Proc. Natl. Acad. Sci. USA
Vol. 83, pp. 125–129, January 1986
Genetics

Origin and evolution of the transcribed repeated sequences of the Y chromosome lampbrush loops of *Drosophila hydei*

(functional repeats/nonconserved sequences/evolutionary implications)

DANA HAREVEN, MATHI ZUCKERMAN, AND ELIEZER LIFSCHYTZ

Department of Biology, Technion Haifa 32000, Israel

Communicated by Dan L. Lindsley, August 12, 1985

ABSTRACT **The molecular evolution and patterns of conservation of clones from four Y chromosome lampbrush loops of *Drosophila hydei* were investigated. Each loop contains a discrete family of transcribed repeats that are only slightly conserved even in the *hydei* subgroup species. Sequencing of clones from the four *D. hydei* loops indicates that all transcribed repeats evolved from A+T-rich elements of the genome. Evidence is presented that suggests a Y-specific family evolved as a result of the transposition of repeated sequences from an autosomal site to the Y chromosome with the concomitant acquisition of transcriptional activity and loss of non-Y sequences. The results support a structural role for the loops in shaping a spermatocyte-specific nuclear organization. Transcribed heterochomatic sequences could play a similar role in nuclear organization in many cell types.**

Particular chromosomal lampbrush loops originate from specific sites along the Y chromosome of *Drosophila* species. They are found only in the primary spermatocyte nuclei and are probably essential for normal differentiation of male germ-line cells (1). Y loops have been most thoroughly studied in *D. hydei*, where six different loop pairs were mapped (1). The number of loops in *D. hydei* conforms with the number of Y-linked male fertility complementation units in *D. melanogaster* (2–4).

Analysis of cloned sequences from four different loops of *D. hydei* has shown that the loops consist of tandemly arranged repeated sequences. These sequences are transcribed, the transcripts varying in size from 10S to 60S. The transcripts are confined to the nuclei of primary spermatocytes and disappear, along with the loops, prior to first meiotic metaphase (refs. 5, 7; unpublished data).

We have proposed that the loops help to regulate spermatogenesis by providing the primary spermatocyte nucleus with a matrix required for the proper compartmentalization of gene activity and sequestration of gene products for postmeiotic differentiation (ref. 5; unpublished data).

Even closely related sibling species of *Drosophila* have morphologically different loops and even different nuclear shape (refs. 1, 8, and 9; Fig. 1). Because of this morphological diversity, and assumed similar function, we undertook a comparative sequence analysis of clones from each of four different loops of *D. hydei*. We also analyzed their sequence conservation by homology tests between the *D. hydei* loop clones and other *Drosophila* species.

MATERIALS AND METHODS

Materials, *Drosophila* Stocks, and Y-Specific Clones. Restriction enzymes were from New England Biolabs, Boehringer Mannheim, and Bethesda Research Laboratories; [α-^{32}P]dATP (800 Ci/mmol; 1 Ci = 37 GBq) and [^{3}H]dATP (100 Ci/mmol) were from Amersham. *D. neohydei* and *D. eohydei* were kindly provided by O. Hess (Düsseldorf) and *D. hydei* by H. Gloor (Geneva). *D. bifurca* was from the Bowling Green stock center. The origin of the Y-specific clones has been described (5, 10). Y23Ns is 0.67 kilobases (kb) long; Y20Ns, 1.75 kb; Y18CI, 2.0 kb, and Y22Tr, 1.6 kb. All are *Eco*RI genomic fragments.

Nucleic Acids. Restriction fragments were subcloned in M13 mp8 or M13 mp9 (11) for sequencing by the chain-termination method (12). Single-stranded probes used in hybridization experiments were prepared according to Hu and Messing (13).

Blot hybridizations were done according to Southern (14), and dot tests and RNA transfers as described by Thomas (15). The DNA·RNA *in situ* hybridization procedure (16) was detailed by Lifschytz *et al.* (5), and *in situ* hybridizations to metaphase chromosomes (17) were as described by Lifschytz and Hareven (18). All *in situ* hybridization experiments were conducted under the stringent conditions specified in the above references for intraspecies tests. This is especially important in light of the high A+T content of all sequences.

