Proc. Natl. Acad. Sci. USA
Vol. 90, pp. 10904–10906, December 1993

Commentary

Conventional protein coding genes in the *Drosophila* Y chromosome: Is the puzzle of the fertility gene function solved?

Wolfgang Hennig

Department of Molecular and Developmental Genetics, University of Nijmegen, Toernooiveld, 6525 ED Nijmegen, The Netherlands

The Y chromosomal fertility genes of *Drosophila melanogaster* have been studied for almost 80 years (1) but the secrets of their function have remained elusive. Even their number and their location on the Y chromosome were recently revised (2, 3), as the Y chromosome poses serious difficulties for genetic mapping due to the absence of marker genes. It is now accepted that two fertility genes (*ks-1* and *ks-2*) reside in the short arm and four (*kl-1*, *-2*, *-3*, and *-5*) [contrary to the originally supposed five (4)] reside in the long arm of the Y chromosome of *D. melanogaster*.

A new area of research on the fertility genes started when Meyer *et al.* (5) discovered that the activity of these genes is accompanied by the formation of large lampbrush loops in primary spermatocyte nuclei. Spermatocytes are the only cell type in which the Y chromosomal fertility genes are active. Soon it became clear that Y chromosomal lampbrush loops are a universal feature of spermatocytes in the genus *Drosophila* (6). In *D. melanogaster*, these lampbrush loops are relatively small and difficult to distinguish. Therefore, most of the recent studies of the fertility genes have been carried out in *Drosophila hydei*. While *D. hydei* has advantages for biochemical work, the genetics of this species is only poorly developed (for a recent review see ref. 7).

For *D. hydei*, it has been concluded that each Y chromosomal lampbrush loop is related to a single complementation group (8). By hybridization experiments it has been shown that the Y chromosome codes for transcripts of repetitive DNA sequences specifically expressed in spermatocytes (9, 10). In both *D. melanogaster* and *D. hydei*, however, attempts to identify protein products related to the expression of the fertility genes have not been very successful (11, 12). However, evidence was presented (13) that the *D. melanogaster* Y chromosome affected a dynein-like protein, although these experiments could not distinguish between the presence of a structural gene or a regulatory effect.

Also, extensive studies of the DNA composition of the lampbrush loop-forming fertility genes did not provide any evidence for protein coding of these genes. In *D. melanogaster*, only satellite DNAs have been found as transcribed constituents of the lampbrush loops (14). The molecular structure of some of the lampbrush loops has been investigated in much more detail in *D. hydei* (for review, see ref. 15). These loops, similar to those of *D. melanogaster*, are composed of a satellite-like type of DNA present in small tandem repeat clusters, which are interspersed by members of other repetitive DNA sequence families that for a large part, if not entirely, are defective transposable elements (16–21). In these studies, no DNA sequences coding for functional proteins were discovered, although considerable stretches of DNA were sequenced.

The report of Gepner and Hays (22) in this issue of the *Proceedings* that the ATP-binding domain of a presumably functional dynein gene is found associated with the *kl-5* fertility gene of *D. melanogaster* therefore provides solid evidence for the presence of a protein-coding gene not only on the Y chromosome but also related to one of the fertility genes. The assignment to the *kl-5* fertility gene is based on the use of a deficiency of the region of the Y chromosome accommodating the *kl-5* gene.

Some indirect earlier evidence for the presence of potentially protein-coding genes came from my laboratory (21). Leoncini had isolated a series of temperature-sensitive alleles of the fertility genes of *D. hydei*. However, as conditional tRNA mutants are known, the possibility remained that the conditional mutations may be caused by defects other than in protein structure. We have not been able to identify the target sequences in the DNA responsible for the conditional mutations. The observations of Gepner and Hays imply that the mutants recovered by Leoncini may indeed be mutations in protein-coding genes.

The results reported by Gepner and Hays raise the important question of how the Y chromosomal dynein relates to the fertility genes. If it does, what are the consequences in the context of our knowledge of the molecular structure of the transcription units in the lampbrush loops? The data provided by Gepner and Hays imply that a defective dynein gene results in male sterility due to defects in development caused by the absence of dynein. However, they do not give any information concerning a relationship between the transcription of the dynein gene and the transcription units appearing as lampbrush loops. By Miller spreading experiments it was earlier shown that the lampbrush loops formed by the fertility genes of *D. hydei* are giant transcription units (23, 24). From indirect cytogenetic evidence, similar large sizes have later been postulated for the fertility genes of *D. melanogaster*.

