Identification of novel genes in Drosophila reveals the complex regulation of early gene activity in the mesoderm

(gastrulation/twist/snail)

José Casal* and Maria Leptin†

Max-Planck-Institut für Entwicklungsbiologie, Speemannstrasse 35, D-72076 Tübingen, Germany

Communicated by A. García-Bellido, Universidad Autónoma de Madrid, Madrid, Spain, May 13, 1996 (received for review February 29, 1996)

ABSTRACT Two zygotic genes, twist and snail, are indispensable for the correct establishment of the mesoderm primordium in the early Drosophila embryo. They are also needed for morphogenesis and differentiation of the mesoderm. Both genes code for transcription factors with different, albeit complementary, functions. Therefore, to understand the early development of the mesoderm, it will be necessary to identify and study the genes regulated by twist and snail. We have searched for downstream genes using a subtractive cDNA library enriched in sequences expressed in the mesoderm. We have isolated sequences that correspond to 13 novel early mesoderm genes. These novel genes show a variety of expression patterns and also differ in their dependence on twist and snail functions. This indicates that the regulation of early gene activity in the mesoderm is more complex than previously thought.

A cascade of events involving maternal and zygotic gene products is responsible for establishing the primordium of the mesoderm in the ventral part of the Drosophila blastoderm (for review, see ref. 1). Two genes, twist and snail, are activated in the most ventral nuclei of the syncytial blastoderm; these nuclei contain the highest concentration of the maternal transcription factor dorsal. Both genes code for transcription factors; twist is a member of the basic helix–loop–helix family and snail contains zinc fingers (2, 3). Gastrulation in embryos mutant for either twist or snail does not take place properly and the mesoderm does not form (4–6). However, the regulatory roles of these two genes are different. Whereas twist is necessary for the activation of genes expressed in the mesoderm, the main function of snail is to repress in the mesoderm anlage genes that are otherwise expressed in the adjacent ectodermal regions (7–9).

The combined action of twist and snail is thought to activate a specific subset of genes whose functions are to carry out the early morphogenesis of the mesodermal germ layer and the regional specification events that lead to the formation of different mesoderm derivatives. Several genes regulated by twist and snail have already been shown to have crucial roles during these processes. However, most of these genes were identified by sequence similarities to other known gene products and therefore they do not constitute a representative sample of gene activities (10–13). We have started a systematic search for genes acting during early mesoderm development and have constructed a subtractive cDNA library in which cDNAs corresponding to mesoderm-expressed genes are enriched. The only assumption this strategy makes is that many genes needed for the development of a particular tissue are expressed specifically in that tissue. A severe limitation of this strategy is often the difficulty of obtaining enough amount of starting material—tissues that express the genes of interest versus tissues that do not. Because of the variety of mutant phenotypes shown by different maternal genes in Drosophila, it is feasible to obtain homogeneous populations of embryos that lack particular cell fates or in which all cells show the same fate. Thus, we made a subtractive cDNA library by using as a source the mRNA embryos in which all cells show mesodermal characteristics and removing nonspecific messages by hybridization to cDNAs obtained from embryos in which the mesoderm is missing. From this library, we isolated cDNAs that correspond to novel early mesodermal genes. Although all these novel genes are expressed in the mesoderm primordium, none of them show expression patterns that precisely coincide with those of twist or snail. Moreover, when their position in the hierarchy of regulatory genes was examined, five of these novel genes showed an unexpected result: they depart from the canonical situation in which twist, but not snail, is necessary for gene expression in the mesoderm anlage. Thus, the regulation of early gene expression in the primordium of the mesoderm appears more complicated than was previously thought.

