTCGAAACCAAAAATGCTTTAAATTGCATTTTAAAACATATCTATGTATTATA
                          *****               *   *                       *
                    GAAAATAGCCTACTATGATCAAAACCAAAAATGCTTTAAATTGCATTTTAAAACATATCTAGGTATTATA

                   620                 640                 660                 680                 700
AATGATAAAAAATGCTAATATAGATATAAAATCGATATTTTTAGCGCATTTTTGAAACTTGTTTCGAAAAGGTACACTCATAGAAAAAAAGTAACAAAAT
                 *                   *****
AATGATAAAGAATGCTAATATAGATATAAAATCGATATTTTTTTTTT

                   720                 740                 760                 780                 800
GTTTGTAACTTGTATGGAAAATGTTAACAAACAATTCGTTAAAATAGTGTTTACGAACTTATTTGTTAAATTTAACCAAATTATTTGTAAGGTTGAATTG

                   820                 840                 860                 880                 900
TTAACTATAAAAATTGTTAATTTAACAAACTTTTTGTTAAATTTTTTTTTTACAAGAAAAAAAACAAATTTTCGGTTAATATTAACCAAGAAATTGTTAAT
                                                                                                      3'
```

Fig. 1. A Restriction map of pPC 170. The sequenced region, containing the Cla-element, is boxed. Restriction enzymes are as follows: C, ClaI; E, EcoRI; O, HaeIII; ●, HindIII; □, HpaI; S, SalI; T, TaqI; X, XhoI; △, XmaIII; ■, Cla-element; ▨, pBR322. **B** Nucleotide sequence of the pPC 170 Cla-element and its flanking DNA. The Cla-monomer (boxed) is compared with the Cla-consensus sequence (in italics). Differences are marked by dots. The ClaI site comprises nucleotides 631–636. A large imperfect palindrome in the 5'-Cla-flanking region is underlined (→). Hypothetical target site duplications are marked by small arrows (→).

ed by its flanking sequences, could be isolated and analyzed. All previously obtained Cla-elements from *Ch. th. thummi* and *Ch. th. piger* were part of tandem-repetitive clusters (Schmidt 1984).

The monomeric Cla-element of pPC 170 has a length of 107 bp and thus is 10 bp shorter than the consensus Cla-element sequence (Fig. 1B). The length difference is due to deletions within oligo-dT/dA regions at the 5'-end and the 3'-end of this Cla-element. Sequence comparison between the *piger*-Cla-element in pPC 170 and the Cla-consensus sequence (Schmidt 1984) reveals a divergence of 3.7%.

At both ends of the Cla-element two duplications of 6 bp can be seen (Fig. 1B). At the 5'-end of the Cla-element, these duplications are tandemly arranged and overlap by 1 bp with the Cla-element sequence itself. At the 3'-end the first 6-bp duplication is completely homologous to the Cla-element. Here, the second duplication is separated from the first by a few nucleotides. These duplications might be an indication for a transpositional event involving the monomeric Cla-element, which has occurred during evolution.