RESULTS

Clones representing transcribed repeated sequences from four lampbrush loops of *D. hydei* (Y20Ns and Y23Ns of the two "nooses" loops, Y18CI of the "clubs" loop and Y22Tr of the "tubular ribbons" loop) were tested for homology with other species. The sequences of *D. hydei* loops bear no homology to spermatocyte RNA, or male genomic DNA, of the distantly related species *D. melanogaster* and *D. virilis* or even of the closely related species *D. repleta* (data not shown). We therefore proceeded to the analysis of homologies with three species within the *hydei* subgroup. The primary spermatocytes of these species are shown in Fig. 1. They are arranged according to their phylogenetic distance from *D. hydei* (Fig. 1*a*), with *D. neohydei* (Fig. 1*b*) being the closest, and *D. bifuca* (Fig. 1*d*) the most remotely related species (19).

The "clubs" clone, Y18CI, includes a 0.9-kb (*Eco*RI–*Hin*dIII) fragment, which is transcribed, and a 1.1-kb nontranscribed fragment (Fig. 2*a* and unpublished data). The transcribed *Eco*RI–*Hin*dIII fragment is characterized by a high A+T/G+C ratio and reiteration of a 7-base consensus sequence, GATTGAT. A perfect 40-base palindrome at the beginning of the satellite-like (i.e., short repeats) section is underlined (Fig. 2*b*). The 1.1-kb nontranscribed sequence (Fig. 2*c*) is also repeated and is clubs-specific, but in contrast has a 48% A+T composition.

The results of *in situ* hybridization of the clubs clone to RNA of primary spermatocytes and DNA of metaphase chromosomes of the four species from the *hydei* group are shown in Fig. 3. Strong homology with nuclear RNA is found in *D. hydei* (Fig. 3*a*) and is at least as strong in the closely

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: kb, kilobase(s); bp, base pair(s).

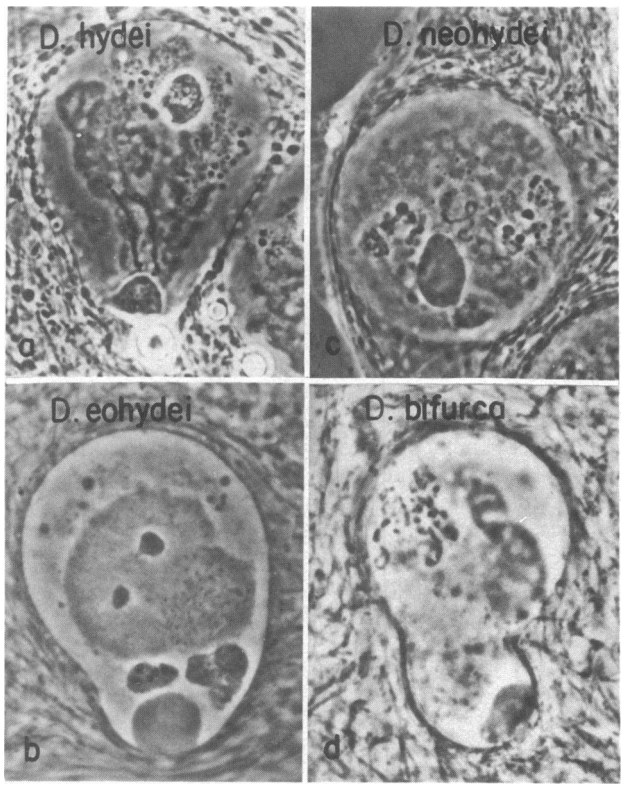


FIG. 1. Living primary spermatocytes of four *hydei* species. Note the variation among species in loop structure and nuclear shape.

related species *D. neohydei* (Fig. 3*c*). No silver grains were observed in the primary spermatocytes of *D. eohydei* or *D. bifurca* (Fig. 3*g*) after a comparable exposure time (10–14 days). A weak signal for *D. eohydei* (Fig. 3*e*) was obtained only after 50 days of exposure.

Hybridization to metaphase chromosomes revealed major clusters of clubs sequences on the long arm of the Y chromosome of *D. hydei* (Fig. 3*b*) and *D. neohydei* (Fig. 3*d*). No major site on any chromosome was observed in *D. eohydei* (Fig. 3*f*) but, most surprisingly, clusters of homologous sequences were detected in the dot chromosomes of the least-related species, *D. bifurca* (Fig. 3*h*).

