Evolution of segmentation genes in insects

DIETHARD TAUTZ AND RALF J. SOMMER

Systematic genetic analysis of the segmentation process in Drosophila has established a paradigm for the molecular control of the formation of metameric segments. However, it has been suggested that some of the mechanisms involved in this process in Drosophila are uniquely adapted to the syncytial mode of embryogenesis in such higher dipterans. A particularly contentious problem is the role of early segmentation genes in short-germ insects, in which development proceeds by sequential addition of segments in a cellular environment. However, analysis of the expression of presumptive homologs of segmentation genes in holometabolous short-germ insects suggests that they do indeed have a role in segmentation and that the Drosophila paradigm may be more widely applicable than is usually assumed. Most interestingly, these results suggest that the molecular mechanisms of pattern formation in noncellular and in cellular environments may not be as radically different as it is often thought to be.

The choice of *Drosophila* as an experimental system for studying early development[1] was entirely governed by the unique suitability of this organism for genetic analysis. Up to that point, *Drosophila* was not a classical system for studying insect embryogenesis. Instead, most experiments were performed with species that show the more ancestral mode of embryogenesis, namely short-germ development[2]. The major difference between short- and long-germ development concerns the timing of segmented pattern formation: in short-germ development, the more posterior segments are formed only after completion of the blastoderm stage, whereas in the long-germ mode, the segmented pattern is already completely established by the blastoderm stage* (Fig. 1). This is an important difference, because the first nuclear cleavages in insect development proceed without concomitant cell divisions and early embryonic development thus occurs in a syncytium[3]. This allows large proteins, such as transcription factors, to diffuse relatively freely and form gradients in the early embryo. In *Drosophila*, these transcription factors have concentration-dependent regulatory effects and act as morphogens[4–6].

The cascade of early patterning decisions is started by the maternal gene products, which are asymmetrically localized in the egg or become locally activated[4]. These maternal gene products delimit the expression domains of the gap genes, which encode transcription factors and themselves form short-range gradients[5]. These gradients in turn instruct the expression domains of the primary pair-rule genes, which are responsible for initiating the metameric segmentation process[6]. The final segmental borders are then set by interactions among the segment polarity genes[7]. These final stages of segmental subdivision occur at the time when the embryo becomes cellularized and several of the segment polarity genes appear to be involved in cell–cell signalling[7]. However, the basic outline of the body pattern is determined at the syncytial stage and is governed by gradients of transcription factors. It has therefore been argued that early *Drosophila* development is a highly specialized form of development that cannot serve as a paradigm for most other forms of embryogenesis[8].

Finding homologs of segmentation genes

The early expressed segmentation genes, the gap genes and the pair-rule genes, encode transcription factors. Most of their protein-coding regions evolve rather fast and are fairly well-diverged in comparisons between species that have evolved separately for 100 million years (Myr) or more. Nonetheless, they also contain blocks of highly conserved sequences; these sequences usually encode the DNA-binding domains of the transcription factors. This allows potential homologs of these genes to be cloned from various organisms by PCR or low-stringency hybridization. However, functional homology cannot be established on the basis of

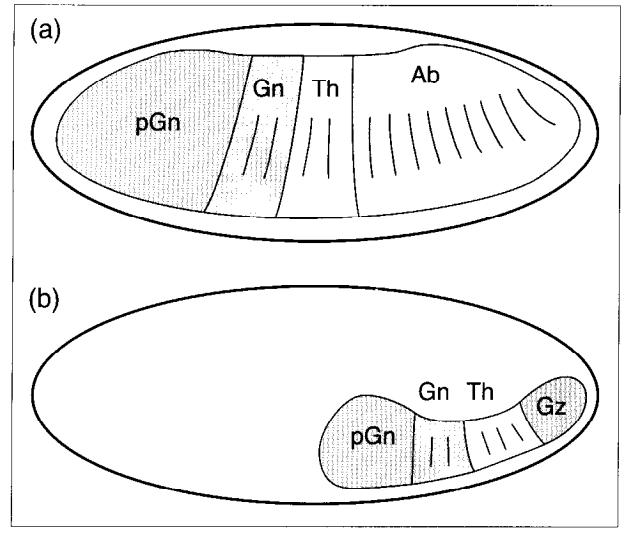

FIGURE 1. Schematic drawing showing the main differences between development of (a) long-germ and (b) short-germ embryos. Depicted is a schematic fate map shortly after the formation of the early germ-band. In long-germ embryos, all segments are already determined at this stage, while in short-germ embryos, the posterior segments are formed secondarily from a growth zone. Note that the total number of posterior segments that are specified early in development varies in different species[11,17]. pGn, pre-gnathal segments; Gn, gnathal segments; Th, thoracic segments; Ab, abdominal segments; Gz, growth zone.

