THOU art sworn as deeply to effect what we intend  
As closely to conceal what we impart.  
Richard III

How can a gene organize the growth of an animal? What characteristics of master regulatory genes allow them to effect the growth of a brain or a wing, and what information do the genes impart to constituent cells that will take on widely different fates? Homeotic genes organize large arrays of cells so that each cell carries out its proper role in the formation of an overall pattern. In flies and in mammals, each homeotic gene is active in a limited region of the embryo, thus making that region distinguishable from others (for reviews see Duncan, 1987; Kaufman et al., 1990). Accordingly, mutations in homeotic genes cause one part of an animal to develop in the likeness of another part. In Drosophila, where these genes have been most extensively studied, mutations in homeotic genes cause dramatic changes: wings develop where eyes should be or legs develop in lieu of antennae. Mutations resulting in loss-of-function of homeotic genes affect the fates of cells where the gene is normally expressed, whereas mutations resulting in gain-of-function, which cause the gene to be turned on in places where it should be off, affect cells in regions where the gene is not normally required. The products of many of the homeotic genes are DNA-binding transcription factors (for reviews see Levine and Hoey, 1988; Scott et al., 1989). Thus the products of homeotic genes appear to specify the development of structures unique to each body segment by regulating the activities of subordinate target genes.

A remarkably wide range of functions are affected by the misplaced production, or lack of production, of a single homeotic protein. How can one protein have such a powerful influence? The structures characteristic of different body segments, for example antennae, legs, wings, and halteres, are very different in size, shape, and organization. Homeotic proteins act in multiple germ layers to coordinate numerous processes such as cell division, cell movement, production of specialized proteins on the surface and inside the cell, cell polarization, and the formation of asymmetric cell structures. To reveal how animals coordinate the activities of many cell types and cellular functions within the different segments, it is important to identify and characterize the target genes regulated by homeotic proteins. Yet, surprisingly few target genes have been found. The few examples of genes thought to be directly regulated by homeotic proteins include the homeotic genes themselves and two growth factor genes. Here we present a close examination of the interactions between the homeotic genes and the few identified targets. The features of these interactions may be general and useful in designing approaches for the identification of additional target genes. We also discuss the advantages and problems associated with current strategies for finding more target genes.
THE HOMEOTIC SELECTOR GENES

In *Drosophila*, the homeotic selector genes are located in two gene clusters. The genes controlling head and thoracic development, *labial* (lab), *proboscipedia* (ph), *Defor med* (Dfd), *Sex comb reduced* (Scr), and *Antennapedia* (Antip), lie in the Antennapedia complex (ANT-C; Kaufman et al., 1990); the genes controlling posterior thoracic and abdominal development, *Ultra bithorax* (Ubx), *abdominal A* (abd-A), and *Abdominal B* (Abd-B), lie in the Bithorax complex (BX-C; Duncan, 1987). For unknown reasons the order of homeotic selector genes along the chromosome parallels the anterior-posterior order of their expression and function along the body axis. Mutations in some other genes result in homeotic phenotypes; however, these genes are not localized in the homeotic gene clusters, and they can be distinguished from the homeotic selector class because their functions are not limited to one or a few segments. Many of these other homeotic genes are regulators of the ANT-C and BX-C homeotic selector genes (Lewis, 1978; Ingham and Whittle, 1980; Duncan, 1982; Struhl, 1983; Kennison and Tamkun, 1988). These genes will not be discussed further here.

Homeotic selector gene complexes also exist in mammals. The *Drosophila* and mammalian homeotic genes are similar on the basis of both order of expression along the anterior-posterior body axis and protein sequence homologies (Duboule and Dolle, 1989; Graham et al., 1989). However, unlike the *Drosophila* genes, which are located in two complexes, mammalian homologs, the Hox genes, are clustered in a single complex that has been replicated during the course of evolution to yield four clusters. Mutations in the murine homeotic genes, engineered in transgenic mice, have dramatic and specific effects on development (Ballinger et al., 1989; Chisaka and Capel, 1991), and some parallels can be drawn between the phenotypes observed in mice and those resulting from homeotic mutations in *Drosophila*. Thus, the mechanisms by which *Drosophila* homeotic genes act through their target molecules to coordinate cell fate may be a paradigm for mammalian development.