The Cla-element-flanking DNA from the *piger*-clone pPC 170 was compared by computer analysis with the flanking sequences of previously analyzed Cla-element clusters from *Ch. th. thummi* (Schmidt 1984). This analysis revealed that a 32-bp region, immediately flanking the 3'-end of the *piger*-Cla-element of pPC 170, is nearly identical to the se-

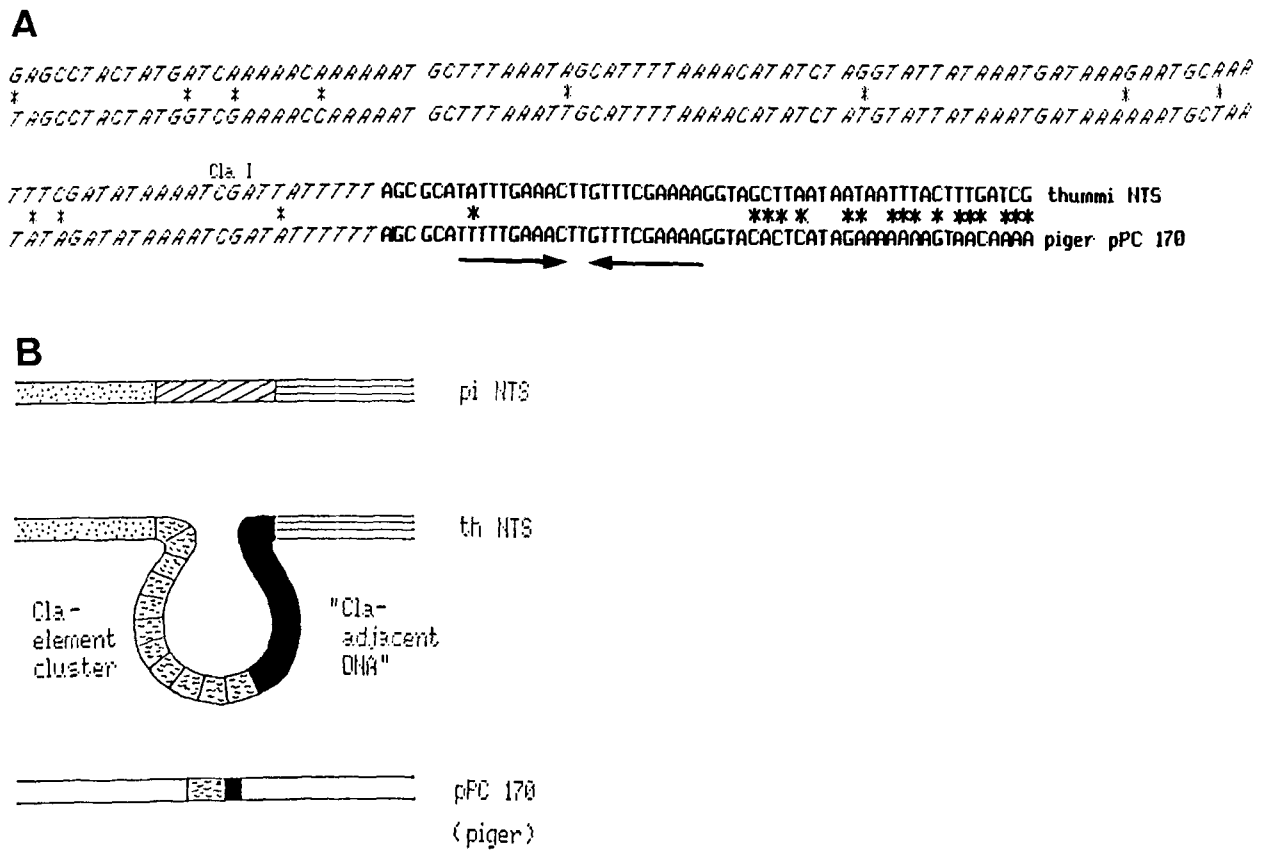


Fig. 2. **A** Sequence homology between the Cla-element/Cla-flanking DNA of the ribosomal NTS and of pPC 170 (Cla-element in italics, "Cla-adjacent DNA" in bold face). A short 3′-flanking imperfect palindromic sequence within the "Cla-adjacent DNA" is indicated by arrows. **B** Compiled schematic comparison between the ribosomal NTS regions of *Ch. th. thummi* and *Ch. th. piger* and the nonribosomal *piger*-clone pPC 170. Homologous sequences are hatched identically.

quence that was found flanking the 3′-end of the Cla-cluster present in the ribosomal NTS of *Ch. th. thummi* (Fig. 2A, B). Following Schmidt and Godwin (1983), this Cla-element-flanking sequence of the *thummi*-NTS is called "Cla-adjacent DNA."

The nearly perfect homology to the "Cla-adjacent DNA" of the *thummi*-NTS is confined to the 32 bp flanking the 3′-end of the pPC 170-Cla-element. Further downstream as well as in the 5′-flanking region of the Cla-monomer there is no homology detectable to any ribosomal sequences. These results indicate that at least the 32 bp of "Cla-adjacent DNA" present in pPC 170 and the NTS of *Ch. th. thummi* were linked to a Cla-element before the two subspecies separated (see Discussion).