*The terms long germ and short germ were introduced by Krause[11], who also used the term semi-long germ for intermediate forms. Originally, these terms were meant to be purely descriptive and at this level there is a continuum of all three forms in various species. However, we now use the terms short germ and long germ in a more functional sense to indicate whether there is a secondary growth process[16]. In our usage, the term semi-long or intermediate has no particular relevance: all embryos that show some form of secondary growth are called short germ. This definition is used herein. For a more extensive discussion of this problem, see Ref. 24.

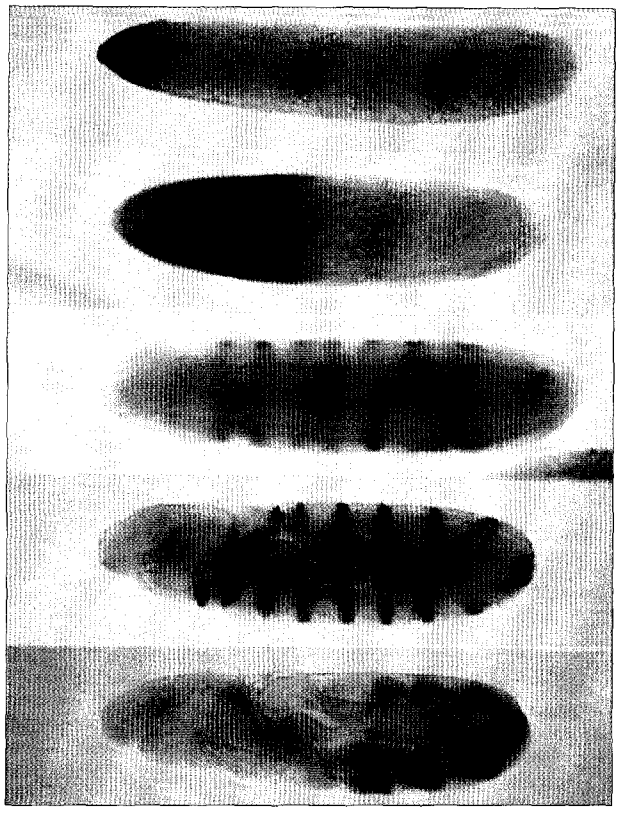

FIGURE 2. Hierarchy of segmentation gene function in the house-fly, *Musca domestica*. The expression pattern of a representative gene from each level of the hierarchy is shown. From top to bottom: *bicoid* (maternal gene), *hunchback* (gap gene), *hairy* (pair-rule gene), *engrailed* (segment polarity gene; note that the segmented part of the embryo has a U-shaped form at these later stages) and *Ultrabithorax* (homeotic gene). The expression pattern of all these genes is basically the same as in *Drosophila* at the equivalent stages. All patterns were revealed by whole-mount *in situ* hybridization using *Musca* specific probes[10], apart from the *hairy* pattern, where an antibody against the *Drosophila hairy* gene product was used.

sequence similarity alone, as genes may lose or gain some functions during evolution or may become duplicated and diversified in function. An additional criterion is therefore required to infer that the genes play the same role in development. In the context of pattern formation, this criterion could be that the genes not only share sequence similarity, but also occupy the same place in a regulatory network. In the case of segmentation genes, this inference can be derived from a comparative analysis of expression patterns. A certain pattern may be called homologous between species if it shows the same spatiotemporal profile in both. This approach, which can be called 'molecular comparative embryology', is similar to the classical approaches of inferring homology between morphological characters in different species. It therefore has the same pitfalls (in that the argument used is necessarily somewhat circular), but also shares its conceptual strengths[9]. Furthermore, in the context of our background knowledge of gene function in *Drosophila*, we can for the first time link embryological comparisons with comparisons of the genetic machinery that underlies particular traits.