Each of the homeotic selector genes encodes a protein, or related set of proteins, containing a highly conserved, 61-amino acid long domain, the homeo-domain, which binds to DNA (Scott et al., 1989). The homeodomain, a billion-year-old structure conserved from yeast (Laughon and Scott, 1984; Shepherd et al., 1984) to plants (Vollbrecht et al., 1991) to man (Levine et al., 1984) binds DNA in vitro (Desplan et al., 1985), has been co-crystallized with DNA (Kissinger et al., 1990), and its structure when bound to DNA in solution was shown with the use of nuclear magnetic resonance methods (Otting et al., 1990). The homeodomain is required for transcriptional regulation by homeotic proteins in cell culture assays (Han et al., 1989; Jaynes and O’Farrell, 1988; Krastnow et al., 1989; Winslow et al., 1989). Thus, the homeotic selector genes probably effect the development of segment-specific structures by binding to DNA and regulating the transcriptional activities of a host of downstream targets.

WHAT TYPES OF TARGET GENES SHOULD BE EXPECTED?

What types of target molecules are likely? One class of target molecules may be required for refinement and maintenance of the expression patterns of the homeotic genes. Many of the early-acting genes that set the initial patterns of homeotic gene expression, such as members of the segmentation gene classes gap and pair-rule, are only transiently expressed and cannot maintain the patterns of homeotic gene expression throughout development. Auto-regulation and “cross-regulation” of the homeotic genes may constitute one such maintenance mechanism.

Other potential target genes may control where and when cells divide. It has been suggested that the complexity of the expression pattern of *string*, a gene that controls cell divisions, depends on earlier acting regulatory genes (Edgar and O’Farrell, 1989). The *string* product is a cell cycle regulator related to the CDC25 protein found in the yeast *Schizosaccharomyces pombe*. Similarly, *Drosophila* cyclins have intricate expression patterns (Lehner and O’Farrell, 1989; Raff et al., 1990; Whitfield et al., 1990) that might indicate regulation by segmentation and homeotic proteins. Cell division must be coordinated using cell communication so the homeotic genes may activate communication pathways among cells. For instance, the two growth factor target genes that will be discussed appear to carry positional information from the mesoderm to the endoderm.

WHAT CRITERIA CAN BE USED TO SHOW THAT A PUTATIVE TARGET IS DIRECTLY CONTROLLED BY A HOMEOTIC PROTEIN?

It is difficult to establish that a gene is directly regulated at the transcriptional level by a homeotic protein. A complete demonstration that an interaction is direct should meet at least some of the following criteria: First, expression of the target gene should change in response to a mutation in a homeotic gene in vivo, that is, in a tissue in a living animal. This result is most convincing if expression of the target gene changes in response to both the loss of the homeotic product and its ectopic expression. Second, there should be evidence for a direct association between the regulatory homeotic protein and the target gene DNA.
Typically, this means that the homeotic protein should bind the target DNA in vitro with relatively high affinity. Finally, one should be able to demonstrate that the target gene sequences bound in vitro confer regulation in vivo. This is frequently done by fusing target gene regulatory sequences to a reporter gene, such as lacZ, and expressing the reporter gene in flies. Deletions and mutations of regulatory regions of the target gene can be used to locate both the sequences relevant to homeotic regulation in vivo and the sequences required for binding in vitro. By far the most difficult experiment is also the best one: to show that a protein with altered binding specificity acts in vivo upon an engineered promoter with a correspondingly altered binding site. Only some of the experiments outlined here have been done for any of the potential targets identified to date. Nonetheless, it is very likely that some of the targets we discuss in this review are directly bound and transcriptionally regulated by homeotic proteins.

**TARGETS OF THE HOMEOTIC SELECTOR GENES INCLUDE OTHER HOMEOTIC SELECTOR GENES**

What can be learned from the few targets now being examined? The first genes identified as potential targets of homeotics were the homeotic genes themselves (Table 1). The regulatory interactions reveal the specificity of the genes for their targets and the importance of cellular environment in governing which interactions occur and whether they are negative or positive. Through autoregulatory and crossregulatory interactions, the homeotic genes regulate both the levels and patterns of homeotic gene expression. Expression of at least two homeotic genes, Dfd and Ubx, is autoregulated (Bienz and Tremml, 1988; Kuziora and McGinnis, 1988). A 1.4-kb upstream region was shown to be sufficient to mediate Ubx autoactivation in the visceral mesoderm (Müller et al., 1989), and specific sequences involved in the autoregulation of Dfd have been identified in the Dfd promoter region (Regulski et al., 1991). In vitro binding sites for the DFD protein have also been identified within this autoregulatory region of the Dfd promoter. The Dfd sequences are the best defined of any target gene sequences as direct sites of homeotic protein action.