The Cla-adjacent DNA and the sequences downstream thereof are of typical simple sequence organization. The nucleotides 650–900 (Fig. 1B) can be written as (imperfect) repeating units of the consensus sequence

$$\begin{matrix} GAAA \\ TTTT \end{matrix} TTGTT \begin{matrix} TG \\ AA \end{matrix}$$

(Fig. 3). Twenty-two such units could be identified. It is worth noting that the length of the repeats often coincides with the helical repeat of DNA. Furthermore, variation of the repeating units at the nucleotide level can be noticed: along with increasing distance to the first repeat, the poly-A sequence proximal to the central guanosine position is replaced by poly-T. Additionally, the cytosine and guanosine residues at position 4 distal to the central G only occur in the first few repeats and are then mainly replaced by adenosine.

The satellite-like sequence also seems to be present in the 5′-Cla-flanking sequence. However, only a few related repeating units can be detected here (Fig. 3), which are not tandemly arranged.

The main feature of the 5′-flanking region of the Cla-element is a relatively large, imperfect inverted repeat, which has a 167-bp stem and a loop of 45 bp (Fig. 1B). In the stem, 132 out of 167 nucleotides are paired. The large hairpin is flanked by a 7-bp direct duplication with one base of mismatch. The 3′-duplication is an integral part of the inverted repeat, whereas the 5′-duplicated sequence immediately flanks the stem (Fig. 1B).

```
3'                          661
12          G A A ACT T  G T T T C
                            672
10          G A A A A    GG T A   C
   ACTCATA                  691
13          G A A A A AAAG T A A C
                            701
10            A A A A T  G T T T G
                            712
11          T A A C T T  G T A T G
                            723
11          G A A A A T  G T T A A
                            732
10          G A A   T T  G T T T G   (rc)
                            738
 5                        G T T A A
                            749
11            A A T AGT  G T T T A
                            758
 8                T A A  G T T C G   (rc)
                            765
 9                T T T  G T T A A A
                            776
10                T T T GG T T A A A (rc)
                            787
10        T T   A T T T  G T   A A
                            800
13        T T G A A T T  G T T A A C
      GG                    816
12        A A A A A T T  G T T A A
     TAT                    826
10              G T T T  G T T A A A (rc)
                            836
11            T T T T T  G T T A A A
   TTTTTT                   852
12        T T T T C T T  G T   A A A (rc)
                            864
 8                T T T  G T T T T   (rc)
                            874
12              T T T TCGG T T A A T
                            885
13          T T T C T T GG T T A A T (rc)
                            895
 7                  T T  G T T A A

5'                          49
11          G A A A T T  G C C A T
                           400
11          G A A A T T  G A A G T

CON-      G A A A                T G
SENSUS    T T T T    T T  G T T    A A
```

Fig. 3. Possible "simple sequence" repeating units of the 3'- and 5'-Cla-flanking DNA. Some repeating units are present as reverse complements (rc). The lengths of the repeats are given at the left margin. The numbering of the sequences corresponds to Fig. 1B. A similar simple sequence organization has also been found in the Cla-element-flanking region of the *thummi*-Cla-element cluster i6 (Schmidt 1984; Rohwedder 1986).