MATERIALS AND METHODS

Flies. Stocks of mutant flies were obtained from the Tübingen collection. The snailHG cn bw sp chromosome carries a strong snail allele (4–6) and Df(2R)S60, b pr cn bw carries a deficiency for the twist gene (4, 14). A recombinant between these two chromosomes (15) was used as a double twist snail mutant chromosome and maintained as the single mutant chromosomes in a balanced stock over SM6B, P[γ+1, eve–lacZ8.0] (a gift from S. Panzer, Yale University, New Haven, CT). For the construction of the subtractive library, we used the following Toll stocks; T(1;3)OR60/mwh e T106/TM3, Sb Ser (16), e TPTRE ca/TM3, Sb Ser (17, 18), and st T644/TM3, Sb Ser (17–19). As wild type, we used white flies.

cDNA Subtractive Library. We used the method described by Wang and Brown (ref. 20; see Fig. 1A for a scheme of the strategy). Ventralized and lateralized embryos were collected at stages 5 to 9 of development (21) and frozen in liquid nitrogen. poly(A)+ mRNA purification and double-strand cDNA synthesis were done according to Sambrook et al. (22). After digesting the cDNAs with Alul and Rsal, adapters containing an EcoRI site were attached to their ends and the resulting cDNA fragments were amplified using polymerase chain reaction (PCR). A portion of each amplified cDNA, which was digested with EcoRI and biotinylated, was hybridized with a tracer amount of the complementary cDNA (20). The biotin-containing molecules were then removed from the mixture by incubation with streptavidine and phenol extraction, and the remaining cDNAs were hybridized again with another aliquot of the complementary biotinylated cDNAs.

*Present address: Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, England.
†Present address: Institut für Genetik, Universität zu Köln, Weyertal 121, D-50931 Köln, Germany.

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.
The resultant enriched cDNAs were reamplified by PCR and submitted to further cycles of enrichment. The enriched ventral cDNA fragments were digested with EcoRI and ligated to a plasmid vector for transformation of bacteria. Colonies were transferred to nylon filters and differentially screened using ventral- and dorsolateralized-enriched cDNAs as probes.

Longer cDNA and genomic clones were obtained by screening a 4–8 hr embryonic cDNA library (23) and a cosmid library (24) with positive clones from the subtractive library.

**In Situ Hybridization.** In situ hybridization to wild-type and mutant embryos was carried out essentially as described (25). To distinguish between homozygous mutant embryos and their heterozygous siblings, a probe for the lacZ gene was included in the hybridization solution. Genome mapping was carried out by standard in situ hybridization to polytene chromosomes (26). The positions of some of the genes were confirmed by in situ hybridization to embryos of deficiency stocks (gene A, 67D-E; gene C, 30C-D; gene D, 58D-E). The others have not been confirmed and are therefore not included here. Their tentative locations, however, can be requested from the authors.

**Sequencing.** Sequence data were obtained using the Sequenase II kit (United States Biochemical) according to the manufacturer’s instructions. Sequence comparisons were carried out using the default parameters at the BLAST (National Center for Biotechnology Information, Bethesda) and BLITZ (European Molecular Biology Laboratory Data Library and Computer Group, Heidelberg) electronic mail servers.

**RESULTS AND DISCUSSION**

**The Subtractive cDNA Library.** All the embryos laid by $T^{10b}$ females show a ventralized phenotype (16). All cells in these embryos show mesodermal characteristics such as the expression of the twist gene (refs. 27 and 28; Fig. 1A). However, the mRNA prepared from these embryos (ventral mRNA) contain, in addition to mesodermal RNAs, nonspecific sequences from “housekeeping” genes. We removed these nonspecific sequences by hybridization with messages obtained from embryos laid by $TPQRE/Tr^{444}$ females (dorsolateral mRNA). These embryos lack mesodermal cell fates (17), and therefore none of their nuclei express twist (ref. 29; Fig. 1A).

We carried out several cycles of subtraction consisting of series of PCR amplifications followed by cross-hybridizations between cDNAs synthesized from messages obtained from the two types of embryos (see Materials and Methods and ref. 20). We controlled the extent of enrichment by examining the relative amount of sequences from the genes twist, decapentaplegic (a gene expressed dorsally in the early embryo; ref. 30), and $\beta$Tub56D (an uniformly expressed tubulin isoform; refs. 31 and 32) in the amplified cDNAs after each cycle of enrichment. As shown in Fig. 1B, twist is enriched in the ventral cDNAs, while after six rounds of enrichment the tubulin sequences are no longer detected. Similarly, decapentaplegic sequences are enriched in the case of the dorsolateral cDNAs while the common tubulin sequences are removed (Fig. 1C).