In contrast to its positive autoregulation in the visceral mesoderm, UBX represses itself in the embryonic epidermis and nerve cord, and in the precursors to the adult cuticle, the imaginal discs (K. Irvine, J. Botas, and D.S. Hogness, personal communication). The two opposite responses of Ubx expression to the UBX protein indicate the importance of cellular environment in determining the response of the target gene. Similarly, the ability of ectopic DFD, whose expression is driven by an inducible heat shock promoter, to induce expression of the endogenous Dfd gene is restricted to a small number of cells. This restricted response to the DFD protein is another indication of factors that limit the susceptibility of target genes (Kuziora and McGinnis, 1988).

<table>
<thead>
<tr>
<th>Table 1.</th>
<th>Homeotic selector genes and their potential targets.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homoeotic selector gene</td>
<td>Potential target locus</td>
</tr>
<tr>
<td>labial (lab)</td>
<td>None known</td>
</tr>
<tr>
<td>proboscipedia (pb)</td>
<td>None known</td>
</tr>
<tr>
<td>Deformed (Dfd)</td>
<td>Deformed (Dfd)</td>
</tr>
<tr>
<td>Sex comb reduced (Scr)</td>
<td>None known</td>
</tr>
<tr>
<td>Antennapedia (Antp)</td>
<td>Sex comb reduced (Scr)</td>
</tr>
<tr>
<td>Ultrabithorax (Ubx)</td>
<td>Sex comb reduced (Scr)</td>
</tr>
<tr>
<td>Antennapedia (Antp)</td>
<td>Sex comb reduced (Scr)</td>
</tr>
<tr>
<td>Ultrabithorax (Ubx)</td>
<td>Sex comb reduced (Scr)</td>
</tr>
<tr>
<td>decapentaplegic (dpp)</td>
<td>Sex comb reduced (Scr)</td>
</tr>
<tr>
<td>labial (lab)</td>
<td></td>
</tr>
<tr>
<td>wingless (wg)</td>
<td></td>
</tr>
<tr>
<td>abdominal A (abd-A)</td>
<td></td>
</tr>
<tr>
<td>Antennapedia (Antp)</td>
<td></td>
</tr>
<tr>
<td>Ultrabithorax (Ubx)</td>
<td></td>
</tr>
<tr>
<td>decapentaplegic (dpp)</td>
<td></td>
</tr>
<tr>
<td>wingless (wg)</td>
<td></td>
</tr>
<tr>
<td>Antennapedia (Antp)</td>
<td></td>
</tr>
<tr>
<td>Ultrabithorax (Ubx)</td>
<td></td>
</tr>
<tr>
<td>Abdominal B ( Abd-B)</td>
<td></td>
</tr>
<tr>
<td>Ultrabithorax (Ubx)</td>
<td></td>
</tr>
<tr>
<td>abdominal A (abd-A)</td>
<td></td>
</tr>
</tbody>
</table>

*Abbreviations are as follows: NC, nerve cord; VM, visceral mesoderm.
by mutations in Antp. In the visceral mesoderm, SCR protein expression is reduced in Antp loss-of-function mutants; this observation suggests that expression of Antp positively regulates that of Scr in this tissue (Reuter and Scott, 1990). Regulation in the visceral mesoderm is likely to be indirect since the SCR-expressing cells never show detectable levels of the ANTP protein. Thus, cross-regulation of Scr by ANTP occurs only in some tissues, and ANTP can act directly or indirectly, and negatively or positively, on Scr.

All three genes of the BX-C complex regulate Antp expression in at least one tissue type. The UBX protein sets the posterior border of Antp expression in all germ layers where they are both expressed, that is, the visceral mesoderm, epidermis, and nerve cord (Hafen et al., 1984; Harding et al., 1985; Carroll et al., 1986). In Ubx mutants, ANTP expression extends further posteriorly into the Ubx expression domains. High-level neural and ederal expression of the ANTP protein can also be extended into both the abd-A and Abd-B domains in animals deficient for Ubx and abd-A, or for all three BX-C genes. Similarly Ubx expression is repressed by the products of the abd-A and Abd-B genes in posterior abdominal regions (Struhl and White, 1985). Some evidence suggests that regulation of Antp by UBX is direct: the UBX protein binds to Antp sequences that confer regulation by UBX protein in transient cotransfection assays in Drosophila tissue culture cells (Beachy et al., 1988; Krasnow et al., 1989). A 2-kb region near one of the Antp promoters is responsive to the BX-C genes in vivo (Boulet and Scott, 1988).