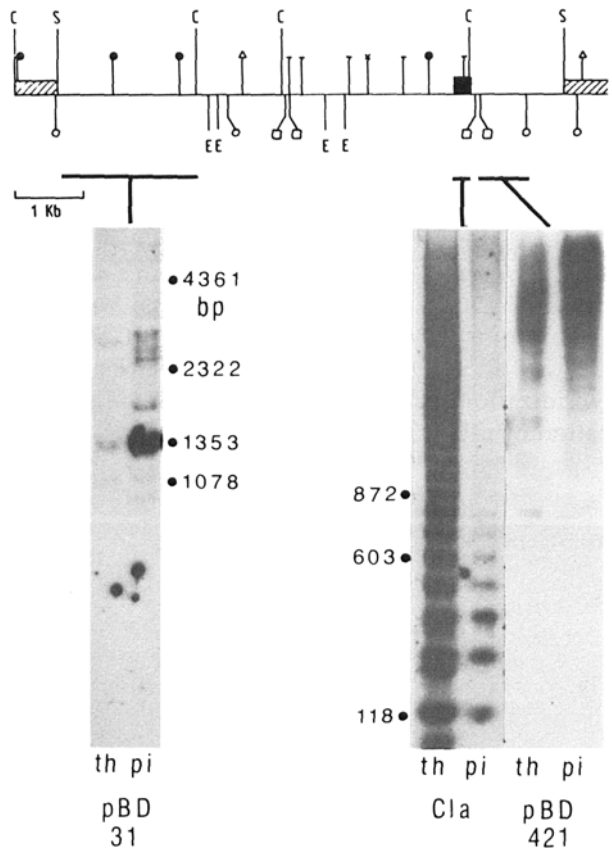


Fig. 4. Southern blot hybridization of Cla-elements and subclones pBD 421 and pBD 31 against ClaI-restricted total genomic DNA of *Ch. th. thummi* (th) and *Ch. th. piger* (pi) larvae. For the symbols used on the restriction map, see Fig. 1.

Genomic Organization of pPC 170

In order to investigate the genomic organization of the Cla-containing region several "subregions" of pPC 170 were analyzed by Southern blot hybridization against total genomic DNA from *thummi* and *piger* larvae (Fig. 4). The subclone pBD 31, which spans the "leftmost" part of pPC 170, hybridizes to only a few bands after ClaI restriction of *thummi*- and *piger*-DNA, with a weaker hybridization to the DNA of *Ch. th. thummi*, indicating a lower copy number in the *thummi* genome.

Subclone pBD 421, containing the Cla-adjacent DNA and its 3'-flanking region, however, produces a broad "smear" in genomic DNA of *Ch. th. thummi* and *Ch. th. piger*, typical for highly repetitive DNA. This region of pPC 170 spans the aforementioned "simple-sequence" DNA, which is found flanking the 3'-end of the Cla-element. The Cla-element produces the characteristic ladder of bands of a highly tandem-repetitive DNA with a significantly different intensity of hybridization of *thummi*- and *piger*-DNA. The sequences in the immediate 5'-Cla-flanking region show a complex hybridization pattern (not shown) and therefore contain dispersed, moderately repetitive DNA.

In Situ Localization

For in situ localization, the Cla-element-free subclone pBD 31 was used as a probe. pBD 31 represents the "left" end of pPC 170 and contains se-

quences that are of low copy number in the genome of *Ch. th. piger* and probably of single copy in the genome of *Ch. th. thummi.* The result of in situ hybridization is shown in Fig. 5. In the chromosomes of *Ch. th. thummi,* only a single band is labeled, corresponding to position C2g on the right arm of chromosome I (according to Hägele 1970). This hybridization is found in both *Ch. th. piger* and *Ch. th. thummi* at the same site. In the chromosomes of *Ch. th. piger* a few additional bands on chromosome I and the centromeres of chromosomes II and III are labeled (Fig. 5). These in situ hybridization results do not allow a definite localization of pPC 170 in *Ch. th. piger.*

Discussion

Cotransposition of Cla-Elements Together with Adjacent DNA

The most remarkable feature of pPC 170 is that the *piger*-Cla-element of this clone is flanked by the same 32-bp DNA sequence that has previously been found flanking the Cla-cluster of the ribosomal NTS of *Ch. th. thummi.* This is important, because the comparative analysis of the ribosomal nontranscribed spacers of *Ch. th. thummi* and *Ch. th. piger* has led to the idea that Cla-elements might have been transposed as a complex together with flanking DNA (Schmidt and Godwin 1983).

In the case of the NTS of *Ch. th. thummi,* 528 bp of so-called "Cla-adjacent DNA" are flanking the 3′-end of the Cla-element cluster. This DNA is clearly nonribosomal and has probably integrated into the NTS, eventually together with the Cla-elements (Schmidt and Godwin 1983).