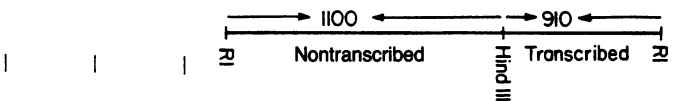
Southern blot hybridizations gave very similar banding patterns (Fig. 4*c*) for male and female *D. bifurca*, where Y18CI DNA is homologous with nontranscribed, autosomal sequences (Figs. 3 *g* and *h*). Male and female DNA of *D. eohydei* (Fig. 4*b*) showed nearly identical patterns, but a few *Hin*fI bands (arrows in Fig. 4*b*) were detected in male but not female DNA. We presume that they represent a small fraction of clubs sequences that are Y-linked but poorly transcribed in the primary spermatocyte of this species (see Fig. 3*e*). In *D. neohydei* almost all clubs sequences are Y-linked (Fig. 4*a*) and transcribed in the primary spermatocyte (Fig. 3*c*). This is similar to the situation in *D. hydei* (5).

Clones Y20Ns and Y23Ns were localized to separate domains of the nooses loops on the short arm of the Y chromosome and shown to have a 60% A+T content (ref. 20; unpublished data).

The results of hybridization with the nooses clone Y23Ns are shown in Fig. 5. Compared with the response in the spermatocytes of *D. hydei* (Fig. 5*a*), a much weaker signal was found in the primary spermatocytes of the sibling species, *D. neohydei* (Fig. 5*b*). Yet an appreciable cluster of sequences occur on short arm of the Y chromosome (Fig. 5*d*). Very little homology (or few sequences) were observed with nuclear RNA of *D. eohydei* (Fig. 5*e*) and none was detected, even after long exposure, with the *D. bifurca* spermatocyte RNA (Fig. 5*f*). We failed to detect hybridization of the nooses sequence with *D. eohydei* or *D. bifurca* metaphase chromosomes. Southern blots and dot-blot tests showed no homology in the genomic DNA of *D. bifurca* males but indicate that an appreciable amount of nooses-like sequences are Y-linked in *D. eohydei*.

The tubular ribbons clone, Y22Tr, is composed of three erratically arranged A+T-rich repeats. Repeat I is 73 bp long

a



b Y18, Clubs, transcribed

```
  1  AGCAAAGCTTAAAGTTGTTTTATTCAGAAACGATTGAAGAAGTGGGTGAATAAAGAGTGATGTATAGAGAGATGATTGATCATAATAAGGGACAGATGTA
101  TGTTAGACGACGATGATGACTGATTTATAAAATAATTGATATAAGAATACATTGAAAATGAAAATTGGTTGATCAAACTGATTAAAGGATTGAGGAACTA
201  AAGGATAAGGGACGGTCAATTAATAGATTGATGTAGATTGATGAAAAAATTCGATTATAAATGGGTAATGGATAGATGATTAATTGATGGATTGAATGAT
301  GATTTACGATTGATTGATTGATGATTAATGATTGAATGATAATTAATGATTGATTGATTGACTGATGATTGATGGTTATGATTGATGGTTATGATTGATG
401  ATTGATGATTAATGATTGATGATTGATGATTGATGATTGATTGATGATATTATGATTGATTGACTG/ATTGATGATTGATTGACTGATTGATTGATGATT
501  GATGGTTGATGATTGATGATTAATGATTGAAGATTGATTGATGGTTGATGATTGATGATTGATGATTGAGAAGAGATGATTGATGATTGATGGTTGATGA
601  TTGATGATGGTCGATTGATAATTGATTGATTAATGATTGATGATTACTATGAATGGATTGATATATATATATATATATCTATAGATATATATATATATAT
701  AATAGGAAGTAAAGTGATGGGATGGGTGGAAAGTTCGCTGATTGACTCACGCACAGGAATTTAAATTGATGAACAAATTTCTGAGTTTTGATTGATTAGT
801  GCGTAGTGAGAATGTGATTGAATAATGCACAGTAAATTTCCAGATCAATCATGTAAGAATTC
```