Long-germ insects

As a first step in the comparative approach, we have analysed the expression of segmentation genes in the house-fly, *Musca domestica*. The house-fly has the same long-germ mode of embryogenesis as *Drosophila*, but the two are evolutionarily separated by more than 100 Myr. We have cloned fragments of several genes from *M. domestica* by PCR and studied their expression pattern[10]. We found that the primary segmentation gene hierarchy is indeed very similar in *Drosophila* and *Musca* (Fig. 2), although a number of aspects of secondary gene expression are different[10]. This demonstrates that at least the first steps in the segmentation hierarchy follow conserved molecular pathways.

A second insect that was studied in this way is the moth *Manduca sexta* (Lepidoptera). Shortly after the blastoderm stage, Lepidopteran embryos form a structure resembling that of an early short-germ embryo; they were therefore originally described as having the characteristics of both long- and short-germ insects[11]. However, cytological reanalysis[12] and blastoderm fate-mapping experiments in a related species[13] suggested that the early embryogenesis of Lepidoptera is more akin to that of long-germ embryos. To analyse this problem at the molecular level, Kraft and Jäckle have cloned several fragments of gap, pair-rule and segment polarity genes and studied their expression in *Manduca*[14]. They found that the expression patterns of *hunchback* (*hb*) and *Krüppel* (*Kr*) at the blastoderm stage look very similar to those in *Drosophila*, with *hb* being expressed in an anterior and a posterior domain and *Kr* in a more central domain. In addition, the pair-rule gene *runt* was found to form eight stripes in the early germ band. These results suggest that the whole segmental pattern becomes specified at the molecular level at the blastoderm stage, very much as in *Drosophila*. *Manduca* is thus clearly a long-germ insect, even though the morphological processes in the early embryo look rather different from those seen in *Drosophila* and *Musca*. This study shows that molecular segmentation pathways can be conserved, even when the overt morphological pathways differ.

Holometabolous short-germ insects

The flour beetle *Tribolium castaneum* is a holometabolous insect like *Drosophila* and *Manduca*. However, it is evolutionarily separated from *Manduca* by at least 250 Myr and is a typical representative of a short-germ insect[15] in which only part of the segment pattern is specified at the blastoderm stage (Fig. 1). Analysis of the expression patterns of homologs of gap, pair-rule and segment polarity genes supports this inference. Although these patterns are distinctly different from those in *Drosophila*, they can nonetheless be reconciled with the assumption that the process as a whole is conserved. In *Drosophila*, the gap gene *Kr* is expressed in a central domain at the blastoderm stage, while in *Tribolium* it is expressed in a posterior cap[16] (Fig. 3). This result suggests two interpretations. One is that *Kr* does not have a homologous function in *Tribolium*. The second possibility, which is supported by comparative analysis of the expression patterns of pair-rule genes, is that the expression pattern reflects

the fact that segments posterior to the *Kr* domain are not yet specified at the blastoderm stage.

The pair-rule genes *hairy* (*h*) and *even skipped* (*eve*) are each expressed in seven stripes in the *Drosophila* blastoderm. Their expression pattern in *Tribolium* is different: they are expressed in only two stripes and in a posterior cap[16,17]. In *Drosophila*, the first two stripes lie anterior to the *Kr* domain, while the third and the fourth lie within the *Kr* domain. In *Tribolium*, the two stripes of expression in the blastoderm are also anterior to the *Kr* domain, while the posterior cap overlaps the *Kr* domain and can be interpreted as a fusion of *Drosophila* stripes three and four[16]. Thus, these results clearly support the morphological inference that only the anterior segment pattern becomes specified in the short-germ embryo at the blastoderm stage.