ROLES OF HOMEOTIC GENES IN MIDGUT MORPHOGENESIS

The homeotic genes are expressed in distinct patterns in multiple germ layers: in the ectoderm and its derivatives, in the visceral and somatic mesoderm, and in the endoderm. Historically, roles of homeotic proteins in development have been studied mostly in adult and larval cuticles, because they are the easiest to see. More recently, investigators have begun to examine actions of the proteins in internal tissues, particularly in the nervous system (Ghysen and Lewis, 1986; Mann and Hogness, 1990; Weinzel et al., 1987) and midgut. The midgut studies have been revealing about target genes. Studies of the homeotic proteins in the midgut are facilitated by its relative simplicity. The midgut tube contains yolk and is composed of only two cell layers, an outer visceral mesoderm and the internal endoderm. Midgut morphogenesis involves the evagination of blind-ended tubes called gastric caeca at the anterior end of the midgut and the formation of three circumferential constrictions that divide the midgut into four compartments along its length. The formation of the gastric caeca and each of the three constrictions has been shown to depend on homeotic genes expressed in the visceral mesoderm at, or adjacent to, the sites of these differentiative events.

At least six homeotic selector genes are expressed in discrete regions of the developing midgut (Diederich et al., 1989; Tremml and Bienz, 1989; DeLorenzi and Bienz, 1990; Karch et al., 1990; Reuter and Scott, 1990). Scr is expressed in the anterior-most region of the visceral mesoderm adjacent to the site of gastric caeca formation; gastric caeca fail to form in Scr mutants. Antp is expressed in the visceral mesoderm surrounding the site of the most anterior constriction, and the absence of Antp results in the failure of this constriction to form. Similarly, Ubx and abd-A expression domains meet at the site of the second constriction, and the second constriction depends on the wild-type activities of both genes. The third constriction, which forms in the center of the abd-A expression domain, is dependent on abd-A function. Abd-B is expressed in the most posterior region of the visceral mesoderm, but its role in gut morphogenesis is unknown. The lab gene is expressed in the endoderm underlying the second Ubx/abd-A-dependent constriction, but there is no obvious change in gut morphology in lab mutants.

The homeotic genes interact in the visceral mesoderm, but not according to the same rules that apply in other tissues. Ubx is derepressed in more posterior regions by abd-A mutants but, in contrast to the situation in epidermis and nervous system, is unaffected by additional Abd-B mutations. Antp is derepressed in the Ubx domain by Ubx mutants, and this shift in Antp expression changes midgut morphology: the first constriction is shifted more posteriorly, maintaining its position near the center of the Antp expression domain (Tremml and Bienz, 1989). Unlike Antp expression in the epidermis and nerve cord, the visceral mesoderm Antp expression is unaffected by either abd-A or Abd-B.

TARGETS OF HOMEOTIC PROTEINS IN THE MIDGUT INCLUDE TWO GROWTH FACTOR–LIKE PROTEINS

Two potentially direct targets of homeotic proteins in the midgut are not themselves homeotic genes. The wingless (wg) and decapentaplegic (dpp) genes were both identified as potential targets on the basis of their patterns of expression in the visceral mesoderm (Immerglück et al., 1990; Panganiban et al., 1990a; Reuter et al., 1990). wg is a segment polarity gene, which, like other genes in this class, affects pattern within each segment (Nüsslein-Volhard and Wieschaus, 1980; van den Heuvel et al., 1989), and is the Drosophila homolog of the int-1 mammalian oncogene (Rijswijk et al., 1989).
1987). dpp is required for normal gastrulation and dorsal-ventral axis polarity in the embryo and the adult precursor cells (Spencer et al., 1982; St. Johnston and Gelbart, 1987). dpp is a member of the transforming growth factor-β gene family (Padgett et al., 1987). The wg and dpp gene products are secreted proteins (Panganiban et al., 1990b; van den Heuvel et al., 1989). Both the expression patterns of wg and dpp in various homotic mutants and initial DNA binding studies with UBX protein provided evidence that the two genes are targets of homotic proteins.

