**Dachsous encodes a member of the cadherin superfamily that controls imaginal disc morphogenesis in *Drosophila***

Hilary F. Clark,1,3,5 Doris Brentrup,2,5 Kay Schneitz,2,4 Allan Bieber,3 Corey Goodman,1 and Markus Noll2,6

1Howard Hughes Medical Institute, Department of Molecular and Cell Biology, University of California, Berkeley, California 94720 USA; 2Institute for Molecular Biology II, University of Zurich, CH-8057 Zurich, Switzerland; 3Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907 USA

Mutations in the *dachsous* gene of *Drosophila* lead to striking defects in the morphogenesis of the thorax, legs, and wings. The *dachsous* gene has been cloned and shown to encode a huge transmembrane protein that is a member of the cadherin superfamily, similar to the *fat* gene reported previously. Both the *Dachsous* and *Fat* proteins contain large tandem arrays of cadherin domains—27 and 34, respectively—as compared with 4 cadherin domains in classic vertebrate cadherins. In addition, *Dachsous* and *Fat* each has a cytoplasmic domain with sequence similarity to the cytoplasmic β-catenin-binding domain of classic vertebrate cadherins. The *dachsous* gene is expressed in the ectoderm of embryos, whereas its expression in larvae is restricted to imaginal discs and specific regions of the brain. The phenotypes of, and genetic interactions between *dachsous* and *fat* are consistent with a model in which cell proliferation and morphogenesis of imaginal structures depends on the coupled equilibria between homo- and heterophilic interactions of the *Dachsous* and *Fat* cadherin proteins.

[Key Words: Cadherin, cell adhesion; *dachsous*, disc morphogenesis]

Received March 6, 1995; revised version accepted May 3, 1995.

Our interest in undertaking a molecular genetic analysis of cadherin function in *Drosophila* was prompted by its apparent importance during vertebrate morphogenesis. Cadherins are glycoproteins that mediate Ca2+-dependent cell adhesion at adherens junctions and desmosomes (for review, see Magee and Buxton 1991; Takeichi 1991; Kemler 1993). All members of the large cadherin gene family share the characteristic cadherin domains thought to be responsible for the Ca2+-dependent cell adhesion. These cadherin domains are ~100 amino acids long and 30% homologous to each other. Classic vertebrate cadherins consist of four tandemly repeated extracellular cadherin domains, a single transmembrane domain, and a conserved cytoplasmic domain that interacts via specific catenins with the actin filaments of the cytoskeleton. In humans, cadherins have been implicated in the suppression of metastasis (e.g., Fri xen et al. 1991; Vlminckx et al. 1991) and in the autoimmune skin disease pemphigus vulgaris (Amagai et al. 1991).

Two loci in *Drosophila* were found to encode cadherin domains, the *fat* (*ft*) gene at chromosomal position 24D and an unidentified gene at 21D (Mahoney et al. 1991).

The *ft* tumor suppressor gene encodes a novel member of the cadherin superfamily. In contrast to typical vertebrate cadherins that contain four tandem cadherin domains, the extracellular region of the Ft protein consists of 34 tandem cadherin domains, followed by five epidermal growth factor (EGF)-like repeats and two laminin A G-domain-like repeats (Mahoney et al. 1991; Patthy 1992). Recessive lethal mutations in *ft* cause hyperplastic, tumor-like overgrowth of larval imaginal discs in a cell-autonomous fashion, defects in differentiation and morphogenesis, and death during the pupal stage (Bryant et al. 1988; Mahoney et al. 1991). Recently, a third cadherin, more closely related to classic vertebrate cadherins, was isolated in *Drosophila* (Oda et al. 1994). The unidentified cadherin gene at 21D will be described here and shown to be encoded by the *dachsous* (*ds*) gene, which had been cloned by chromosomal walking in an independent approach to isolate this morphogenetic gene.

The *ft* and *ds* genes have long been known to play important roles during imaginal disc development and morphogenesis. A spontaneous mutation at the *ds* locus, *ds*1, was discovered in 1917 by Calvin Bridges (Stern and Bridges 1926). The first recessive and dominant mutant alleles of *ft*, *ft*1 and *Gull*, were isolated 2 years later, and the similarity between their phenotypes and that of *ds*1 suggested that *ds* and *ft* might function in the same ge-
Dachsous cadherin controls morphogenesis

Results

Cloning of a cadherin-like gene in the chromosomal band 21D1,2

Two different DNA segments encoding cadherin domains were amplified by PCR from genomic Drosophila DNA. One DNA sequence, located in the chromosomal interval 24D, was shown to originate from the ft gene, and the other from an unknown gene at 21D [Mahoney et al. 1991]. Independently, a chromosomal walk including the aristales [all] gene [Schneitz et al. 1993] was extended proximally into the 21D1,2 region to clone the ds gene. Hybridization of cDNA clones, isolated by the use of the PCR product from 21D as probe, to DNA of the chromosomal walk showed that it mapped to the region of the ds locus. As described below, the cadherin domain was subsequently shown to be encoded by the ds gene.