Intriguingly, the pair-rule genes seem to play a role in the segmentation of the embryo even after the blastoderm stage, during secondary growth of the abdominal segments. Both *h* and *eve* are expressed in stripes in the growth zone of the elongating embryo[16,17]. It has been shown that these stripes of *eve* expression represent a double segmental pair-rule pattern, since they split into a segmental pattern at later stages[17]. A third pair-rule gene, *runt*, has a similar expression pattern (S. Brown, pers. commun.). Thus, all three primary pair-rule genes, which are thought to respond directly to the positional cues of the gap genes[6] in *Drosophila*, appear to be used in the secondary segmentation process of the *Tribolium* embryo, even though this process occurs in a cellular, rather than syncytial, environment (see below).

Genes from the next level of the segmentation gene hierarchy, the segment polarity genes *engrailed* (*en*) and *wingless* (*wg*), have very similar expression patterns in *Tribolium* and *Drosophila*[18,19]. In *Drosophila*, these genes are expressed in adjacent rows of cells and it is clear that interplay between *wg* and *en* is crucial for the formation of the segmental borders[7]. In *Tribolium*, the two genes are also segmentally expressed in neighbouring domains, suggesting that their functional interaction is conserved[19] (Fig. 3).

Not only does the process of anterior–posterior pattern formation seem to be highly conserved between *Drosophila* and *Tribolium*, dorsal–ventral pattern formation also seems to be conserved. Among the first zygotic target genes that respond to maternal dorsoventral cues in *Drosophila* are *snail* (*sna*) and *twist*[4] (*twi*). Both are expressed in the prospective mesodermal region of the embryo and are required for formation of the mesoderm. *Tribolium* homologues of *sna* and *twi* have comparable expression patterns to those of the *Drosophila* genes. Both genes are expressed in similar domains at the ventral side of the blastoderm embryo, in the region of the prospective mesoderm. Furthermore, their expression persists in the developing mesodermal region of the growing germ-band until segmentation is complete[20].

In contrast to the anterior–posterior and dorsal–ventral pattern forming systems, genes of the terminal system have not yet been identified in *Tribolium*. In *Drosophila*, the terminal system determines the acron and the telson and spatially delimits expression of the segmentation genes[5]; for example, *Kr* is negatively regulated by the terminal system from both ends of the embryo at the blastoderm stage. Clearly, the situation in *Tribolium* must be different, as *Kr* is expressed at the posterior terminus and there is no obvious indication of repression. On the other hand, the expression pattern of *wg* hints that there may be a terminal activity in *Tribolium*, even at the blastoderm stage. In *Drosophila*, *wg* has a domain of secondary expression at the posterior terminus that is regulated by the terminal system. This secondary domain is also seen in the blastoderm-stage *Tribolium* embryo, suggesting that an early terminal activity is indeed present[19] (Fig. 3). If there is such an activity, its functions must be uncoupled from the primary segmentation process. Perhaps