Expression of wg is regulated by ABD-A, and possibly UBX, but only in the visceral mesoderm (Immerglück et al., 1990; Reuter et al., 1990). wg expression is unaffected by abd-A or Ubx mutations in other embryonic tissues. Normally the protein encoded by wg is first apparent in the visceral mesoderm in some of the cells expressing abd-A and migrates to the underlying endoderm cells. Therefore, this protein may be involved in communication between the two germ layers. In abd-A mutant embryos the loss of wg midgut expression suggests that ABD-A is necessary for activation of wg in this tissue. Moreover, the presence in wild-type embryos of the wg gene product in only some of the cells that produce ABD-A protein indicates that another factor must limit the region in which wg expression is activated. It was recently shown that this factor is not ABD-B (Reuter et al., 1990). Expression of Ubx may also be required indirectly for the activation of wg in the midgut, since with some, but not all, Ubx mutant alleles wg expression in its normal midgut domain is significantly reduced. Why Ubx alleles behave differently is unclear.

In the visceral mesoderm cells dpp RNA is made at high levels at two locations in the midgut: the base of the gastric caeca and the second constriction (Immerglück et al., 1990; Panganiban et al., 1990a). Expression of dpp is regulated in opposite ways by two different homotic proteins (Immerglück et al., 1990; Reuter et al., 1990). Its expression near the second constriction is activated by the UBX protein and repressed by the ABD-A protein. In Ubx loss-of-function mutants, dpp expression is reduced in the second constriction, whereas in abd-A loss-of-function mutants it is extended posteriorly, presumably in response to the extended domain of Ubx in abd-A mutants. In animals carrying a Ubx cDNA driven by the promoter of the heat shock gene hsp-70 (HS-Ubx), a heat shock treatment causes the UBX protein to be expressed in every cell. This leads to the expansion of dpp expression in the midgut from its normal site to the second constriction throughout all of the anterior midgut. Surprisingly, high levels of UBX protein in posterior midgut cells fail to activate dpp. Heat-shocked HS-Ubx animals that are deficient for Abd-A express the dpp gene product at high levels throughout the visceral mesoderm, indicat-

ing that posterior midgut dpp expression is prevented by ABD-A. Thus, the presence of one homotic protein can prevent another homotic protein from activating its target. Also, Ubx expression in tissues other than the midgut fails to activate dpp.

DPP is necessary for at least two morphogenetic events in the midgut, the formation of gastric caeca and of the second constriction. In the shortvein class of dpp mutations (dpp<sup>av</sup>), the dpp gene product is not expressed in the visceral mesoderm (Immerglück et al., 1990; Panganiban et al., 1990a) and both the gastric caeca and the second constriction fail to form. Therefore, dpp is not only a possible target of UBX; it is essential for a morphogenetic event that is under the influence of UBX. In dpp<sup>av</sup> mutants, WG protein levels in the adjacent visceral mesoderm cells are reduced to the levels observed in some UBX mutants, so UBX effects on wg expression in adjacent visceral mesoderm cells could be mediated by DPP.

Like the WG protein, DPP protein is made in the visceral mesoderm and transported to the underlying endoderm; there DPP induces lab expression (Immerglück et al., 1990; Reuter et al., 1990). In wild-type embryos, lab is expressed in the nuclei of endoderm cells that underlie the Ubx- and dpp-expressing visceral mesoderm. In both loss-of-function Ubx mutants and dpp<sup>av</sup> mutants, expression of lab in the endoderm is lost. Correspondingly, ectopic expression of lab can be induced in the endoderm by expressing UBX more anteriorly using a heat-shock construct or by expressing UBX protein further posteriorly in abd-A mutants. Since LAB protein accumulation parallels dpp expression in the midgut, its expression is presumably a consequence of DPP protein moving into these cells.

DPP also regulates homotic genes (Panganiban et al., 1990a), indicating that a target gene can have a feedback effect on homotic gene expression and function. At the anterior end of the midgut DPP function normally prevents Scr expression in the cells that actually form the gastric caeca. In the second constriction, DPP is required in the visceral mesoderm to maintain high level Ubx expression. Therefore, a mutation in a homotic gene may affect its own expression, and this regulation is mediated by one or more of the targets of the homotic gene. Indeed, UBX activates its own expression in the visceral mesoderm (Bienz and Treml, 1988), although it is not yet known whether this is direct or mediated by DPP.

The growth factors DPP and WG are both excellent candidates for downstream targets of UBX and ABD-A. However, evidence for a direct relationship exists only in the case of the dpp gene and the UBX protein. Partially purified UBX protein binds dpp DNA in vitro and seven copies of the “optimal” UBX binding site, determined by binding competition of random double-stranded twelvemers centered within a
70-nucleotide DNA fragment, are found in the shortvein region of the dpp gene (Ekker et al., 1991).