Starting with the initial PCR product, many overlapping cDNAs were isolated from oligo(dT)-primed as well as randomly primed 9- to 12-hr embryonic cDNA libraries until the 10.7-kb open reading frame (ORF) was covered by a composite cDNA length of 12.4 kb (Fig. 1). Overlapping cDNAs and the corresponding 12 exons of genomic DNA were sequenced to derive a protein of 3503 amino acids. The entire ORF contains two methionines that are located 40 and 53 amino acids upstream of the initiator methionine of the protein shown in Figure 2. However, because an unusually long signal peptide would result if the first or second methionine were used as initiator, we assume that the protein begins with the third methionine and a signal sequence that is probably cleaved after 20 amino acids. After amino acid 3043, another stretch of 25 hydrophobic amino acids, presumably a transmembrane domain, is followed by several basic amino acids on the cytoplasmic side, suggesting that ds encodes a transmembrane protein. The amino-terminal extracellular domain includes 27 tandemly repeated domains of ~110 amino acids each that are similar in sequence to those found in all other members of the cadherin superfamily. In particular, most of the key amino acids of the consensus sequence of cadherin domains and their putative Ca2+-binding sites [Ringwald et al. 1987; Ozawa et al. 1990a] are conserved in the Ds protein (Fig. 2).

Only a short stretch of amino acids separates the last cadherin domain from the transmembrane domain (Fig. 2). This is in contrast to the Ft protein (Fig. 3A), in which, after the last of its 34 cadherin domains, the extracellular portion includes five EGF-like repeats [Mahoney et al. 1991], the last being flanked by two G-domain-like repeats of laminin A [Patthy 1992; Wodarz et al. 1993].

The cytoplasmic domain of the Ds protein has sequence similarity to that of the classic vertebrate cadherins, such as E-cadherin [Fig. 3B]. This sequence, which corresponds to the β-catenin-binding region in classic cadherins [Nagauchi and Takeichi 1989; Ozawa et al. 1990b], is interrupted in the cytoplasmic domain of both Ds and Ft (Fig. 3). The intervening peptide between the two conserved regions largely accounts for the larger cytoplasmic domain of Ds as compared with that of vertebrate cadherins.

The cadherin-like protein is encoded by the ds gene

The ds gene was cloned by extending a 240-kb chromosomal walk, comprising the deficiency Df(2L)al at 21C [Schneitz et al. 1993], by 185 kb into the chromosomal band 21D1,2. Mapping the ds cDNAs to the overlapping clones of chromosomal DNA shows that the primary ds transcript extends from a centromere proximal start site over >65 kb toward the left telomere of the second chromosome (Fig. 1). Eleven introns were identified by mapping and sequencing all intron-exon boundaries of the genomic DNA. The largest intron is located ~1.7 kb downstream from the translational start codon and consists of ~40 kb (Fig. 1).

Several breakpoints of chromosomal rearrangements and restriction fragment length polymorphisms (RFLPs)
Clark et al.

Figure 1. The ds gene encodes a huge transmembrane protein. (Top) Several rearrangements of the ds locus (cf. Table 1) are mapped with respect to a genomic EcoRI map (a shorter vertical line separates two neighboring EcoRI fragments whose order has not been determined) at chromosomal band 21D1,2 (brackets and open bar indicate the limits of the regions that include the breakpoints). The scale underneath indicates the distance [in kb] from the origin of a chromosomal walk (located at 21C7,8; Schneitz et al. 1993) that includes the clones Z6-2 to Z18-2 isolated from a Ks/CyO library in EMBL4. Below the genomic clones, the exon/intron structure of the ds transcript is shown with the corresponding coding regions of the Ds protein consisting of a signal peptide (SP), 27 extracellular cadherin domains (stippled), a transmembrane domain (TM), and a cytoplasmic domain including two regions (R1, R2) that are homologous to the β-catenin-binding domain of vertebrate cadherins (cf. Fig. 3). (Bottom) The location of several sequenced ds cDNAs are shown with respect to the encoded Ds protein with vertical lines indicating the positions of unspliced intron sequences [scale underneath in kb].

of ds mutant alleles (Table 1) have been mapped throughout the gene encoding the large cadherin-like protein [Fig. 1]. Two deficiencies uncovering ds, Df(2L)ast1 and Df(2L)S2, delete at least 65 kb of the ds transcript and its entire upstream region. In contrast, Df(2L)ast10, which complements other ds mutant alleles, is located proximal to clone Z18-2 of our chromosomal walk [Fig. 1]. Moreover, two inversion breakpoints associated with ds, ds55 on the SM5 balancer chromosome, and ds8260b on the In(2LR)Pm chromosome, are located within the large intron of the ds transcript.

A P-element-lacZ enhancer trap line, 2D60 (kindly provided by D. Cimbora and S. Sakonju, University of Utah, Salt Lake City), that exhibits a strong ds phenotype in trans over other ds alleles, contains two P elements on the second chromosome whose flanking DNA was recovered by plasmid rescue [D. Cimbora and S. Sakonju, unpub.]. One of these insertions, 2D60b, maps to the same 4.7-kb genomic EcoRI fragment as the 3' end of the second exon and is located 1.9 kb downstream from the 5' end of the 40-kb intron [Fig. 1]. The other P element was removed by recombination to produce a line retaining a single P element at 21D1,2. The chromosome containing this single P element in the 40-kb intron fails to complement ds and thus is itself a ds allele, ds8260b. This conclusion is corroborated by the observation that excision of the P element in 15 independent lines was accompanied by the loss of the ds phenotype.