c Y18, Clubs, nontranscribed

```
  1  CACAGTCCGAGCGCGGGACTGAGTGCGGCTGAGCAAGCCGTTGACATGACGGGCGACATTGAGGTGAGTGCATAAGACCGATAAAAAAAAAAGCAGCAAT
101  CATTAAGCTAGCTCTTGTTATTGGCTTTTTACAGCCTGTGTCATCAGCGCTGGCTGAACGAGGCTCTTGCCACGTCGTACAGGAGAGTCGCGGCGCCCAG
201  ATCTATGAGTGCTACAATAATACGACGGGCGCCAGCGATAATCTCAGCGAACAGCTGCTGGGCCAATATGTCACCAGATGCTGGCCACGGCACGTGCTCT
301  CCACTTCACACGGCAGACGGAGAAGGCCATTGGTGGCATGTCGTTGGCCAAGCTGCAGCAGCATAATGTGATGAGCCCCAGCGTTGGCATTGGCTGAGTC
401  AGCTGCGATACTTACATGGGCGGCGGCACAACAAGGCGGCCGATGCCGCAGGTGTGGGTGCGCATGGTTTATGCTGAATTTGAATACGCATACGAATTTC
501  TCGGACATGCCAATAAGCTCGTGCACATTTGGCTCACACACAAACTGCTATCTAATACTCACACAGGCCATGCATATGGCCTTGCGGTAATCGTTCGGTC
601  CGCGGCACGGGCAAAACGGAGAGCGTTAAGGCCCTGGGCGGCATGCTGGGAACGCCTAGTTCTGGTCTTCAACTGTGATGAGGTGAGGGTCCCAGGCTCG
701  TCGTGTCGATTTGTCATAATATTTCTCTGTGTGTGCTTGCAGAACGTGTACACCGAATCGATGAGCCTAATCCTAACTGGTTTGGCCGCTGCGGCCCTGG
```

FIG. 2. Sequence analysis of the clubs clone. (*a*) Major restriction sites of the Y18CI clone. Lengths are given in base pairs (bp). RI, *Eco*RI. (*b*) Nucleotide sequence of the transcribed 900-bp *Eco*RI–*Hin*dIII segment of Y18CI. The 7-base consensus repeat and a complete 40-base palindrome are underlined. About 50 bases in the middle of the segment where the GATTGAT repeats overlap were not sequenced. (*c*) Sequence of the nontranscribed region of Y18CI. All 1100 nucleotides were sequenced, but an 800-bp segment is illustrated.