Strong support for this hypothesis comes from the sequencing data of clone pPC 170. In pPC 170 we found a situation which is very similar to that in the NTS of *Ch. th. thummi.* The Cla-monomer is associated with a 32-bp flanking sequence, which is homologous to part of the "Cla-adjacent DNA" found in the *thummi*-NTS. Since pPC 170 has been cloned from *Ch. th. piger* and does not contain rDNA sequences, the "Cla-element/Cla-adjacent DNA" complex has now been found independently in both subspecies at different chromosomal loci. In fact, this means that Cla-elements and 32 bp of the "Cla-adjacent DNA" were associated already before the two subspecies separated and before these sequences jumped into the nontranscribed spacer of the *thummi*-rDNA.

However, the major part of the *thummi*-NTS "Cla-adjacent DNA," which is clearly nonribosomal, too, is not present in the pPC 170 clone. Furthermore, in pPC 170, the 32-bp sequence shares a repeated motif with sequences extending into the 3′-direction. Thus, it probably shares a common origin with these 3′-sequences.

This is not the case in the NTS, because here the 32-bp sequence does not seem to be related to the rest of the "Cla-adjacent DNA." At present we do not know how the linkage between the two parts of the NTS "Cla-adjacent DNA" can be explained. A possible explanation would be that the 528-bp "Cla-adjacent DNA" is the result of two independent transpositional events.

It is not known whether pPC 170 is the genuine origin of the "Cla-element/Cla-adjacent DNA" complex. Genomic Southern hybridizations show that the Cla-adjacent DNA (as well as the Cla-elements) is a highly repetitive and interspersed DNA sequence. The association of Cla-elements with the "Cla-adjacent DNA" thus might be more widespread than suggested by the two samples analyzed so far. It could be that the Cla-elements have a strong target site specificity, and that the "Cla-adjacent DNA" represents a typical target site for Cla-element integration. A more detailed analysis is necessary here.

There are also several points of evidence that the Cla-elements alone are, or have been, transposable elements. All Cla-element clusters investigated so far are surrounded by short sequence duplications, as they are typically found after transpositional events (Schmidt 1984). The same is true for the monomeric Cla-element found in pPC 170: there are six base-pair duplications at both ends of the *piger*-Cla-element, indicating an integration of the Cla-monomer at this site in pPC 170. The