The location of these insertion and inversion breakpoints within the large intron, however, does not exclude the possibility that ds is located entirely within the 40-kb intron and is different from the cadherin-like gene. Therefore, additional ds alleles were screened for breakpoints or RFLPs by genomic Southern blot analysis, using cloned cDNAs as probe. In these screens two alleles were found to affect the coding region of the ds transcript [RFLPs of ds alleles that lie entirely within introns would have escaped detection]. First, a reciprocal translocation between the second and third chromosome that breaks at 21D of the second chromosome, ds1524 (Craymer 1980), translocates the 5' portion, including part of the 40-kb intron of ds, to the third chromosome, and thus
Figure 2. The Ds protein belongs to the cadherin superfamily. The amino acid sequence of the putative Ds protein is shown to consist of a signal sequence, 27 cadherin domains, a transmembrane domain, and a cytoplasmic domain. The cadherin domains with their two characteristic Ca$^{2+}$-binding regions (Ringwald et al. 1987; Ozawa et al. 1990a) are aligned for optimal homology between themselves, which results in excellent alignment with the cadherin domains of Ft (Mahoney et al. 1991) and classic vertebrate cadherins (Hatta et al. 1988). To illustrate the extent of homology among the Ds cadherin domains, one or few amino acids that occur at a certain position with high frequency are highlighted as dark boxes, emphasizing the high conservation of a motif rather than of a unique amino acid sequence among cadherin domains. Note that if few amino acids are preferred at the same position, they belong frequently, yet not always, to a group of amino acids with similar properties (S/T; I/L/V; E/D; K/R). The amino acid sequence has been derived from the longest ORF of overlapping cDNAs and their corresponding genomic DNAs. Translation of the Ds protein is assumed to start at the third methionine of the ORF, which immediately precedes the putative signal sequence and, hence, is designated as position 1 of the amino acids numbered at right. The 23 amino acids that are deleted from the eleventh cadherin domain in the Ds$^s$ mutant protein are boxed.

separates the region encoding the first five cadherin domains from the remaining 3' portion of ds.

In addition, RFLPs of the ds$^s$ allele were detected. Isolation of ds$^s$ DNA from a genomic library and sequencing revealed a 7.6-kb insertion of the 412 transposable element (Finnegan et al. 1978; Will et al. 1981; Shepherd and Finnegan 1984) within the region encoding the eleventh cadherin domain very close to the 3' end of the fifth exon (Fig. 4). Because ds$^s$ is a weak allele, the question arises why the insertion of a large transposon in the middle of the ORF does not cause a more severe effect on the phenotype. The answer comes from sequence analysis of ds$^s$-cDNA, which reveals in embryos and third-instar larvae three types of splice products that lack the 412 insertion. Apparently, the wild-type splice donor site of the fifth intron is inactivated by the close proximity of the 412 insertion. As a result, three cryptic splice donor sites are uncovered, two upstream of and close to the 412 insertion and a third site 30 bp within the 412 element [Fig. 4]. Whereas one of the two upstream ds$^s$ splice products generates a frameshift resulting in a truncated, secreted protein, the reading frame of the other product remains unchanged and encodes a protein in which merely 23 amino acids of the eleventh cadherin domain are deleted (Figs. 2 and 4). The use of the third cryptic splice donor site also results in the premature release and secretion of a truncated protein (Fig. 4). Because ds$^s$ is a spontaneous mutation (Stern and Bridges 1926), it is not surprising that it is caused by the insertion of a transposon. Interestingly, the spontaneous ft allele Gull is also an insertion of the 412 transposon in the ORF of a cadherin domain (Mahoney et al. 1991).

Expression of ds in embryonic ectoderm, larval brain, and imaginal discs

Consistent with a composite cDNA length of 12.4 kb, Northern blot analysis revealed a very large transcript...
well above 9.5 kb that persisted throughout embryogenesis and was also present in third-instar larvae and adults.

The spatial distributions of ds transcripts were analyzed by in situ hybridization to whole-mount embryos (Fig. 5). The first ds transcripts are detectable during gastrulation in a pair-rule pattern of six weak epidermal stripes and in a prominent stripe at the amnioserosal invagination (Fig. 5A). During germ-band extension, ds RNA accumulates in a segmentally repeated pattern of 14 stripes and in the procephalon (Fig. 5B). The most pronounced expression of ds is observed during the extended germ-band stage mainly in the forming tracheal pits (Fig. 5C). At the beginning of head involution, ds RNA appears in the anterior spiracles and again in stripes of the segmental grooves and buds while it remains weakly expressed in the remnants of the tracheal pits (Fig. 5D). In addition, ds RNA is first detected in the primordial leg discs that form in the ventral posterior part of each thoracic segment (Bate and Martinez-Arias 1991; Cohen et al. 1991). At late stage 14, ds is expressed strongly in the nearly fused labial buds and at invaginations of the maxillary segment while it continues to be expressed in the leg disc primordia (Fig. 5E, H), along the segmental folds (Fig. 5E, G), and probably in the appendages (Fig. 5G). After dorsal closure, ds expression persists only in the appendages and in the head region (Fig. 5F, I).