After six cycles of enrichment, an aliquot of the ventral cDNAs was inserted into a plasmid vector and used to trans-
form bacterial cells. Colonies (n = 5000) were plated at low density and transferred to nylon filters. A differential screening was then carried out using the enriched ventral and dorsolateral cDNAs as probes; 234 colonies gave a detectable signal only with the ventral cDNA probe. Nonradioactive probes were made from these colonies and their pattern of expression was examined by in situ hybridization to wild-type embryos. Some of the positive clones (n = 48) showed a distinguishable pattern of expression and were then hybridized to dot blots containing the original 234 clones. In situ hybridization with probes from the remaining clones did not produce any signal or only a very weak uniform background. The identity of the positive clones was further confirmed by hybridization to dot blots containing DNA of genomic and long cDNA clones. A total of 16 different classes was thus obtained and a representative clone of each group was chosen for sequencing (Table 1). Three groups showed sequences that were already present in the data bases. They correspond to three known Drosophila mesodermal genes: tinman (36), DFR1 (33), and zfh–1 (34, 35). These three genes are all expressed in the primordium of the mesoderm at blastoderm stage and later in different mesoderm derivatives. We did not find in our screening sequences of other genes such as twist or snail known to be expressed early in the mesoderm, although it is clear that they have been enriched during the subtraction procedure (Fig. 1C). In addition, most of the genes we have found were represented by a single clone (Table 1). Both facts together indicate that the screen is not saturated and, therefore, we cannot estimate at present the number of different genes represented in the library.

Novel Mesodermal Genes. The other 13 groups had sequences not contained in the data bases and, therefore, they correspond to novel genes. As expected from the subtraction strategy, all 13 genes are expressed in ventral cells of the blastoderm and/or in cells of ventral origin during later stages. Four genes are expressed in mesodermal tissues, but their earliest expression start at the late embryonic stages 9 to 12 (data not shown). The other nine genes are first expressed during the blastoderm stage (Fig. 2). However, their early expression patterns differ from each other and from those of twist and snail.

Early anterior and posterior borders. In the blastoderm two genes, B and G, have an anterior border of expression that coincides with the anterior limit of the ventral furrow (Fig. 2). The expression of the other genes extends up to the anterior pole of the embryo. Four genes (B–D and F) show a posterior border of expression that, as in the case of snail (37), coincides with the posterior limit of the ventral furrow. It has been shown that the expression of huckebein in the posterior pole of the embryo delimits the expression of snail as well as the ventral furrow (37). These four genes are regulated in a similar fashion by huckebein to establish the posterior limit of their expression domains in the blastoderm (data not shown). Gene A begins to be expressed during nuclear cycle 12 and its posterior limit of expression lies anterior to the posterior border of the ventral furrow. In this case, tailless, a gene that is expressed in a posterior domain broader than huckebein (38), is responsible for this limit (data not shown). Two genes (G and I) are strongly expressed in the part of the ventral furrow that lies anterior to the cephalic furrow. The expression of gene G posterior to this area, in the trunk region, is very weak and seems to reach the posterior border of the ventral furrow. Genes E, H, and I are expressed beyond this border and, as in the case of twist (3, 8, 37), their domains of expression reach the posterior pole of the embryo.

Expression during gastrulation and later stages. During invagination of the mesoderm and germ-band extension, the nine genes are expressed in patterns that evolve directly from their early expression domains (third column in Fig. 2), although with some variations. During germ-band extension, genes E and F show high expression levels at the head, whereas their expression in the invaginated mesoderm is much weaker. Further differences arise at the extended germ-band stage (fourth column in Fig. 2). Four genes are expressed in other germ layers. The mesoderm expression of gene B is reduced while a new expression in cells of the nervous system can be observed. Gene C shows expression in tracheal pits. Gene E is expressed in the mesectoderm and its early expression in yolk vitellophages (see above) becomes restricted to the periphery of the yolk, often along internal ridges (Fig. 2). At the mid blastoderm stage, another gene, I, is also expressed in yolk cells. During late stages of germ-band extension, a new expression of gene F appears in the mesoderm in a segmental repeat fashion. During later stages of development, the differences in expression become even more pronounced. Two genes, A and D, are not expressed after the retraction of the germ band. Gene B is strongly expressed in the ventral cord, while gene C continues to be expressed in the tracheal system and in parts of the visceral mesoderm. Three genes, F, H, and I, are expressed in the fat body, and gene G is expressed in migrating hemocytes. Both types of tissue are of mesodermal origin (39). twist and snail Dependence. It has been shown that twist and snail are the main zygotic regulators of early gene expression in the primordium of the mesoderm (7, 8). Gene expression in the region normally occupied by the mesoderm anlage of