presence of more than two duplications can be explained, if at least two copies were already present prior to the integration events. Döring and Starlinger (1984) have found such multiple duplications at the insertion sites of maize transposable elements. The hypothetical target site duplications of the Cla-monomer are partially homologous to the sequence of the Cla-element itself. The same situation has been found in the case of inserted snRNA pseudogenes (van Arsdell and Weiner 1984; Bark et al. 1985). An integration mechanism involving homologous recombination between parts of the transposed element and the target locus could be a reasonable explanation for this finding. Additionally, the palindromic structure of the flanking "Cla-adjacent DNA" (see Fig. 2a) might play a role in targeting Cla-elements. The loop sequence itself displays a 9-bp homology to the oligo-dT/dA region at the end of a Cla-repeat (the ends of a Cla-repeat are defined by the ClaI restriction site; cf. Schmidt 1984). This organization strongly resembles that of the short *Strongylocentrotus* mobile repetitive element D881, where a 15-bp homology exists between the repeat

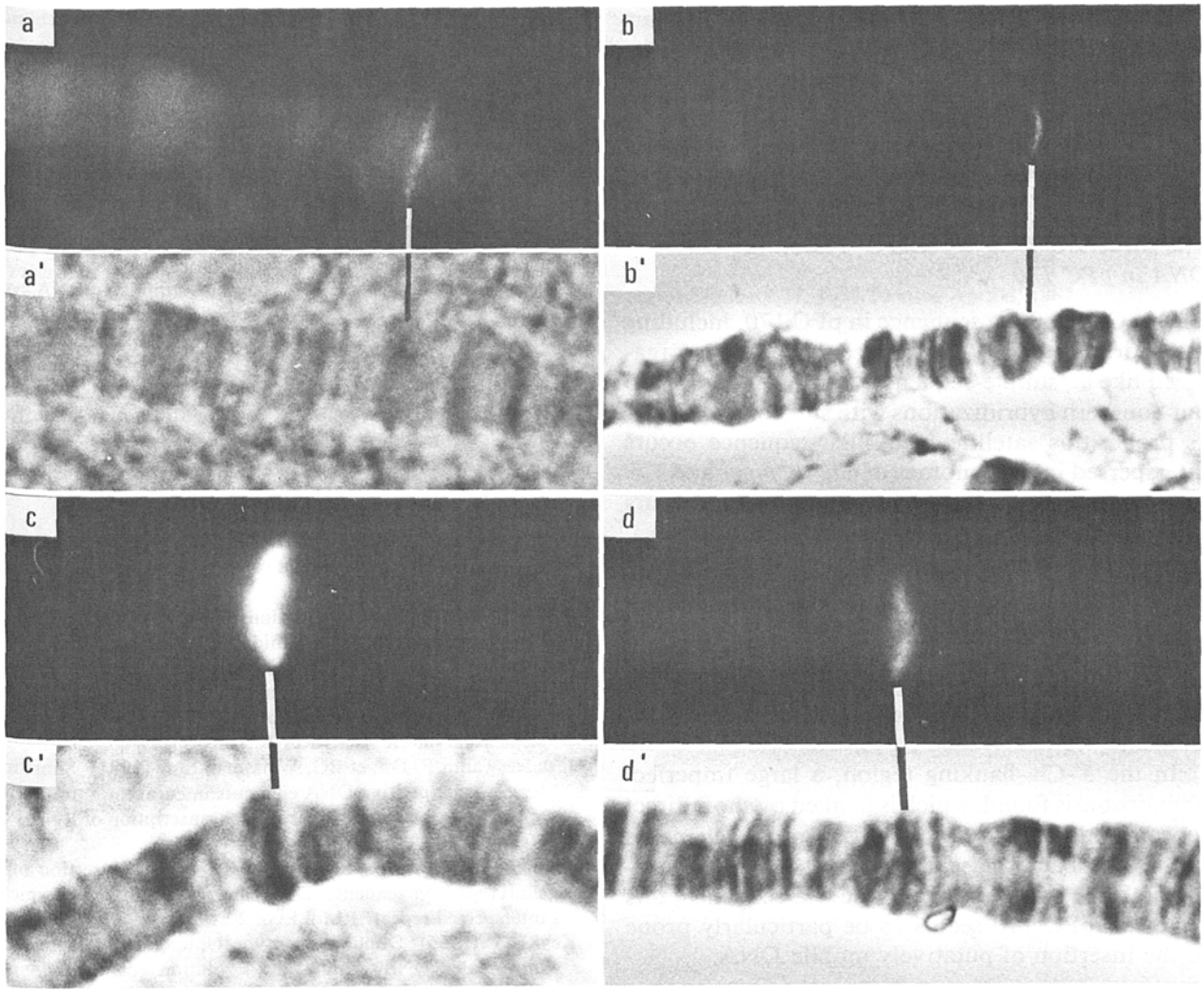


Fig. 5. In situ hybridization of biotin-labeled pBD 31 DNA to the polytene chromosomes of *Ch. th. thummi* (a) and *Ch. th. piger* (b, c, d). A single hybridizing band at position C2g of *thummi*-chromosome I can be seen (a), which is also present in *piger*-chromosome I (b). In addition to this single band present in both subspecies, a weak hybridization is found at three positions in the centromeric region of *piger*-chromosome I (b). Furthermore, the centromeres of *piger*-chromosomes II (c) and III (d) are labeled. (a', b', c', d') Phase-contrast photographs of the chromosomes shown in a, b, c, and d.

itself and its immediately flanking DNA (Johnson et al. 1984). Homologies between transposed DNA and the surrounding target sequence, involving short palindromic structures, have recently been reported for the insertion element Tgm1 found in soybeans (Rhodes and Vodkin 1985).

In view of their length, the Cla-elements might be classified as short interspersed repetitive elements (SINEs) (Singer 1982). Families of SINEs are the dominating repetitive sequences in the genomes of higher vertebrates. Several SINE families, including the Alu family of humans, are thought to be transposed via an RNA intermediate and reverse transcription (Jagadeeswaran et al. 1981) and thus termed "retroposons" (for review see Rogers 1985).