In third-instar larvae, ds transcripts are found in the imaginal discs and the brain (Fig. 6). In the supraoesophageal ganglion, ds is expressed in two areas of the optic lobe and in a region that might belong to the mushroom body (Fig. 6A). In imaginal discs, strong ds expression occurs frequently along folds separating the anlagen of distinct imaginal structures (Fig. 6B–F). In the antennal disc, ds is expressed in the arista and first and second antennal segment anlagen while, in the eye disc, ds transcripts are abundant along the folds of the future bristle region of the eye disc (not shown). In leg discs, ds is expressed strongly in the anlagen of the tarsal joints and, particularly, of the most proximal leg segment (Fig. 6D). Similarly, ds expression is strongest in the pleural, dorsal...
The remaining stocks were from the Bloomington Stock Cen-
hinge, and prescutal regions of the wing disc, whereas
unknown Garcia-Bellido N.D.

Table 1. ds alleles

<table>
<thead>
<tr>
<th>Allele</th>
<th>Origin</th>
<th>Discoverer</th>
<th>Eclosion</th>
</tr>
</thead>
<tbody>
<tr>
<td>ds¹</td>
<td>spont.</td>
<td>Bridges, 17k12</td>
<td>100%</td>
</tr>
<tr>
<td>dsšw</td>
<td>spont.</td>
<td>Bridges, 29d24</td>
<td>40%</td>
</tr>
<tr>
<td>dsš38k</td>
<td>spont.</td>
<td>Bridges, 33k28</td>
<td>N.A.</td>
</tr>
<tr>
<td>dsš38k</td>
<td>spont.</td>
<td>Waddington, 38k</td>
<td>40%</td>
</tr>
<tr>
<td>dsš4k</td>
<td>unknown</td>
<td>Garcia-Bellido</td>
<td>N.D.</td>
</tr>
<tr>
<td>dsš55š</td>
<td>X-ray</td>
<td>Craymer</td>
<td>N.A.</td>
</tr>
<tr>
<td>dsšF31B</td>
<td>EMS</td>
<td>Postner</td>
<td>N.D.</td>
</tr>
<tr>
<td>dsšM6</td>
<td>EMS</td>
<td>Postner</td>
<td>N.D.</td>
</tr>
<tr>
<td>dsšM114</td>
<td>EMS</td>
<td>Postner</td>
<td>N.D.</td>
</tr>
<tr>
<td>dsšM16</td>
<td>EMS</td>
<td>Postner</td>
<td>N.D.</td>
</tr>
<tr>
<td>dsšM208</td>
<td>EMS</td>
<td>Postner</td>
<td>N.D.</td>
</tr>
<tr>
<td>dsšM212</td>
<td>EMS</td>
<td>Postner</td>
<td>40%</td>
</tr>
<tr>
<td>dsš2D60b</td>
<td>P-element</td>
<td>Cimborna</td>
<td>N.D.</td>
</tr>
<tr>
<td>Df(2L)ast</td>
<td>X-ray</td>
<td>Lewis</td>
<td>N.A.</td>
</tr>
<tr>
<td>dfp(2L)ast</td>
<td>X-ray</td>
<td>Lewis</td>
<td>N.A.</td>
</tr>
<tr>
<td>dsšD4c</td>
<td>X-ray</td>
<td>Sigmund (1978)</td>
<td>N.A.</td>
</tr>
</tbody>
</table>

A large number of alleles define the ds locus. The allele design-
"On Im(2LR)bw¹ also known as Pm balancer; cytology, Im(2LR)
"On Im(2LR)/SM5, also known as SM5 balancer, cytology, Im(2LR)
"On T(2;3)D¹, cytology, T(2;3)21D;70-71.

tic expression in the notum remains relatively low (Fig. 6F).

We also examined the patterns of lacZ expression in the dsšD60b
enhancer trap P-element insertion stock (Fig. 7). The patterns reveal that the enhancer detected by the
P-element insertion in the large first intron [Fig. 1] con-
trols expression both in embryos and larvae. However, while the patterns of lacZ and ds expression are virtually
indistinguishable in third-instar larvae, they differ si-
nificantly during embryogenesis [cf. Figs. 5 and 7]. For
example, lacZ is expressed in a weak anterior stripe dur-
ing cellularization [Fig. 7A] and, during gastrulation, in
the ventral furrow, along the cephalic furrow, and in the
procephalon [Fig. 7B] but fails to show the early pair-rule
pattern of ds stripes [Fig. 5A]. During germ-band elonga-
tion and the extended germ-band stage, lacZ fails to be
expressed in stripes or in the tracheal pits (Fig. 7C,D).
Only after dorsal closure, lacZ and ds expression are
nearly congruent except that lacZ is expressed ectopi-
cally in the eighth abdominal segment, but fails to be
expressed in the frontal sac (cf. Figs. 5F, I and 7F, G). The
differences in embryonic expression patterns between ds
and lacZ might result from an incompatibility of some of the
embryonic ds enhancer elements with the basal pro-
moter combined with lacZ [Li and Noll 1994].