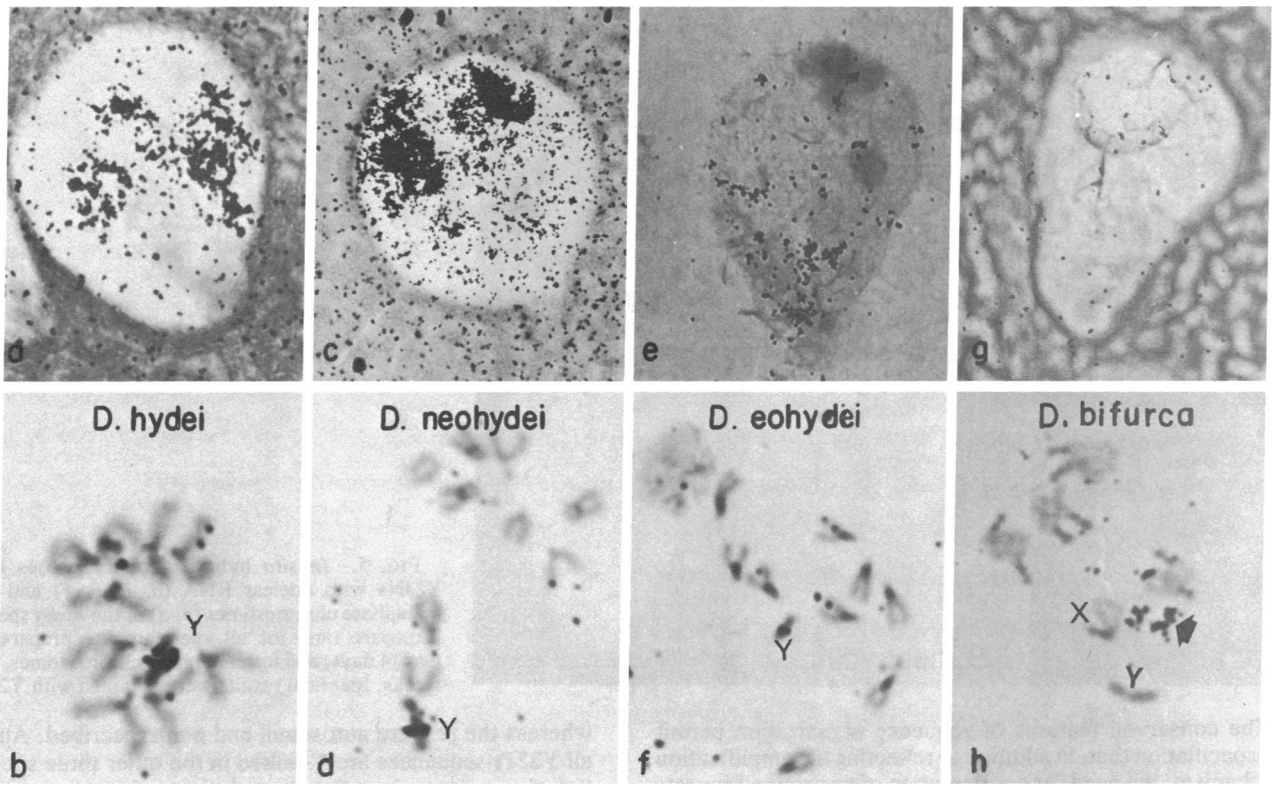


FIG. 3. *In situ* hybridization of the clubs clone (Y18CI) with nuclear RNA (*a*, *c*, *e*, and *g*) and with metaphase chromosomes (*b*, *d*, *f*, and *h*) of the *hydei* species. The slight signal shown by nuclei of *D. eohydei* (*e*) was obtained after a 50-day autoradiographic exposure; no signal was seen after the usual 14-day exposure.

and 63% A+T, whereas II and III are 55 and 57 bp long, respectively, and both are 75% A+T (Fig. 6). The proportions of I/II/III repeats in the cloned sequence are 8:3:2; two typical arrangements are I-I-III-I-I-II-I or I-II-I-III-I-II-I. As with Y18CI, all repeat units are cooriented, no head-to-head, or tail-to-tail arrangements were found.

Y22Tr is the only clone that hybridizes to nuclear RNA sequences in all four species of the *hydei* group (Fig. 7). The extent of homology is progressively reduced from *D. hydei* to *D. bifurca*, reflecting the evolutionary relatedness of the species. Male and female DNA of *D. bifurca* responded weakly but equally in dot-hybridization with Y22Tr probe, but attempts to localize the homologous sequences on chromosomes failed, probably because of homology. Unexpectedly, a cluster of homologous sequences was found on the X chromosome of *D. neohydei* (Fig. 7*d*) but not in *D. hydei* or *D. eohydei* (Figs. 7 *b* and *f*). We do not know whether the X-linked Y22Tr cluster in *D. neohydei* is transcribed in the primary spermatocyte.

In *D. hydei*, all repeats of a loop-sequence family are transcribed from the same strand (unpublished data). Since Y22Tr is the only clone that is homologous to spermatocyte RNA of all *hydei* species, we have tested the homology of both its strands with spermatocyte RNA of the other three species. Only one strand was complementary with spermatocyte RNA of all the species (as in Fig. 7), showing that through the evolution of the *hydei* subgroup, Y22Tr coding sequences have been retained in the same single strand.

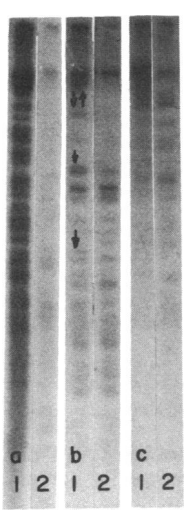


FIG. 4. Evolution of the Y18CI family from autosomal repeats to Y-specific loops. DNA from *D. neohydei* (*a*), *D. eohydei* (*b*), and *D. bifurca* (*c*) were digested with *Hin*fI, and 5-μg samples were electrophoresed in 1.2% agarose gel and transferred according to Southern (14). The hybridization probe was ^{32}P-labeled (nick-translated) Y18CI. Lanes 1: male DNA (partial *Hin*fI digest for *D. bifurca*). Lanes 2: female DNA.

DISCUSSION

Different Y-Loop Domains Evolved as Tandem Repeats of A+T-Rich "Founder" Sequences. The transcribed sequences of the four loops are characterized by a high A+T composition and by relatively short repeating units: a 7-bp repeat in Y18CI, a 55–75 bp repeat in the tubular-ribbons clone Y22Tr, and an ≈580-bp repeat in the nooses domain (20). The only exception is the nontranscribed but repeated clubs sequence.

Although all the transcribed sequences are A+T-rich, they are not related, on the basis of their sequence, to one another. The idea that different loops have a common origin (1) is substantiated only in the case of the nooses loops (20) but should be discarded as a general concept.

Sequence analysis reveals that all repeated elements within a given clone are cooriented, and it has been shown before that all transcripts are made from only one strand (unpublished data). Furthermore, only the transcribed strand of the Y22Tr clone is homologous with the nuclear RNA of all the *hydei* species (Fig. 7).