<table>
<thead>
<tr>
<th>Gene</th>
<th>Representative clone</th>
<th>Different fragments (no. of clones)</th>
<th>mRNA size, kb</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>B3</td>
<td>2 (22, 4)</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>#32</td>
<td>6 (3, 3, 2, 2, 1, 1)</td>
<td>&gt;10</td>
</tr>
<tr>
<td>C</td>
<td>#47</td>
<td>1 (4)</td>
<td>5.5 ± 4.4</td>
</tr>
<tr>
<td>D</td>
<td>#57</td>
<td>1 (2)</td>
<td>0.85</td>
</tr>
<tr>
<td>E</td>
<td>1a5</td>
<td>1 (3)</td>
<td>4.7 ± 4</td>
</tr>
<tr>
<td>F</td>
<td>1a11</td>
<td>1 (2)</td>
<td>4.7 ± 4.4</td>
</tr>
<tr>
<td>G</td>
<td>1c2</td>
<td>1 (2)</td>
<td>1.5</td>
</tr>
<tr>
<td>DFR1</td>
<td>1e5</td>
<td>1 (3)</td>
<td>2.8*</td>
</tr>
<tr>
<td>H</td>
<td>1d7</td>
<td>1 (6)</td>
<td>2</td>
</tr>
<tr>
<td>zfh–1</td>
<td>1f4</td>
<td>1 (1)</td>
<td>7.5*</td>
</tr>
<tr>
<td>I</td>
<td>2a9</td>
<td>1 (4)</td>
<td>2.6</td>
</tr>
<tr>
<td>J</td>
<td>1h11</td>
<td>1 (1)</td>
<td>7.5</td>
</tr>
<tr>
<td>K</td>
<td>2c9</td>
<td>1 (2)</td>
<td>2</td>
</tr>
<tr>
<td>L</td>
<td>2c11</td>
<td>1 (1)</td>
<td>&gt;10</td>
</tr>
<tr>
<td>M</td>
<td>1c12</td>
<td>1 (1)</td>
<td>2</td>
</tr>
<tr>
<td>tinman</td>
<td>2d5</td>
<td>1 (1)</td>
<td>1.7*</td>
</tr>
</tbody>
</table>

*Ref. 33.
†Ref. 34 and 35.
‡Ref. 36.
embryos mutant for both genes is indistinguishable from the adjacent ectoderm (8). In the absence of twist, a number of genes that are normally expressed in the mesoderm are not activated. snail function is required for the repression in the mesoderm of genes that are otherwise expressed in the adjacent neuroectodermal regions of the blastoderm. Therefore, we would expect the novel genes we have found not to be expressed in twist embryos, whereas their early expression would be unaltered in snail mutant embryos.

We have examined the expression of the nine early expressed genes described above in embryos mutant for twist or snail and in embryos carrying mutations in both loci. Four genes (B, C, E, and I) do not show any expression in the ventral part of twist embryos at the blastoderm stage (data not shown). However, the expression detected outside of the mesoderm, such as in vitellophages or tracheal pits, is unaffected in twist embryos. In snail embryos, the level of expression of these four genes in the blastoderm as well as during the early phases of germ band extension is normal. In some cases, the expression appears to decay prematurely (data not shown), but this is most likely an indirect effect: twist expression is not maintained in snail embryos (8, 28). None of these four genes are expressed ventrally in twist snail double mutant embryos (data not shown).

The expression of the other five genes is regulated in an unconventional fashion. The five cases are, however, different from each other. The earliest expressed gene, gene A, is expressed at normal levels in twist embryos, not only at the blastoderm stage, but also during germ band extension (Fig. 3A). Its expression in early snail embryos is also normal, although the ventral region anterior to the ventral furrow does not show normal levels of expression. During gastrulation, this region seems to be void of expression. Double mutant embryos hybridized with the same probe show a signal identical to the one observed in snail embryos (data not show).