Mutations in the ds gene lead to defects
in morphogenesis

The adult phenotype of ds mutants is consistent with the
high levels of ds transcript in imaginal discs. In all
known alleles (Table 1), defects are seen with 100% pen-
etrance in the legs, wings, and thorax. In contrast, eye
defects, apparent as rough patches, occur at a low fre-
cuency. On ds wings, the anterior cross-vein is displaced
distally, that is, closer to the posterior cross-vein (Fig.
8A, C, E, G), the legs are stubby with a reduced number of
tarsal joints in stronger alleles (Fig. 8B, D, F, H), and the
thorax is broadened (Waddington 1943). These adult phe-

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure4.png}
\caption{The ds¹ allele generates an altered form of the Ds protein. The ds¹ allele is an insertion of the 412 transposon
generating a deletion of 23 amino acids from the eleventh cad-
herin domain due to an altered splicing pattern. The genomic
DNA sequence surrounding the insertion in ds¹ of the 412 transposon
at the 3' end of exon 5 is shown. Above the DNA
sequence the corresponding amino acid sequence of the elev-
enth cadherin domain is derived. Only one of the three cryptic
splice donor sites of intron 5 that are observed in ds¹ is in-frame
with the wild-type splice donor site and hence results in the
deletion of the 23 shaded amino acids. The two other splice
donor sites of ds¹ generate truncated proteins because of prematu-
ture termination.}
\end{figure}
The *ds* transcript is expressed in ectodermally derived tissues of wild-type embryos. Whole-mount embryos, hybridized with a digoxigenin (DIG)-labeled *ds* cDNA probe (combined inserts of HC8, 119, and 143 cDNA clones in Fig. 1), are shown at stage 6 (A), early stage 10 (B), stage 11 (C), early (D) and late stage 14 (E), or stage 16 (F) as lateral (A–D), ventral (E), or dorsal view (F). Enlarged optical sections illustrate parts of embryos in G–H. (G) Dorsal view of the posterior thoracic and anterior abdominal segments (T2-A2) of a stage 15 embryo, demonstrating *ds* expression along both sides of the intersegmental grooves; (H) lateral view of thorax of stage 14 embryo, showing *ds* expression in leg disc primordia; (I) dorsal view of head of stage 16 embryo, illustrating *ds* expression in derivatives of the clypeolabral (lb, dpo, and epiphysis), labial (labial segment; (mir labial segment; (mdl mandibular segment; (mmd mandibular segment; (mx) maxillary segment; (pc) pole cells; (st) stomodeum; (T1) prothoracic segment.

Nototypes are more pronounced in stronger mutant alleles, with the addition of duplicated bristles on the notum and wings stiffly held out with broken and ectopic cross-veins. Several of the *ds* alleles listed in Table 1 have been examined in homozygous and transheterozygous combinations for the severity of their mutant phenotypes. Clearly, *ds* is the weakest known *ds* allele, as it is the only fully viable allele when combined with itself or any of the other alleles, including the two deficiencies *Df(2L)S2* and *Df(2L)astl*, and displays the weakest nototype [Fig. 8C–F]. All other allelic combinations exhibit various degrees of pupal lethality and eclose with a correspondingly reduced efficiency. In the strongest combinations, only a very low percentage (on the order of 1%) of extremely disfigured escapers emerges from the pupal cases. These adults have difficulty walking, cannot jump or fly, do not reproduce, and usually die within a few days after eclosion.

It is uncertain whether any of the known *ds* alleles are null alleles, although we expect the molecular lesions of *ds*^D4^, *ds*^33k^, and *ds*^55^ to generate a secreted, truncated Ds protein (Fig. 1) that might be without function. Combinations of *ds*^33k^ and *ds*^55^ with each other or with the two deficiencies eclose with the lowest frequency and show very strong deformations of thorax, legs, and wings. Because the P-element insertion *ds*^2D60B^ behaves very similarly, it might also be a null allele. We conclude that the strongest *ds* alleles are probably null alleles but are not completely lethal before eclosion. Additional strong alleles are *ds*^38k^ and *ds*^M213^ which, when combined with one of the two deficiencies, eclose only rarely (<5%). These are probably not null alleles, as they eclose with a much higher efficiency (~40%) in homozygous or transheterozygous combinations, although the escapers are highly deformed [Fig. 8G,H] and die within a few days after eclosion. Surprisingly, no embryonic phenotype has been observed even of the strongest allelic combinations.

The *ft* gene has been termed a tumor suppressor gene because recessive null alleles give rise to mutants with overgrown imaginal discs [Bryant et al. 1988; Mahoney et al. 1991]. Because of the similarity between the overall structures of the Ds and Ft proteins, the expression of their transcripts, their viable adult mutant phenotypes, and the lethal stage of their putative null alleles, we wondered whether strong *ds* alleles would also give rise to overgrowth phenotypes. To test this possibility, imaginal discs of homozygous or transheterozygous combinations of two strong *ds* alleles, *ds*^M213^/*ds*^M213^ and *ds*^M213^/*ds*^33k^, were compared in size with those from *CyO* and wild-type Canton-S larvae and examined for a possible overgrowth phenotype. All of the *ds* discs ex-
Dachsous cadherin controls morphogenesis

Figure 6. The ds transcript is expressed in specific regions of the larval brain and imaginal discs. Whole-mount late third-instar imaginal discs and larval brain were hybridized with the same DIG-labeled ds cDNA probe as used in Fig. 5. Expression of ds transcripts is shown in the supraesophageal ganglion or brain (A), and in imaginal discs of the eye–antenna (B), humerus (C), metathoracic leg (D), wing (E), and haltere (F). (ans) Antennal segments; (at) arista; (c) capitellum; (dhi) dorsal hinge; (fe) femur; (hu) humeral disc; (no) notum; (ol) optic lobe; (p) pedicel; (pl) pleura; (ps) prescutum; (s) scabellum; (spg) supraesophageal ganglion; (ta) tarsus; (th) thorax; (ti) tibia; (tr) tracheae; (vg) ventral ganglion; (wb) wing blade.

amined were within the size range observed in wild-type larvae and exhibited no overgrown imaginal discs. Therefore, in contrast to mutations in the ft gene, ds mutations appear to alter imaginal disc morphogenesis exclusively without affecting cell proliferation.