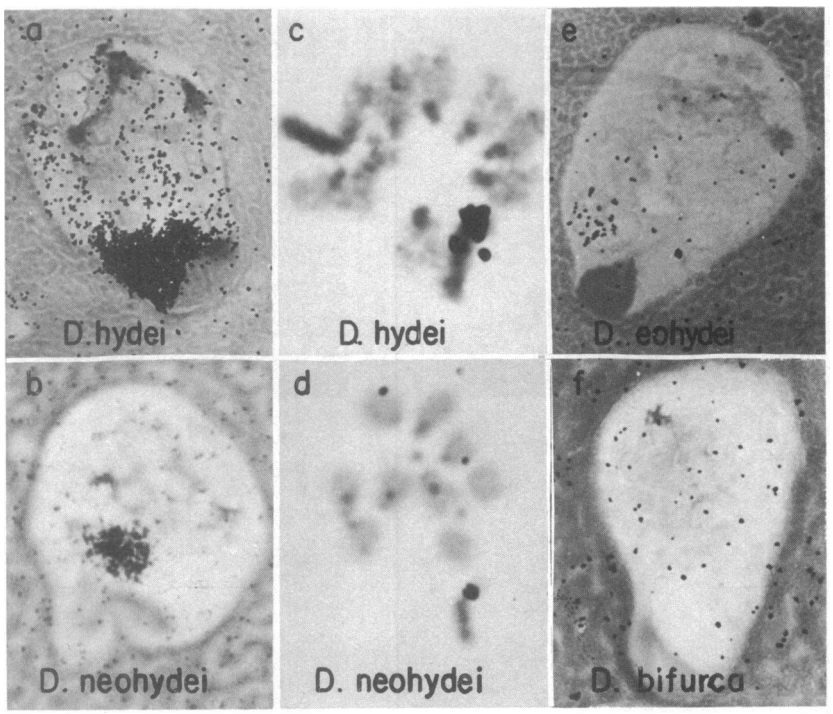


FIG. 5. *In situ* hybridization of nooses clone Y23Ns with nuclear RNA (*a, b, e, f*) and with metaphase chromosomes (*c, d*) of the *hydei* species. Exposure time for all spermatocyte preparations was 14 days, and for metaphase chromosomes, 8–12 weeks. Identical results were obtained with Y20Ns.

The conserved features of sequence organization permit the speculation that, in addition to reflecting the amplification mechanism, the need for coorientation of transcribed repeats played a role in the evolution of the loop domains from founder sequences.

The Dynamic Evolution of Loop Sequences. Clubs transcripts are abundant in the nuclear RNA of *D. hydei* and *D. neohydei* but are rare in *D. eohydei* and nonexistent in *D. bifurca*. The majority of Y18CI chromosomal repeats in *D. hydei* and *D. neohydei* are Y-linked; there are few repeats on the Y chromosome of *D. eohydei* (Figs. 3 and 4) and they are exclusively autosomal in *D. bifurca*. The proportion of sequences that are Y-linked decreases concomitantly with the level of spermatocyte transcripts of these sequences.

We suggest that the change in the fate of the clubs family was initiated by a transposition of some sequences from an autosomal site to the long arm of the Y chromosome, where they acquired the ability to abide by the rules of Y-chromosome activation and evolved as a chromosomal loop by repeated duplications. At the same time, the autosomal sequences were eliminated, presumably by selection of variant karyotype with the least non-Y repeats.

The Y22Tr clone is the only one with homology to spermatocyte nuclear RNA of all *hydei* species. The gradual decrease in hybridization signal (Fig. 7) follows the taxonomic relations of the four species (19). Since male and female DNA from *D. bifurca* respond equally in dot-hybridization (data not shown), it is possible that the state of the Y22Tr family in this species is similar to that of Y18CI in *D. neohydei*: some sequences may be Y-linked and transcribed, whereas the rest are autosomal and nontranscribed. Almost all Y22Tr sequences are Y-linked in the other three species, but the occurrence of a homologous cluster on the X chromosome of *D. neohydei* underscores the possibility that evolved loop families were derived from mobile elements in the genome.

From the pattern of homology with transcripts of the four species we surmise that the tubular-ribbons sequences represent the most ancient transcribed loop sequence. The nooses loops must be the newest addition to the loops of the *hydei* species, as they are virtually nonexistent in *D. bifurca* and are represented but scarcely transcribed in *D. eohydei*. They are approximately equally abundant in the two sibling species, *D. neohydei* and *D. hydei* (data not shown), but are relatively weakly transcribed in the spermatocytes of *D. neohydei* (Fig. 5).