TARGET GENES HAVE MULTIPLE ROLES IN DEVELOPMENT

The known targets of homeotic genes are involved in more than one developmental process. Both WG and DPP have other roles in embryogenesis. WG is required for the formation of anterior structures in every embryonic segment and for proper wing development. DPP is required for the establishment of the dorsal-ventral polarity axis in early embryogenesis and for the formation of all the appendages in adults. Moreover, the expression of WG and DPP is not restricted to the cells where they are controlled by homeotic genes. Rather, in some tissues and some stages, the growth factor-like proteins are regulated by homeotic genes; at other times and places, different regulators give rise to quite distinct WG and DPP expression patterns. In keeping with the expression patterns, the phenotypic effects of complete loss-of-function mutations in wg or dpp are neither strictly nor completely homeotic. The similarity in phenotype observed in the midgut of Ubx and dpp mutants is only revealed by a special class of dpp mutations, the shortvein alleles. In the shortvein mutants the effect of DPP on midgut expression can be separated from most of those of the other dpp phenotypes. dpp complete loss-of-function mutants do not develop sufficiently for defects in gut morphogenesis to be observed. Therefore, in seeking other target genes it will probably not be fruitful to identify only genes with homeotic phenotypes.

Thus far, the known target genes appear to be regulatory proteins; the few examples of potential target molecules suggest that many genes directly controlled by homeotic proteins may themselves regulate the activities of genes further downstream. Therefore, the homeotic genes may be at least one step removed from genes whose products affect the final morphology and physiology of differentiated cells. However, extrapolation from the few known candidate target genes, a set biased by their own interesting phenotypes, may be misleading. It is certainly possible that the full set of direct targets will also include structural proteins and enzymes.

APPROACHES TO TARGET GENE IDENTIFICATION

Various approaches are proving to be useful in searching for target genes, including genetic screens, enhancer traps, DNA binding in vitro, polytene chromosome localization, and immunoprecipitation. These approaches have different strengths and weaknesses that are discussed briefly here.

Genetic Screens

A gene dedicated to a single process, for example sex determination, can be identified because mutations in the gene affect only the sex of the animal and nothing else. Many and perhaps most of the genes that control sex determination in flies (Baker, 1980; Baker and Belote, 1983; Clime, 1980; Clime, 1983) and in worms (Hodgkin, 1987; Hodgkin, 1990; Meyer, 1988; Villeneuve and Meyer, 1990) were identified because of mutations that switch the sex of the organism to either the opposite sex or some intermediate. The relevant genes were ordered within each hierarchy by epistasis tests, and fall into largely linear pathways. The fly sex determination gene hierarchy has been characterized all the way from the genes that respond to, and perhaps define, the sex-determining chromosome ratio (Clime, 1988; Cronmiller and Clime, 1987; Oliver et al., 1988; Parkhurst et al., 1990; Steinmann-Zwicky, 1988; Torres and Sánchez, 1989) to specialized differentiation functions such as the production of female-specific yolk proteins (Belote et al., 1985).

Some of the inferred interactions among sex determination genes may not be direct, simply because some components of the sex determination hierarchy are required in other processes as well. Perhaps for this same reason, classical genetic approaches have yet to be successful in identifying genes downstream of the homeotic genes. Mutations in target genes may have many effects, only some of which might be recognizable as homeotic. Also, homeotic transformations of segment fates may involve the concerted actions of large arrays of genes, so mutating any one target gene will not give a homeotic phenotype.

One genetic approach to identify mutations in homeotic target genes is to look for mutations that either suppress or enhance a homeotic phenotype. In this way, subtle effects of genes may be detected. This approach will potentially detect regulators of homeotic gene expression, cofactors for homeotic protein function, and genes regulated by homeotic proteins. Such a screen (Kennison and Russell, 1987; Kennison and Tamkun, 1988) has identified a novel group of genes, some of which may be directly downstream. One of the genes in the group, Distal-less (also known as Brista) encodes a homedomain protein and is therefore presumably itself a regulator (Cohen et al., 1989). Genetic screens carried out in animals sensitized to further perturbations in the homeotic pathway may lead to the discovery of target genes where other screens would fail because of the lethality of null mutations in the target genes or roles of the target genes in other developmental events.
DNA Binding

Because homeotic genes encode DNA-binding proteins, one approach to finding direct targets is to isolate sequences that preferentially bind to homeotic proteins in vitro. Although this approach seems relatively straightforward, homeodomain-containing proteins bind to many sequences in the test tube, including, for example, sites in bacteriophage lambda DNA, which are presumably not regulated by any homeotic protein. Even though to date most studies of DNA binding have employed sequences not known to be regulated by the homeotic protein used, valuable information has been obtained about the capabilities and structures of the proteins (Affolter et al., 1990; Beachy et al., 1988; Billete et al., 1990; Ekker et al., 1991; Laughon et al., 1988; Otting et al., 1990; Percival-Smith et al., 1990).