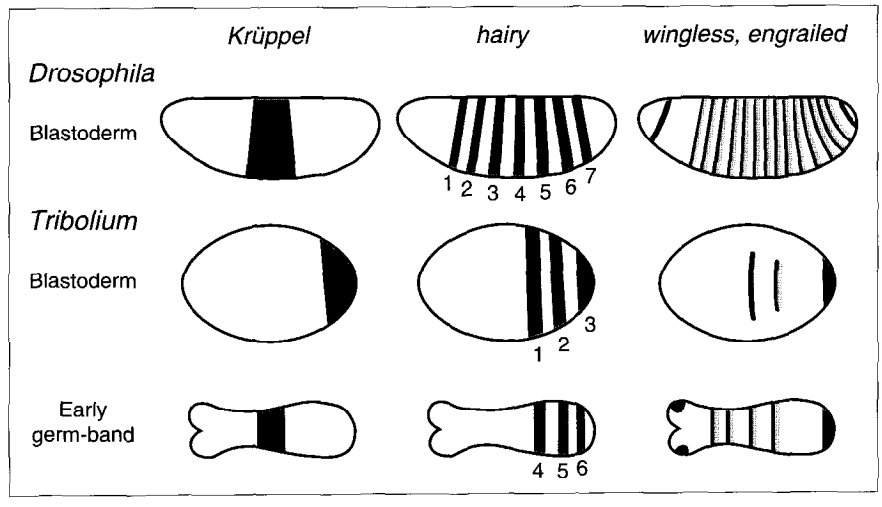

FIGURE 3. Comparison of the expression patterns of three segmentation genes in *Drosophila* and in the flour beetle *Tribolium*. At the blastoderm stage, *Krüppel* (*Kr*) is expressed in a central domain in *Drosophila* and in a posterior domain in *Tribolium*. However, in the early germ-band, expression is more central. At the blastoderm stage, hairy (*h*) is expressed in *Drosophila* in seven striped domains and in *Tribolium* in three stripes. In the early germ-band, the more anterior stripes (shown as light green) can no longer be detected, at least at the level of the transcript[16]. Accordingly, the stripes seen represent stripes 4–6 (darker green). *wingless* (*wg*, red) and *engrailed* (*en*, yellow) are expressed next to each other in fourteen stripes in the segmented region of the *Drosophila* embryo. In addition, *wg* is expressed in an anterior and a posterior domain (purple) that is not involved in primary segment formation, but marks the head region and the posterior terminal region (note that the *Drosophila* picture is highly schematic, since the terminal *wg* stripes are already present before the segmental stripes form and since the full pattern of fourteen stripes is seen only after the beginning of gastrulation). The head and the posterior domain of *wg* expression are also seen in *Tribolium* at an early stage, while the segmental stripes develop progressively. Note that the striped expression of the pair-rule genes in the growth zone of the *Tribolium* embryo precedes the formation of the stripes of segment polarity gene expression.

a particular group of cells receives a maternally derived terminal signal and keeps on transmitting this positional information until the growth process is completed. These cells might be identical to those that form the hindgut anlage (proctodaeum), which can already be identified at the blastoderm stage in short-germ embryos[21].

Hemimetabolous short-germ insects

Among the hemimetabolous insects, only in the grasshopper *Schistocerca* has the expression of segmentation genes been studied in detail. Interestingly, the pair-rule genes studied, *eve* and a potential homolog of *fushi tarazu* (*ftz*), seem not to be involved in the segmentation process: only later in development, during formation of the central nervous system, are their expression patterns homologous to those in *Drosophila*[22,23]. It was therefore suggested that the segmentation process is radically different in hemimetabolous short-germ insects. However, we favor an alternative explanation: that *Schistocerca* represents a special case. In this insect, the early germ-band is extremely short and consists only of the head lobes and a growth zone. However, this is not typical of other hemimetabolous insects, in particular, more ancestral ones; for example, embryonic development of the damsel-fly *Platycnemis* is much more reminiscent of that seen in *Tribolium*[2,21]. We therefore argue that the form of embryogenesis seen in *Schistocerca* has evolved secondarily, and may not be representative of the ancestral state of insect embryogenesis[21,24]. Clearly, further comparative studies of other hemimetabolous insects are necessary to resolve this problem. However, at present, results obtained in studies of *Schistocerca* should not be interpreted as conclusively refuting the possibility that there is general conservation of the segmentation gene pathways in insects.

Other organisms

Expression of segmentation genes has so far been studied in relatively few other organisms. However, one useful study used a monoclonal antibody against the *en* gene product that cross-reacts with a conserved epitope in homologs throughout the animal kingdom[25]. The results showed that *en* appears to be involved in determining segmental boundaries not only in insects, but also in crustaceans[25]. Another gene of the *engrailed* class is expressed in a segmental pattern in an organism from a different phylum, the leech[26], suggesting that the involvement of *engrailed*-like genes in the segmentation process is an ancestral characteristic. On the other hand, the *en* homologues so far recovered from vertebrates do not show a segmental expression pattern. Again, this is not conclusive evidence that the segmentation hierarchy as a whole does not play a role in vertebrate embryogenesis. At the very least, the structure and expression pattern of the homeotic genes, which in *Drosophila* are regulated by the segmentation genes, are known to be very well conserved in the various phyla[27].