Discussion

Two Drosophila genes, ft [Mahoney et al. 1991] and ds, as shown here, belong to the cadherin superfamily and encode huge transmembrane proteins that contain large numbers of tandem cadherin domains (27 for the Ds and 34 for the Ft protein). Interestingly, both genes have been studied for decades because of the striking similarity in their mutant phenotypes observed in adult structures derived from imaginal discs [Mohr 1923; Stern and Bridges 1926], and because of their genetic interactions with each other [Mohr 1929] and with several other genes involved in the morphogenesis of imaginal discs [Waddington 1943].

Although the Ds and Ft proteins are quite similar, they differ in one important aspect. In contrast to Ds, Ft includes five EGF-like repeats and two laminin A G-domain-like repeats in the extracellular portion between the cadherin domains and the transmembrane segment [Mahoney et al. 1991; Patthy 1992]. This structural difference between the products of the two genes is paralleled by an important difference in the spectrum of their mutant phenotypes. Although viable mutations in both genes lead to similar defects in the morphogenesis of imaginal discs, strong mutations in ft, but not in ds, lead to tumor-like hyperplastic growth of imaginal discs [Bryant et al. 1988]. This overgrowth phenotype of lethal ft alleles is cell autonomous [Mahoney et al. 1991], indicating that the Ft protein acts as signal receptor in the control of cell proliferation. Thus, it is attractive to speculate that the EGF-like repeats and the laminin A G-domain-like repeats specific for the Ft protein are responsible for its tumor suppressor function by transmitting a signal regulating cell proliferation [Mahoney et al. 1991]. On the other hand, the protein domains common to Ft and Ds play similar roles in the morphogenesis of imaginal discs.

Do Ds and Ft interact with the cytoskeleton-like classic vertebrate cadherins?

Classic vertebrate cadherins mediate cell adhesion at adherens junctions. Adherens junctions play a central role in morphogenesis by regulating not only cell adhesion but also cellular polarity, changes in cell shape by interaction with the actin cytoskeleton, and signal transduction. Recently, a cadherin was isolated from Drosophila, DE-cadherin, whose structure and properties are very similar to those of classic vertebrate cadherins [Oda et al. 1994]. DE-cadherin is part of a membrane-associated complex that includes the cytoplasmic proteins Dα-catenin and Armadillo (Arm), the Drosophila homologs of
vertebrate α- and β-catenin (Oda et al. 1993; Peifer 1993). The highly polarized subcellular distribution of this complex to the lateral-apical interfaces of epithelial cells suggests that it forms intercellular adherens junctions (Peifer et al. 1993a, b; Oda et al. 1994) homologous to those of vertebrates (Takeichi 1991; Kemler 1993).

Ds and Ft differ from classical vertebrate cadherins and DE-cadherin mainly in the much larger number of extracellular cadherin domains (Fig. 3A). In addition, DE-cadherin and Ft have EGF- and laminin A G-domain-like repeats in their extracellular domain that are absent in Ds and classic cadherins. All three Drosophila cadherins, Ft, Ds, and DE-cadherin, have a cytoplasmic domain with significant sequence similarity to classic vertebrate cadherins in the region that is known to bind β-catenin (Nagafuchi and Takeichi 1989; Ozawa et al. 1990b), an actin-associated protein thought to anchor the classic cadherins at adherens junctions to the cytoskeleton (Kemler 1993). However, in contrast to vertebrate cadherins and DE-cadherin, this putative β-catenin-binding domain is split by a short and a long intervening region in Ds and Ft, respectively (Fig. 3). Despite this interruption in the β-catenin binding site, the cytoplasmic domain might still be able to link Ds and/or Ft to the actin cytoskeleton. It is therefore conceivable that Ds and Ft regulate morphogenesis and cell proliferation as components of adherens junctions in embryos or imaginal discs via catenins such as Armadillo and Dα-catenin. Although we could not demonstrate a genetic interaction between ds and arm in double heterozygotes (H. Clark, unpubl.), an interaction between Ds and Arm is not excluded because this test might not be sensitive enough. It is also possible that the cytoplasmic domains of Ds and Ft interact with yet another β-catenin-like Drosophila protein.