Evolution and Function of Y Loops. The Y chromosome of *Drosophila* species exhibits extreme structural and functional divergence, far beyond that of the rest of the chromosome complement. There are some species in which no Y chromosome is found (21) and its function must be fulfilled by other chromosomes. Within the *hydei* group, the Y chromosome varies from a long acrocentric chromosome in *D. hydei* to a dot element in *D. nigrohydei* (19). It is possible that in intermediate cases some "authentic" Y-linked fertility factors are carried by autosomes (22). We have seen that loop structure and nucleotide sequence varies widely among the species. This variation contrasts with the conservation of unique-sequence genes (23, 24) but is in accord with the

```
I    CCAGGAAACTGCAATCTTATTTTATATTCAGATTAATGCAGGATTGCCTGTTCTGGTCTAACCTTTTCTAT
     ******          * *   *                 *   *  *    *   * * *
     -ATTAGAACTCTGAATCTAA-TTTA-TATCAGATTAATGCAGG-TTGTCTCTTCTGTTCTATC-TTTCTAT
II   AAATGTACCATATATATGAGTTAAAATAGTGTTATCTCTTCTGGTATAATTATTTT
          *   **    *                *  *       *
     AAATGTTGCCATGAATATGTGTAAAATAGTGTTAACACTTCTGGTCTAATTATTTT
III  ATATGAACCATATATATAAGAGGGCTATATATATGTGGAACAGGTTGTTTCTTTATT
                 *          ****      *** *
     ATATGAACCATCTATATAAGAGGC---TATATAGTCGAAACAGGTTGTTTCTTTATT
```

FIG. 6. Sequence analysis of the tubular-ribbons clone (Y22Tr). The *Eco*RI insert is composed of various combinations of three A+T-rich unit sequences: I, 73-bp unit; II, 55-bp unit; III, 55- to 57-bp unit. Comparison of two reiterated sequences of each type are shown. Palindromic sequences capable of forming small stem and loop structure are underlined in the type III units.