It is possible that Cla-elements have been transposed in a way similar to that proposed for retroposon SINE families. This possibility is supported by the fact that (1) at least some Cla-elements are transcribed (as shown by Northern hybridization; Schmidt, unpublished) and (2) Cla-elements possibly have a 3'-oligo-dT/dA end.

The second point is supported by the analysis of the monomeric Cla-element in pPC 170 described here. Because pPC 170 contains a monomeric Cla-element, there is no ambiguity of the ends of the Cla-sequence. The oligo-dA/dT cluster at the 3'-end of the Cla-element can easily fold back in an RNA molecule and thus serve as a priming end for reverse transcription (as proposed also for Alu-elements by Jagadeeswaran et al. 1981). Furthermore, the resolution of the hairpin structure after a possible self-priming could lead to short deletions of variable size, which are actually observed at the oligo-dA/dT end of various Cla-elements.

When the Cla-elements transpose together with

flanking DNA, the reverse transcription could start at structures like the pPC 170 3′-Cla-flanking region: a transcript of an array of repeats with weak palindromes could be the source of a number of different DNA molecules including various stretches of 3′-Cla-flanking DNA.

The Structure of the 5′- and 3′-Flanking DNA in pPC 170

The 3′-Cla-flanking sequence in pPC 170, including the short stretch of "Cla-adjacent DNA," is organized like a "simple-sequence" DNA. According to our Southern hybridizations with subclone pBD 421 as probe, this satellite DNA-like sequence occurs interspersed in the genomes of *Ch. th. piger* and *Ch. th. thummi* and spans a few hundred base pairs downstream of the Cla-element in pPC 170.

Sequence elements of the satellite-like DNA are not confined to the 3′-Cla-flanking region: two homologous sequences can be detected 5′ upstream of the Cla-element. Thus, the 5′- and 3′-Cla-flanking regions may be related, suggesting that this sequence has been interrupted by the integration of the Cla-element.

In the 5′-Cla-flanking region, a large imperfect palindrome is found, which is flanked by short direct repeats and is thus reminiscent of mobile elements of the "foldback" class (Potter et al. 1980). Regarding both the palindrome and the Cla-sequence, this region of pPC 170 seems to be particularly prone to the insertion of putatively mobile DNA.

In Situ Localization of pPC 170

The in situ localization of pPC 170 and of the monomeric Cla-element in the chromosomes of *Ch. th. piger* is difficult, because all Cla-element-flanking sequences are more or less also repetitive and hybridize in situ to a number of different locations. However, the subclone pBD 31, which contains the leftmost region of pPC 170, hybridizes to only one common site in the chromosomes of both subspecies, suggesting that this might be the original location of the whole sequence. A definite decision as to where the pPC 170 clone comes from can be made when the homologous region of the *Ch. th. thummi* chromosomes is analyzed. If the *Ch. th. thummi* homologous region contains the same structure of sequences as found in pPC 170, we can be sure that site C2g is the chromosomal origin.

Acknowledgments. We thank Prof. Dr. H.-G. Keyl for the working facilities, helpful suggestions on the manuscript, and cytological advice; Mrs. B. Weich for skillful technical assistance; Mrs. H. Sommerfeld for typing of the manuscript, and Mrs. S. Becker for the photographic work. This work was carried out with the support of the Deutsche Forschungsgemeinschaft.