**Do Ds- or Ft-like cadherins also exist in vertebrates?**

It is not known whether vertebrates also have large Ft- or Ds-like members of the cadherin superfamily and whether they also have cadherins containing EGF-like and laminin A G-domain-like repeats. It will be of interest to determine whether vertebrates do have large cadherins like Ds and Ft and whether they too play major roles in tissue morphogenesis and growth control. It is possible that Ft and Ds define a new subfamily of the cadherin superfamily and that such large cadherins in both vertebrates and insects play a major role in tissue morphogenesis. Alternatively, if only insects like fruitflies possess such enormous cadherins that are used primarily for imaginal disc morphogenesis, as is the case for Ds and Ft, it will be interesting to learn what constraints of that particular developmental process in insects might have led to the evolution of such unusually large cadherins or, alternatively, what constraints in vertebrates might have suppressed their evolution.
Dachsous cadherin controls morphogenesis

Figure 8. ds mutations cause defects in adult structures derived from imaginal discs. Examples of wing [A,C,E,G] and leg [B,D,F,H] phenotypes of wild-type Canton-S flies (A,B) are compared with those of the homozygous weak ds1 allele (C,D), the strong ds38k allele (G,H), and their transheterozygous combination of intermediate strength (E,F). As the strength of the allelic combination increases, the distance between cross-veins (asterisks), the femur length (arrowheads), and the number of tarsal joints (arrows) are reduced.

A model for the interaction between ds and ft

Two types of recessive ft alleles are known. Viable ft alleles display a morphogenetic phenotype similar to that of ds, whereas lethal ft alleles produce hyperplastic overgrowth of imaginal discs. Gull is a dominant allele of ft with respect to the morphogenetic phenotype but recessive with respect to the hyperplastic overgrowth.
phenotype. Moreover, the dominant effect of Gull is antimo-
phomorphic or dominant negative (Mahoney et al. 1991) and,
hence, results from an interaction and/or competition of mutant Gull protein with wild-type Ft protein, leading to an inhibition of Ft function (for review, see Herskowitz 1987). The observation that this domi-
nant-negative effect of Gull is partially suppressed by one and
largely by two copies of ds1 (Mohr 1929) favors a model in
which both Ft and Gull are able to interact with Ds and
Ds1 proteins, possibly through heterophilic interaction
of their cadherin domains. It is also proposed that Ds
and Ft interact in a homophilic fashion, suggesting
the molecular model illustrated in Figure 9, which
is consistent with the genetics of ds and ft.

As argued above, Ft is proposed to mediate control of
cell proliferation through its extracellular EGF-like and/or
its laminin A G-domain-like repeats, which probably
interact as receptors with a membrane-bound or diffus-
able extracellular signal molecule. The Ds protein lacks
these domains and thus is only able to modify this con-
trol, without disrupting it, through its homophilic and
heterophilic interactions with itself and Ft. Because ho-
mozygous Gull larvae display the hyperplastic over-
growth phenotype of imaginal discs, Gull protein is un-
able to mediate the growth control signal and thus is
presumably mutant in those Ft domains required for this
control. Although the Gull mutation has been shown to be
cause by the insertion of a 412 transposable element
into the region encoding the thirty-third cadherin do-
main of ft (Mahoney et al. 1991), its effect on the Gull
protein is not known. In principle, the Gull product
could consist of three not mutually exclusive mutant
forms of the Ft protein: (1) a secreted, truncated protein
consisting of 32 cadherin domains; (2) a transmembrane
protein with extracellular EGF-like repeats, laminin A
G-domain-like repeats, and less than two cadherin do-
 mains [Mahoney et al. 1991]; and (3) a Ft protein from
which a portion surrounding the site of the 412 insertion
has been deleted because of the use of cryptic splice sites
similar to what has been observed here for the ds1 mu-
tation (Fig. 4). That a truncated, secreted Gull protein
causes the dominant mutation appears unlikely because
very similar truncated, secreted Ft proteins do not ex-
hbit a dominant effect (Mahoney et al. 1991). The two
remaining possible forms of Gull are both membrane
proteins, but only the third seems consistent with its
suppression by Ds1 and its lack of growth control be-
cause it retains most cadherin domains and is mutant for
the Fat-specific domains.

The dominant effect of Gull is explained by a hetero-
philic interaction of Gull with Ds that is stronger than
that of Ft with Ds and thus reduces the number of ho-
mophilically interacting Ds proteins, resulting in a dis-
turbed morphogenesis. The same effect is expected for
ds mutants, which explains their similar phenotype to that
of heterozygous Gull mutants. Conversely, the suppres-
sion of Gull by ds1 may be attributed to a decrease in
strength of Gull–Ds as compared with Gull–Ds inter-
action, which leads to an increase in homophilically in-
teracting Ds1 proteins and thus largely restores control

Figure 9. Model for role of Ds and Ft cadherins in control of
morphogenesis and cell proliferation in imaginal discs. This
model proposes that Ds and Ft mediate cell–cell adhesion by
homophoric and heterophilic interaction of their cadherin domains
and transmit signals regulating morphogenesis and cell prolif-
eration via their cytoplasmic domains to the cell interior and
nucleus. It is possible that the extent of homo- or heterophilic
association between cadherin domains can be modulated and is
not maximal as illustrated here. The morphogenetic signals
transmitted by Ds or Ft are not necessarily the same, although
they might cooperate. Only Ft, and not Ds, mediates signals
controlling cell proliferation, through its specific extracellular
EGF-like (EGF) and laminin A G-domain-like repeats (G) that
act as receptors. Both processes, control of cell proliferation and
morphogenesis, are intimately linked by coupled equilibria be-
 tween homophilic and heterophilic associations of the Ds and
Ft cadherins. For further details of this model, see text.

of morphogenesis. Thus, in this model, morphogenesis
and cell proliferation are regulated by homophilic inter-
actions of Ds and Ft proteins whose equilibria are cou-
ped by heterophilic interactions between the two cad-
herin-like proteins. Accordingly, recessive viable muta-
tions of ft, which show no loss of control of cell
proliferation, are expected to affect the function of the Ft
cadherin domains but not the Fat-specific domains and
thus to shift the equilibria between Ft–Ft, Ft–Ds, and
Ds–Ds interactions.