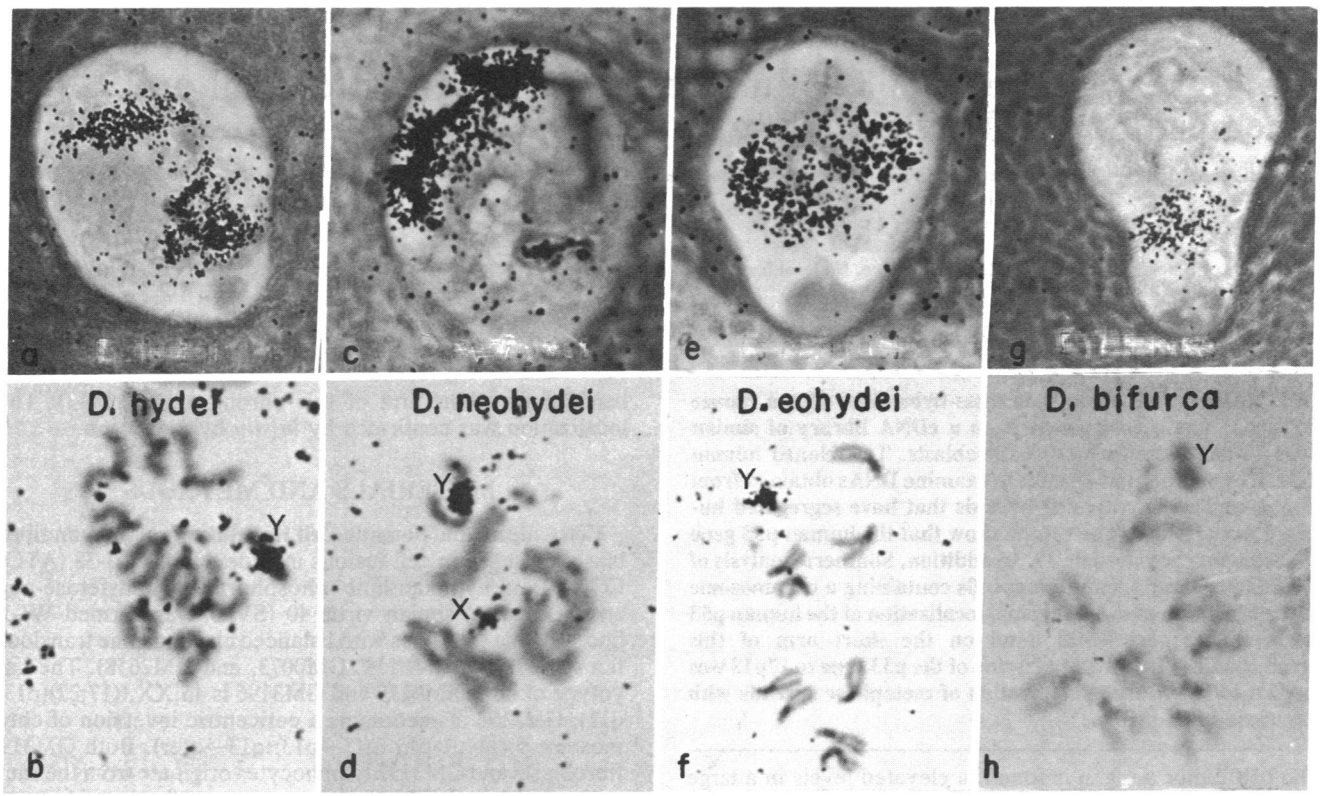


FIG. 7. *In situ* hybridization of the tubular-ribbons clone (Y22Tr) with nuclear RNA (*a, c, e, g*) and with metaphase chromosomes (*b, d, f, h*) of the *hydei* species. Exposure time for spermatocytes 14 days, and for metaphase chromosomes, 7 weeks. Note the small cluster of homologous sequences in the heterochromatic section of the X chromosome of *D. neohydei*.

observation made for mobile middle-repetitive sequences (6).

Presumably similar functions are fulfilled by chromatin structures of diverse morphology and, as shown here, widely different base sequences. This reinforces the view that Y loops regulate spermatocyte development by providing the correct "spermatocyte-specific" nuclear matrix rather than by supporting the expression of loop-specific coded functions (ref. 5; unpublished data).

Due to their unusual size and preparatory roles in the primary spermatocyte nucleus, the Y chromosome lampbrush loops may represent an exaggerated situation. In other differentiating nuclei, chromatin structures formed transiently by the unfolding and transcription of other clustered repeats (heterochromatin?) could have functions similar to those of Y loops.

We thank Drs. W. J. Peacock and T. Peterson for their invaluable help in writing this manuscript. This work was supported by a grant from the Committee for Basic Research of the Israeli Academy of Sciences.

1. Hess, O. & Meyer, G. F. (1968) *Adv. Genet.* **14,** 171–223.
2. Brosseau, G. E. (1960) *Genetics* **45,** 257–274.
3. Kennison, J. A. (1981) *Genetics* **98,** 529–548.
4. Gatti, M. & Pimpinelli, S. (1983) *Chromosoma* **88,** 349–373.
5. Lifschytz, E., Hareven, D., Azriel, A. & Brodsley, H. (1983) *Cell* **32,** 191–199.
6. Dowsett, A. P. (1983) *Chromosoma* **88,** 104–108.
7. Vogt, P., Hennig, W. & Sigmund, I. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 5132–5136.
8. Meyer, G. F. (1963) *Chromosoma* **14,** 207–255.
9. Meyer, G. F. (1965) *Drosoph. Inf. Serv.* **40,** 80.
10. Lifschytz, E. (1979) *J. Mol. Biol.* **133,** 267–277.
11. Messing, J. & Vieira, J. (1982) *Gene* **19,** 269–276.
12. Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H. & Roe, B. A. (1980) *J. Mol. Biol.* **143,** 161–178.
13. Hu, N. & Messing, J. (1982) *Gene* **17,** 271–277.
14. Southern, E. (1975) *J. Mol. Biol.* **89,** 503–517.
15. Thomas, P. S. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 5201–5205.
16. Pukkila, P. J. (1975) *Chromosoma* **53,** 71–89.
17. Pardue, M. L. & Gall, J. G. (1975) *Methods Cell. Biol.* **10,** 1–16.
18. Lifschytz, E. & Hareven, D. (1982) *Chromosoma* **86,** 429–442.
19. Wassermann, M. (1962) *Univ. Tex. Publ.* **6205,** 63–72.
20. Lifschytz, E. & Hareven, D. (1985) *Mol. Gen. Genet.* **199,** 188–194.
21. Patterson, J. T. & Stone, W. S. (1952) *Evolution in the Genus Drosophila* (MacMillan, New York), p. 1952.
22. Lifschytz, E. & Hareven, D. (1977) *Dev. Biol.* **88,** 276–294.
23. Sturtevant, A. H. & Novitski, E. (1941) *Genetics* **26,** 517–541.
24. Brock, H. W. & Roberts, D. B. (1983) *Genetics* **103,** 75–92.