Cell walls and gradients of transcription factors

In *Tribolium*, there appears to be a relatively high degree of conservation of the segmentation gene hierarchy, even against the background of a rather different type of early embryonic development. The most surprising finding in this context probably concerns the role of pair-rule genes, which appears to violate one of the current paradigms of *Drosophila* development, that the striped expression of the primary pair-rule genes is directly controlled by the gradients of transcription factors encoded by the gap genes. In *Drosophila*, it is believed that these gradients are generated by diffusion of molecules in the syncitial blastoderm[5,6]. This is obviously not possible in *Tribolium*, as by the time elongation of the germ-band occurs, closed cells appear to have formed[15]. How then are gradients of the transcription factors set up in short-germ embryos? There are several possibilities. First, in these embryos, pathways of cell–cell signalling might be involved even at the level of the gap and pair-rule genes. We consider this unlikely, since it seems difficult to conceive that such pathways could have been lost in the evolutionary line to the higher insects and that the same transcription factors would still be involved. Second, gradients might not be generated by diffusion, but by serial dilution as the result of progressive cell division. While this is an attractive hypothesis, the visible pattern of cell division does not directly support it. Cell divisions do not occur strictly at the posterior tip of the germ band, but irregularly throughout the whole growth zone (D. Tautz and R.J. Sommer, unpublished), so there is no obvious correlation between cell division and pattern forming processes in this region. Third, transcription factors might cross cell walls to form diffusion-controlled gradients, even in short-germ insects. While it seems unlikely that this could occur via the normal pathway of protein secretion, as this would require the presence of both signal peptides in the proteins and fairly specific receptors in the neighbouring cell, it is interesting to note that *in vitro* studies of at least one transcription factor, the product of the homeotic gene *Antennapedia*, have shown that it can enter cells by a novel pathway[28]. Finally, another possibility might be that direct connections, similar to gap junctions, exist between cells in the growth zone. Gap junctions can directly couple the interior of neighbouring cells, but vertebrate gap junctions allow only the passage of small molecules. While insects have different types of gap junctions to those in vertebrates[29], even these are probably too small to permit passage of proteins as large as transcription factors. Nonetheless, it should be interesting to determine whether transcription factors can be directly exchanged between neighbouring cells; if such a mode of cell–cell communication did exist, it might also be important in understanding pattern formation processes in other systems.

Conclusion

Although systematic analysis of the homologs of segmentation genes in other organisms is at an early stage, it is already clear that the genetic principles established in *Drosophila* may be much more widely applicable than has so far been assumed. However, the current picture is still incomplete and must be enhanced by further comparative studies. In particular, it is still not clear how far the maternal systems[4] are conserved and to what extent the regulation of the

homeotic genes follows conserved pathways. Answering these questions will provide major clues as to the evolution of pattern forming processes and should therefore prove a fruitful task.

Acknowledgements

We thank M. Akam, S. Brown, B. Kraft and N. Patel for providing manuscripts and information before publication, and M. Klingler and our colleagues for discussions and helpful suggestions on the manuscript.