Materials and methods

General procedures

Standard procedures such as the isolation of genomic DNA, the
construction and screening of genomic libraries, chromosomal
walking, whole genome Southern analysis, in situ hybridization
to salivary gland chromosomes, isolation and Northern analysis
of poly(A)+ RNA, and PCR were carried out essentially as de-
Mapping of ds rearrangements and analysis of the ds mutation

The breakpoints of the inversions ds33k, associated with In(2LR)bw^53, and ds55, associated with the SM5 balancer chromosome, of the translocation T(2;3)D^5i, and of the deficiencies Df(2L)st and Df(2L)s2 were mapped with respect to the chromosomal walk in 21D1.2 by in situ hybridization to salivary gland chromosomes and whole genome Southern analysis. The ds mutation was characterized as RFLP by whole genome Southern analysis and isolated from a genomic library of homozygous ds flies. To examine the mRNA products derived from ds, ds^-cDNA was prepared by priming reverse transcription of poly(A)^+ RNA from embryos or late third-instar larvae with the primer 5'-GGTAAATGTTGGCCTGTGTC-3', located 0.46 kb downstream of the 412 insertion. These cDNAs were amplified by nested PCRs, using the additional primer 5'-CTGTGATTCTCAGCTCGGCTC-3', located 0.47 kb upstream of the 412 insertion, in the first PCR, and in the second PCR the primers 5'-GCCAGCGATCTGGACACGG-3' and 5'-CTGTGAATGTGTCCGAATCG-3', located 0.47 kb upstream and 0.05 kb downstream of the 412 insertion, respectively. The various PCR products were cloned and sequenced (cf. Fig. 4).

In situ hybridization to whole-mount embryos and discs and immunostaining of embryos

In situ hybridization to whole-mount embryos and imaginal discs with digoxigenin (DIG)-labeled probes (Tautz and Pfeifle 1989) or immunostaining for ß-galactosidase has been described previously [Li et al. 1993; Schnetzer et al. 1993].

Isolation and reversion of the P-element insertion ds2^{2D60b}

The P-element enhancer trap line 2D60b was isolated by Dan Cimbora. To separate the P element inserted in ds, 2D60b, from a second insertion, 2D60a, the 2D60a chromosome was recombined with al dp b pr c px sp, and ds2^{2D60b} b pr c px sp recombinants were recovered.

The 2D60b P-element was excised by constructing a fly stock ds2^{2D60b} b pr c px sp/Cyo; ry^506/TM2, ry and crossing these flies to Ki p^P(ry^-, Δ2-3), a source of transposase, and selecting progeny that subsequently lost the ry^+ marker carried by the P-element insertion ds2^{2D60b}. To isolate these progeny, the ds2^{2D60b} b pr c px sp + ; ry^506/Ki p^P(ry^-, Δ2-3) or ds2^{2D60b} b pr c px sp + ; TM2, ry/Ki p^P(ry^-, Δ2-3) male progeny were crossed individually to ry^506 or ry^10 ry/TM2, ry virgin females in 127 lines. Flies with rosy eyes could result from loss of the P-element 2D60b from the marked second chromosome or from two wild-type second chromosomes. To isolate only the P-element excisions, every ry^+ male from each of the 127 isolines was tested and the excision chromosome, if present, balanced by crossing individually with ds2^{M213}/Cyo, pr or In(2LR)bw^53, ds33k/Cyo, pr virgin virgins.

Drosophila strains

Most stocks were obtained from the Bloomington Stock Center. Marya Postner of Eric Wieschaus’ laboratory [Princeton University, NJ] kindly provided the ds alleles F31B, M56, M114, M116, M121, M208, and M213, which resulted from a screen for halo mutations, in a cn bw sp/Cyo background. T(2;3)D^5i was supplied by Michael Ashburner (Cambridge University, UK), and ds2^{M213}/Cyo, mwh h by Antonio Garcia-Bellido (Universidad Autónoma de Madrid, Spain).

Acknowledgments

We thank Zaida Traquina, Patrick Spielmann, and Natasha Delay for technical assistance and Fritz Ochsleinbein for the artwork. We are grateful to Paul Mahoney for help in the early phase of this project, to Dan Cimbora and Shige Sakonju for the ds enhancer trap line and the rescue plasmid, to Marya Postner in Eric Wieschaus’ laboratory for providing the ethylmethane sulfonate (EMS)-induced ds alleles, and to Luis Garcia-Alonso, Tom Elkins, and Michael Ashburner for genetics advice. We thank Hans Noll for comments on the manuscript. This work has been supported by a National Institutes of Health [NIH] Genetics Predoctoral Traineeship and NIH grant 5R37 HD21294 to H.F.C., by grant 31-26652.89 from the Swiss National Science Foundation to M.N., and by the Kanton Zurich. C.S.G. is an investigator with the Howard Hughes Medical Institute.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked ‘advertisement’ in accordance with 18 USC section 1734 solely to indicate this fact.

Note

The GenBank accession number for the ds cDNA sequence is L08811.

References