References

Bark C, Hammarström K, Westin G, Petterson U (1985) Non-random integration of human U4 RNA pseudogenes. Mol Cell Biol 5:943–948

Benton WD, Davis RW (1977) Screening λgt recombinant clones by hybridization to single plaques in situ. Science 196:180–182

Bolivar F, Backman K (1979) Plasmids of *Escherichia coli* as cloning vectors. Methods Enzymol 68:245–267

Bonner JJ, Pardue ML (1976) Ecdysone-stimulated RNA synthesis in imaginal discs of *Drosophila melanogaster*. Chromosoma 58:87–99

Boseley PG, Moss T, Birnstiel ML (1980) 5′ Labeling and poly(dA) tailing. Methods Enzymol 65:478–495

Döring HP, Starlinger P (1984) Barbara McClintock's controlling elements: now at the DNA level. Cell 39:253–259

Feinberg AP, Vogelstein B (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal Biochem 132:6–13

Frischauf AM, Lehrach H, Poustka A, Murray N (1983) Lambda replacement vectors carrying polylinker sequences. J Mol Biol 170:827–842

Hägele K (1970) DNS-Replikationsmuster der Speicheldrüsen-Chromosomen von Chironomiden. Chromosoma 31:91–138

Israelewski N, Schmidt ER (1982) Spacer size heterogeneity in ribosomal DNA of *Chironomus thummi* is due to a 120 bp repeat homologous to a predominantly centromeric repeated sequence. Nucleic Acids Res 10:7689–7700

Jagadeeswaran P, Forget BG, Weissman SM (1981) Short interspersed repetitive DNA elements in eucaryotes: transposable elements generated by reverse transcription of RNA Pol III transcripts? Cell 26:141–142

Johnson SA, Davidson EH, Britten RJ (1984) Insertion of a short repetitive sequence (D881) in a sea urchin gene: a typical interspersed repeat? J Mol Evol 20:195–201

Keyl HG (1965) Duplikationen von Untereinheiten der chromosomalen DNS während der Evolution von *Chironomus thummi*. Chromosoma 17:139–180

Keyl HG, Strenzke K (1956) Taxonomie und Cytologie von zwei Subspezies der Art *Chironomus thummi*. Z Naturforsch 116:727–735

Maxam AM, Gilbert W (1980) Sequencing end-labelled DNA with base specific chemical cleavages. Methods Enzymol 65: 499–560

Potter S, Truett M, Phillips M, Maher A (1980) Eucaryotic transposable genetic elements with inverted terminal repeats. Cell 20:639–647

Rhodes PR, Vodkin LO (1985) Highly structured sequence homology between an insertion element and the gene in which it resides. Proc Natl Acad Sci USA 82:493–497

Rigby PWJ, Dieckmann M, Rhodes C, Berg P (1977) Labeling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. J Mol Biol 113:237–251

Rogers JH (1985) The origin and evolution of retroposons. Int Rev Cytol 93:187–279

Rohwedder A (1986) Lokalisation und Aufbau eines Integrationsortes repetitiver DNA-Sequenzen im Genom von *Chironomus thummi*. Diploma Thesis, Ruhr-Universität, Bochum, West Germany

Schaefer J, Schmidt ER (1981) Different repetition frequencies of a 120 base-pair DNA-element and its arrangement in *Chironomus thummi thummi* and *Chironomus thummi piger*. Chromosoma 84:61–66

Schmidt ER (1981) The development of a 120 basepair repetitive DNA sequence in *Chironomus thummi* is correlated to the duplication of defined chromosomal segments. FEBS Lett 129:21–24

Schmidt ER (1984) Clustered and interspersed repetitive DNA sequence family of *Chironomus*. J Mol Biol 178:1–15

Schmidt ER, Godwin EA (1983) The nucleotide sequence of an unusual non-transcribed spacer and its ancestor in the rDNA in *Chironomus thummi*. EMBO J 2:1177–1183

Schmidt ER, Vistorin G, Keyl HG (1980) An AT-rich DNA component in the genomes of *Chironomus thummi thummi* and *Chironomus thummi piger*. Chromosoma 76:35–45

Schmidt ER, Godwin EA, Keyl HG, Israelewski N (1982) Cloning and analysis of ribosomal DNA of *Chironomus thummi piger* and *Chironomus thummi thummi*. Chromosoma 87: 389–407

Singer MF (1982) SINEs and LINEs: highly repeated short and long interspersed sequences in mammalian genomes. Cell 28: 433–434

Southern EM (1975) Detecting of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 98:503–517

Van Arsdell SW, Denison RA, Bernstein LB, Weiner AM (1981) Direct repeats flank three small nuclear RNA pseudogenes in the human genome. Cell 26:11–17

Received January 27, 1987/Revised April 4, 1987