Gene D is not expressed at any stage in twist embryos (Fig. 3D). A very weak ventral expression is detected in the blastoderm in snail embryos. During germ-band extension, no expression is detected at all. The absence of twist and snail functions affects the expression of zfh-1 and DFRI in a similar way. zfh-1 and DFRI are not expressed in twist embryos and their expression in snail embryos is severely reduced (refs. 33 and 34; unpublished observations).

snail embryos hybridized with a probe for gene F show a normal expression (Fig. 3F). The signal, as in wild-type embryos, shows an anterior-posterior gradient that is maintained during the early phases of germ-band extension. In twist embryos, however, although a similar gradient is detected, the signal appears to be weaker than in wild type. The anterior expression of gene G in twist embryos is normal (Fig. 3G). In snail embryos, some anterior expression is detected at the blastoderm stage but it disappears completely during germ-band extension. The expression of gene H in twist embryos at blastoderm stage is reduced almost to the limits of detection (Fig. 3H) and disappears completely during germ band extension. In snail embryos its expression is also reduced—although not so strongly—at blastoderm stage but, as in twist mutants, it dissappears during germ-band extension.

Our observations show that establishing the expression patterns of mesodermal genes is more complicated and in-
volves input from a greater variety of regulators than previously thought. Three aspects of this are worth emphasizing: (i) the surprising variation in early mesodermal gene expression pattern implying multiple regulatory inputs; (ii) the apparent function of *snail* as an activator of mesodermal gene expression; and (iii) the difference in the regulatory role of *twist* and *snail* in the head and the trunk region of the embryo.

If mesodernally expressed genes were controlled exclusively by *twist* and *snail*, then their expression patterns should resemble those of *twist* and *snail*. However, the observed variations in expression patterns argue that mesodermal genes receive regulatory inputs from other genes as well. The most obvious example is gene A, which is expressed in a broad ventral domain independently of both *twist* and *snail* functions, suggesting that it is activated directly by *dorsal*. In addition, unlike any other known ventrally expressed genes, gene A is repressed posteriorly by *tailless*. Genes D and F, on the other hand, represent two genes that do not only receive the necessary inputs for their ventral expression but also receive inputs from genes that are implicated in patterning along the anterioposterior axis: *bicoid* in the case of gene F and a pair-rule gene in the case of gene D. Whether these modulations are functionally significant remains to be seen.

Not only *twist* but also *snail* is required for the expression of genes C, *zfh-1*, and *DRF1* in the mesoderm primordium (see also refs. 33 and 34; unpublished observations). Thus, contrary to its postulated role as a repressor of ectodermal genes (7, 8), *snail* can also have a positive regulatory function. However, gene C, *zfh-1*, and *DRF1* are still expressed, although at reduced levels in very early *snail* embryos. Therefore, *snail* is not essential for their activation, but rather their maintenance at high levels, and this may be due to an indirect function.

Finally, it is worth noticing the case of gene G which requires *snail* but not *twist* for its normal expression in the anterior part of the ventral furrow. Thus, *snail*, but not *twist*, acts as a mesodermal activator in this region. Another unusual case of this kind is *twist* itself not requiring *twist* function for their expression in the head region (8) and genes A and D depending on *snail* as an activator in the anterior region, but not in the trunk. Taken together our observations argue against the simple view that *twist* and *snail* alone, by complementary mechanisms, regulate gene expression in the mesodermal primordium (7, 8) is not longer tenable.

We gratefully acknowledge the help of Caryl Mayes and Heike Schauerte in sequencing. We thank N. Azpiazu, N. Brown, E. Knust, S. Panzer, and B. Shilo for sending flies and reagents, and Z. Wang and D. Brown for hospitality and tuition during the early steps of the project. We also thank M. Granato, P. Lawrence, M. Mullins, and R. Reuter for critical reading of different versions of the manuscript. J.C. had a European Molecular Biology Organization fellowship during part of the work. This project was supported by a grant from the Behrens-Weise-Stiftung to M.L.