The most direct conclusion would be to assume that such a large transcription unit, if coding for dynein, has a large number of introns interspersed between a series of small exons or has few exons spaced by very large introns. The results communicated by Gepner and Hays give information neither on the genomic length of the Y chromosomal dynein gene nor with respect to its exon–intron structure. It will be of much interest to see further data answering these two questions.

But it is not even clear yet whether the dynein gene has any relationship with the lampbrush loop-forming region of the fertility gene *kl-5*. If it has, it might as well be positioned at the very beginning or at the end of the transcription unit in a structure with few (or no) introns. But is it situated in the loop-forming region at all or is it located outside the lampbrush loop transcription unit? The latter possibility seems incompatible with the finding that only single complementation groups can be assigned to the various fertility genes identified. However, we do have to keep in mind that the absence of markers in the Y chromosomes raises substantial difficulties for analysis of the genetic fine structure of genes on this chromosome and that we may end up with unpredicted discoveries.

Some genetic data also seem incompatible with all current models made for the structure and function of the fertility genes. It has repeatedly been claimed that the Y chromosomal genes have very high mutation rates and that this can be explained only by the presence of physically large genes. However, in an exon–intron model as discussed before, this

would also imply that mutations in introns, which must form a major part in any gene model explaining lampbrush loop sizes by an intron–exon structure, must delete the normal function of this gene. For such an idea, no evidence is available from other genetic systems.

Therefore, alternative interpretations seem more adequate. One of the possibilities is the existence of hot spots for mutations or DNA regions comparable to those responsible for the fragile X syndrome caused by DNA sequence expansion in the human X chromosome.

One might as well argue that not all complementation groups in the Y chromosome have yet been discovered. This could be true if sizes of different Y chromosomal genes are widely different. The large sizes of the transcription units in the lampbrush loops may have shielded small genes from detection in mutagenesis experiments if the DNA in the loop regions display similar mutation rates as genes with, for example, 20 times less DNA. I seriously wonder whether a number of Y chromosomal mutations checked to date would suffice to uncover the small gene at all. Hence, the chromosome region carrying fertility gene *kl-5*, which in most experiments including those of Gepner and Hays is defined by a relatively large deletion, may carry additional complementation groups not yet discovered. I do not want to imply that this seems to be the case but it should be considered as a possibility that cannot be entirely excluded. If it is true, the question of lampbrush loop function in spermatogenesis must be kept apart. If not, the situation becomes even more complex, as it appeared until now where the lampbrush loop models are based on the exclusive finding of repetitive DNA elements within the transcription units.

Why does the detection of a protein coding gene not solve the secrets of the fertility genes if it is located within the lampbrush loop-forming region assigned to *kl-5*? The most obvious reason is that the formation of lampbrush loops is a specific but universal feature of Y chromosomal genes of *Drosophila* in primary spermatocytes, at least as far as it can be derived from the few species studied for a linkage of the lampbrush loops to the Y chromosome. Such lampbrush loops are not found in somatic tissues nor are they formed by genes in other chromosomes. Even more intriguing is the molecular structure of the transcription units as it exists for a large portion of satellite-type DNA sequences, which are interspersed with transposable elements (15, 20). This basic molecular structure is maintained in different species even though the DNA sequences involved are totally different, except in closely related species. Whether the interspersion of the satellite DNA with transposon sequences, which also seems to be a general feature, is of functional importance is entirely unclear but not impossible (25). Our recent data show that certain constraints at the DNA sequence level exist—for example, the transcriptional orientation of the different DNA sequence elements is strictly maintained within but not outside the transcription unit (R. Hockstenbach and W.H., unpublished data).

Since one has not obtained any evidence for the interspersion of exons within the lampbrush loop sequence studied so far, a wide dispersion of a protein-coding gene throughout the transcription unit is rather unlikely. However, the formation of a lampbrush loop must functionally be connected to the functional state of the respective fertility gene, as in all cases in which a lampbrush loop is absent due to a mutation, the respective fertility gene is inactive as well. Moreover, in many cases where lampbrush loops are morphologically modified the fertility gene is not functional (8, 21).