One must be able to determine the significance of in vitro DNA binding and its relationship to function in vivo. The conditions for binding in vivo may be quite distinct from those chosen for experiments in vitro. The presence or absence of possible cofactors may affect binding site selection, although as yet little evidence implicates such cofactors for selector gene proteins. In practice it has been most effective to identify cis-acting regulatory elements using tests in vivo before using DNA-binding assays to identify the precise relevant sequences. This procedure has led to the identification of binding sites for the DFD homeotic protein within the region necessary for Dfd autoregulation (Regulski et al., 1991).

cDNA Subtraction

Another approach to finding targets is to identify transcripts in embryos or cultured cells that are preferentially activated or repressed by induced expression of a homeotic protein. That this approach is of value is suggested by studies demonstrating that the UBX protein can repress transcription from the Antp gene when coexpressed in transient cotransfection assays in tissue culture cells (Krasnow et al., 1989) or in embryos that can be induced to express UBX protein by means of an introduced construct (HS-Ubx). The assay will potentially identify both direct and indirect target genes, and will be most useful for abundant transcripts. Once sequences have been isolated in this way, they can be tested by other methods for their relationship to a homeotic protein. The best assay is to use in situ hybridization to ask whether the expression pattern of the identified sequence in embryos is responsive to homeotic mutations. There are several limitations to the method. For example, minor changes in transcript abundance are impossible to detect, yet changes of only twofold in the level of homeotic proteins lead to morphological transformations in flies. In addition, subtractive hybridization methods with the use of embryonic RNA may fail to identify sequences that may be increased by homeotic gene action in one part of the embryo, but compensatorily decreased in another part.

Immunoprecipitation

A variation of the in vitro DNA-binding approach has been quite successful in identifying potential direct targets (Gould et al., 1990). DNA associated with UBX protein was purified from chromatin fractions using a biotinylated antibody against UBX. After a second in vitro selection step, clones containing the consensus binding sites for ANTP and UBX proteins were found to be enriched relative to control clones isolated in the absence of UBX antibody. Four newly isolated genomic DNA clones were mapped to chromosomes and their expression patterns examined by whole mount in situ hybridizations to wild-type and mutant embryos. The expression patterns of two of the clones changed in BX-C mutants, in ways that suggest regulation by UBX. These same clones are selectively bound by partially purified UBX protein in the presence of a high concentration of competitor salmon sperm DNA, so sites bound by the UBX protein in vivo and in vitro may overlap. The expression patterns of the two isolated clones in both wild-type and BX-C mutant embryos are consistent with regulation by either ANTP or UBX, so more work must be done to demonstrate direct regulation by UBX.

Suggestive Gene Expression Patterns

One approach to identifying target genes is to look for genes expressed in some of the same cells where a homeotic gene is active. It is then easy to test whether the potential target gene's expression is affected by mutations in the homeotic gene. This approach enables the identification of both direct and indirect targets. One obvious advantage of identifying already known genes as homeotic targets is that a previously characterized target gene will already be better understood: its products, expression, and some of its developmental roles may be known. The dpp and wg genes were identified as targets of homeotic genes in this way. A related but more powerful approach does not rely on serendipitous previous characterization of target genes. “Enhancer traps” (or “enhancer detectors”) can be used to examine the expression patterns of large numbers of genes without knowing the functions or identities of the genes (O'Kane and Gehring, 1987; Bellon et al., 1989; Bier et al., 1989; Fasano and Kerridge, 1988; Grossniklaus et al., 1989; Wilson et al., 1989). The enhancer traps used to date are P element transposons carrying the lacZ coding region downstream of a short, relatively inactive promoter. The
transposon is mobilized with transposase from a P element that cannot itself jump. lacZ expression occurs as a result of the transposon inserting in the neighborhood of enhancers, and a startlingly high frequency of interesting patterns is obtained (for a review see Wilson et al., 1990). Thousands of such cell lines, each carrying a single P element, have been examined in many laboratories. Some of the enhancer trap lines express lacZ in patterns suggestive of homeotic gene regulation. The responsiveness of the enhancer trap to homeotic regulation can be tested by crossing the transposon insert into homeotic mutant background and observing changes in the lacZ expression pattern. Using the P element sequence as a molecular tag, it is possible to quickly clone the gene. In addition, if the original insert did not mutate the gene, mobilization of the P element can create deletions that may reveal the function of the gene.