References

1 Nüsslein-Volhard, C. and Wieschaus, E. (1980) *Nature* 287, 795–801
2 Sander, K. (1976) *Adv. Insect Physiol.* 12, 125–238
3 Foe, V.E. and Alberts, B.M. (1983) *J. Cell Sci.* 61, 31–70
4 St Johnston, D. and Nüsslein-Volhard, C. (1992) *Cell* 68, 201–219
5 Hülskamp, M. and Tautz, D. (1991) *BioEssays* 13, 261–268
6 Pankratz, M.J. and Jäckle, H. (1990) in *The Development of Drosophila melanogaster* (Bate, M. and Martinez Arias, A., eds), pp. 467–516, Cold Spring Harbor Laboratory Press
7 Ingham, P.W. and Martinez Arias, A. (1992) *Cell* 68, 221–235
8 Davidson, E.H. (1991) *Development* 113, 1–26
9 Patterson, C. (1988) *Mol. Biol. Evol.* 5, 603–625
10 Sommer, R.J. and Tautz, D. (1991) *Development* 113, 419–430
11 Krause, G. (1939) *Biol. Zbl.* 59, 495–536
12 Broadie, K.S., Bate, M. and Tublitz, N.J. (1991) *Roux's Arch. Dev. Biol.* 199, 327–334
13 Lüscher, M. (1944) *Revue Suisse Zool.* 51, 531–627
14 Kraft, R. and Jäckle, H. (1994) *Proc. Natl Acad. Sci. USA* 91, 6634–6638
15 Stanley, M.S.M. and Grundmann, A.W. (1970) *Ann. Ent. Soc. Am.* 63, 1248–1256
16 Sommer, R.J. and Tautz, D. (1993) *Nature* 361, 448–450
17 Patel, N.H., Condron, B.G. and Zinn, K. (1994) *Nature* 367, 429–434
18 Brown, S.J., Patel, N.H. and Denell, R.E. (1994) *Dev. Genet.* 15, 7–18
19 Nagy, L.M. and Carroll, S. (1994) *Nature* 367, 460–463
20 Sommer, R.J. and Tautz, D. (1994) *Dev. Genet.* 15, 32–37
21 Anderson, D.T. (1972) in *Developmental Sytstems: Insects* (Vol. 1) (Counce, S.J. and Waddington, C.H., eds), pp. 96–163, Academic Press
22 Patel, N.H., Ball, E.E. and Goodman, C.S. (1992) *Nature* 357, 339–342
23 Dawes, R., Dawson, I., Tear, G. and Akam, M. (1994) *Development* 120, 1561–1572
24 Tautz, D., Friedrich, M. and Schröder, R. *Development* (Suppl.) (in press)
25 Patel, N.H. *et al.* (1989) *Cell* 58, 955–968
26 Lans, D., Wedeen, C.J. and Weisblat, D.A. (1993) *Development* 117, 857–871
27 McGinnis, W. and Krumlauf, R. (1992) *Cell* 68, 283–302
28 Le Roux, I. *et al.* (1993) *Proc. Natl Acad. Sci. USA* 90, 9120–9124
29 Caveney, S. (1991) in *Physiology of the Insect Epidermis* (Binnington, K. and Retnakaran, A., eds), pp. 23–35, CSIRO Press

D. TAUTZ IS IN THE ZOOLOGISCHES INSTITUT DER UNIVERSITÄT MÜNCHEN, LUISENSTRASSE 14, 80333 MÜNCHEN, GERMANY, AND R.J. SOMMER IS IN THE CALIFORNIA INSTITUTE OF TECHNOLOGY, DIVISION OF BIOLOGY 156–29, PASADENA, CA 91125, USA.

The genetics of meiosis in *Caenorhabditis elegans*

MONIQUE ZETKA AND ANN ROSE

The many features that have made the hermaphroditic nematode Caenorhabditis elegans a good model system for studying development have also attracted investigators to the study of meiosis. Genetic analysis suggests that in C. elegans there are two types of chromosomal sites required for proper meiotic function. The first is needed early in meiosis for recombination and segregation. The second is involved in the mechanisms that establish the normal frequency and distribution of exchange. Genes whose products may interact with these sites have been identified by mutant analysis. Study of these mutations in the nematode is enhancing our general understanding of meiotic functions.

Meiosis in *C. elegans* is fundamentally the same highly conserved process that has been described in other sexually reproducing species. *C. elegans* chromosomes are described as holocentric, as they have a diffuse kinetochore during mitosis. However during meiosis there is a single chromosomal site of attachment to the spindle, and the behaviour of *C. elegans* chromosomes is therefore analogous to that of monocentric chromosomes. The earliest stages of meiosis involve the recognition and alignment of homologous pairs, and this is followed by recombination and chiasma formation. During metaphase I, the chromosomes align axially (an end-to-end association). Kinetic activity is limited to one end of the bivalent, although either end of the bivalent has the potential to assume this function. No kinetochore structure is evident during meiosis and the microtubules project directly into the chromatin[1]. As in other organisms whose chromosomes contain nonlocalized centromeres, the activity and structure of the centromeres differs between meiosis and mitosis[2].

The events involved in recognition and alignment of chromosome pairs during prophase I have traditionally been described as chromosome pairing. However, the term pairing has no precise definition in the literature of meiosis and has been used to describe a wide range of meiotic events. Many pairing events are now understood to be temporally and functionally distinct; these include homolog recognition, alignment for recombination, and synapsis. In this article, we define homolog recognition as the initial lining up of chromosome pairs, alignment for recombination as the matching of homologous sequences that is necessary for recombination to occur, and synapsis as the intimate