What then may be the function of the lampbrush loops and how does this function relate to the presence of a protein coding gene? Before additional data on the Y chromosomal dynein gene are available, the answer can be given only in the form of speculations. To me it seems likely that any protein-coding DNA sequence, if located within a lampbrush loop, is positioned toward one end of the lampbrush loop-forming DNA region, most likely toward the end of the transcription unit as implied by certain mutations. A more plausible assumption to me is that the dynein gene is located outside the lampbrush loop transcription unit. The lampbrush loop itself must serve for functions other than those connected with the production of a normal pre-mRNA, as has been expressed in my previous view that lampbrush loop formation differs between oogenesis and spermatogenesis (26). Contrary to what occurs in oogenesis, no antigenic determinants of small nuclear ribonucleoproteins or small nuclear RNAs (as detected by *in situ* hybridization or with the antibodies directed against their m3 cap; see ref. 27) can be found in the *Drosophila* lampbrush loops. Hence, no major splicing activities of newly synthesized pre-mRNA are expected to occur in the Y chromosomal loops, as would be required if these genes were composed of a large number of alternating exons and introns. Autosomes in primary spermatocytes, on the contrary, display the expected antibody reactions (unpublished observations). One might argue that splicing is delayed and occurs only after completion of the transcripts. This might be considered a mechanism to introduce a clock for the formation of active mRNAs into the spermatocyte as the completion of lampbrush loop transcripts takes hours, thus delaying expression of the respective gene products. However, such a view is not based on much experimental evidence.

A rather unique feature of the Y chromosomal lampbrush loops is their association with several mostly loop-specific antigens as first shown by Hulsebos *et al.* (28). This has led to the proposal that the transcripts formed in the Y chromosomal lampbrush loops, and possibly some of the landmark loops in amphibians such as the giant granular loops, may serve as a target for protein binding and, for example, for intracellular compartmentalization purposes in the differentiation of the male germ cell (29). Other authors (30) have recently joined our view based on observations made on the protein binding of lampbrush loops in *D. melanogaster*.

There may exist some similarities between the properties of Y chromosomal lampbrush loops to other unusual genes. The heat shock locus *93D* of *D. melanogaster* contains alternating blocks of satellite-like DNA sequences and a section of unique DNA sequences, a molecular structure more simple but in principle rather similar to that of the Y chromosomal lampbrush loops. The evolutionary divergence of the DNA sequences in the equivalent sites of other species is similarly strong, as in the Y chromosome. In this locus, no protein seems to be encoded (see, for example, ref. 31). The Responder locus of *D. melanogaster*, which modulates segregation distortion, also displays an intriguing structural similarity. More recently, the human *XIST* locus and the homologous *Xist* gene in mice has been speculated to serve for functions such as modifying the chromatin constitution of the inactive female X chromosome by means of its transcripts. They also contain repetitive sequence elements but do not code for proteins (32, 33). In all these cases, the function of the active genes may relate to RNA–protein interactions at the site of transcription.

Additional evidence supporting the view of specific RNA–protein interactions has recently been obtained from studies of antisera induced against the proteins of synaptonemal complexes (SCs) (W.H. and Heyting, unpublished observations). Even though *Drosophila* males do not form SCs, as they have no meiotic recombination, strong immunological reactions between several antisera against SC proteins and specific Y chromosomal lampbrush loops are observed in different *Drosophila* species. These immunoreactions are restricted to germ cells in meiotic prophase. The intriguing question that arises is whether the Y lampbrush loops serve as a functional substitute for the SCs normally formed during the meiotic prophase. The

absence of crossing-over and the formation of lampbrush loops in *Drosophila* spermatocytes may be functionally related and may answer the question why such lampbrush loops are found only in *Drosophila* spermatocytes.

Whatever the final outcome of the puzzle of Y chromosomal fertility gene function may be, it must explain why large lampbrush loops, mainly visible due to the large amounts of protein accumulated, are formed and what is the biological function of these loops. The documentation that a protein-coding gene resides on the Y chromosome not only adds an exciting aspect to the already complex picture, but it might finally also help to resolve the properties of this fascinating and unusual genetic system.

I am grateful to Dr. Allan Spradling for critically reading this commentary.