**Polytene Chromosome Immunolocalization**

Another approach to identifying genes directly regulated by homeotic proteins is to use antibodies to map binding sites for these proteins on the giant polytene chromosomes from Drosophila salivary glands. Polytene chromosomes provide a highly amplified map of the genome. The maximum resolution in finely banded regions is less than 10 kb, although more typical resolving power is in the 50 to 150 kb range. Because the protein is fixed where it was located on chromosomes in the living animal, the binding sites may indicate regulatory interactions in vivo. A wide variety of proteins have been located on polytene chromosomes, such as RNA polymerase (Greenleaf et al., 1978) and chromatin components of unknown function (Elgin et al., 1977; Silver and Elgin, 1976), including a zinc finger protein, PEP (Amero et al., 1991). Antibodies against three regulatory Drosophila proteins, ZESTE (Z) (Benson and Pirrotta, 1988), POLYCOMB (PC) (Zink and Paro, 1989), and a regulator involved in hormonal control, the ec dysone-induc ed protein E74 (Urness and Thummel, 1990), have been used to detect protein binding to specific sites on the polytene chromosomes. Fewer than 100 binding sites were found for each of these proteins, and these include binding sites at or near genes suspected to be regulated by these proteins.

Transposed genes or chromosomal rearrangements can be used to determine exactly which gene corresponds to a particular binding site. The protein PC is a regulator of the homeotic genes (Lewis, 1978). Bound PC was found at or near several genes it is known to regulate, including the engrailed segmentation gene and the ANT-C and BX-C homeotic genes. Fragments of the Antp gene, transposed to new locations using P elements, were found to attract PC protein to new chromosomal locations (Zink et al., 1991).

It remains to be seen how successfully polytene mapping can be applied to the homeotic proteins. There are potential limitations to the polytene chromosome approach. Salivary glands are the source of the best polytene chromosomes, and if a homeotic protein is not normally made in this tissue it will have to be ectopically activated. Other cell-type specific factors may not be present in the salivary gland cells, so the binding sites observed may not represent the full repertoire of targets and may include nonspecific sites. In addition, some of the sites detected may not correspond to downstream genes; the protein may accumulate there but have no effect on the expression of any gene. Another limitation is that the mapping provides no easy way to obtain mutants or clones for the genes of interest. However, if the same putative target is identified by some of the methods described above as well as by polytene mapping, many of these problems disappear.

**CONCLUDING REMARKS**

The study of homeotic genes is entering an exciting new phase as target genes are pursued to make the link between transcription and morphogenesis. Given the multitude of processes influenced by homeotic genes in the epidermis, mesoderm, and nervous system, there are likely to be many targets for the homeotic genes and each of these genes is probably used for more than one purpose. Over the past decade we have learned that homeotic phenotypes are often due to misexpression of a transcription factor. Over the next decade we can hope to learn how a single homeotic protein can have such profound effects on cell fates and morphogenesis.

**Acknowledgments**

We are grateful to S. Cordes, M. Petitt, S. Pfeffer, and C. Wong for comments on the manuscript. Work in our laboratory is supported by NIH grant no. 18163. We also acknowledge the support by an American Cancer Society postdoctoral fellowship to D.J.A. We thank K. Irvine, J. Botas, and D. Hogness for communication of unpublished results.

**REFERENCES**


Targets of homeotic genes / 13


Cline TW (1988): Evidence that sisterless-a and sisterless-b are two of several discrete "Numerator Elements" of the X/A sex determination signal in Drosophila that switch Sd between two alternative stable expression states. Genetics 119:829-862


Crommoller C, Cline TW (1987): The Drosophila sex determination gene daughterless has different functions in the germline as the soma. Cell 48:479-487


Percival-Smith A, Müller M, Affolter M, Gehring WJ (1990): The interaction with DNA of wild-type and mutant fushi tarazu homeodomain. EMBO J 9:3967-3974
St. Johnston RD, Gelbart WM (1987): Decapentaplegic transcripts are localized along the dorsal-ventral axis of the Drosophila embryo. EMBO J 6:2785-2791
Torres M, Sánchez L (1989): The seute (T4) gene acts as a numerator element of the X-A signal that determines the state of activity of Sex lethal in Drosophila. EMBO J 8:3079-3086

Oliver B, Perrimon N, Mahowald AP (1988): Genetic evidence that the sans fille locus is involved in Drosophila sex determination. Genetics 120:159-171