1. Bridges, C. B. (1916) *Genetics* **1,** 1–52 and 107–162.
2. Kennison, J. A. (1981) *Genetics* **98,** 529–548.
3. Gatti, M. & Pimpinelli, S. (1983) *Chromosoma* **88,** 349–373.
4. Brosseau, G. E. J. (1960) *Genetics* **45,** 257–274.
5. Meyer, G. F., Hess, O. & Beermann, W. (1961) *Chromosoma* **12,** 676–716.
6. Hess, O. (1967) *Chromosoma* **21,** 429–445.
7. Hennig, W. & Kremer, H. (1990) *Int. Rev. Cytol.* **123,** 129–176.
8. Hackstein, J. H. P., Leoncini, O., Beck, H., Peelen, S. & Hennig, W. (1982) *Genetics* **101,** 257–277.
9. Hennig, W. (1968) *J. Mol. Biol.* **38,** 227–239.
10. Hennig, W., Meyer, G. F., Hennig, I. & Leoncini, O. (1974) *Cold Spring Harbor Symp. Quant. Biol.* **38,** 673–683.
11. Ingman-Baker, J. & Candido, E. P. M. (1980) *Biochem. Genet.* **18,** 809–828.
12. Hulsebos, T. J. M., Hackstein, J. H. P. & Hennig, W. (1983) *Dev. Biol.* **100,** 238–243.
13. Goldstein, L. S. B., Hardy, R. W. & Lindsley, D. L. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 7405–7409.
14. Bonaccorsi, S., Gatti, M., Pisano, C. & Lohe, A. (1990) *Chromosoma* **99,** 260–266.
15. Hennig, W., Brand, R. C., Hackstein, J. H. P., Hochstenbach, R., Kremer, H., Lankenau, D.-H., Lankenau, S., Miedema, K. & Pötgens, A. (1989) *Genome* **31,** 561–571.
16. Vogt, P., Siegmund, I. & Hennig, W. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 5132–5136.
17. Vogt, P. & Hennig, W. (1986) *Chromosoma* **9,** 449–458.
18. Vogt, P. & Hennig, W. (1986) *Chromosoma* **9,** 459–467.
19. Huijser, P. & Hennig, W. (1987) *Mol. Gen. Genet.* **260,** 441–451.
20. Hochstenbach, R., Pötgens, A., Meyer, H., Dijkhof, R., Knops, M., Schouren, K. & Hennig, W. (1993) *Chromosoma* **102,** 526–545.
21. Leoncini, O. (1978) *Chromosoma* **63,** 329–357.
22. Gepner, J. & Hays, T. S. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 11132–11136.
23. Grond, C., Siegmund, I. & Hennig, W. (1983) *Chromosoma* **88,** 50–56.
24. deLoos, F., Dijkhof, R., Grond, C. J. & Hennig, W. (1984) *EMBO J.* **3,** 2845–2849.
25. Hennig, W. (1987) in *Results and Problems in Cell Differention*, ed. Hennig, W. (Springer, Heidelberg), Vol. 14, pp. 133–146.
26. Hennig, W. (1989) in *Developmental Genetics of Higher Organisms*, ed. Malacinski, G. M. (Macmillan, New York), pp. 239–274.
27. Wu, Z., Murphy, C., Callan, H. G. & Gall, J. G. (1991) *J. Cell Biol.* **113,** 465–483.
28. Hulsebos, T. J. M., Hackstein, J. H. P. & Hennig, W. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 3404–3408.
29. Hennig, W. (1985) *Adv. Genet.* **23,** 179–234.
30. Pisano, C., Bonaccorsi, S. & Gatti, M. (1993) *Genetics* **133,** 569–579.
31. Pardue, M. L., Bendena, W. G. & Garbe, J. C. (1987) in *Results and Problems in Cell Differention*, ed. Hennig, W. (Springer, Heidelberg), Vol. 14, pp. 121–131.
32. Brown, C. J., Hendrich, B. D., Rupert, J. L., Lafrenier, R. G., Xing, Y., Lawrence, J. & Willard, H. F. (1992) *Cell* **71,** 527–542.
33. Brockdorf, N., Ashworth, A., Kay, G. F., McCabe, V. M., Norris, D. P., Cooper, P. J., Swift, S. & Rastan, S. (1992) *Cell* **71,** 515–526.