The *Drosophila dysfusion* Basic Helix-Loop-Helix (bHLH)–PAS Gene Controls Tracheal Fusion and Levels of the Trachealess bHLH-PAS Protein

Lan Jiang and Stephen T. Crews*

Program in Molecular Biology and Biophysics, Department of Biochemistry and Biophysics, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-3280

The development of the mature insect trachea requires a complex series of cellular events, including tracheal cell specification, cell migration, tubule branching, and tubule fusion. Here we describe the identification of the *Drosophila melanogaster dysfusion* gene, which encodes a novel basic helix-loop-helix (bHLH)–PAS protein conserved between *Caenorhabditis elegans*, insects, and humans, and controls tracheal fusion events. The Dysfusion protein functions as a heterodimer with the Tango bHLH-PAS protein in vivo to form a putative DNA-binding complex. The *dysfusion* gene is expressed in a variety of embryonic cell types, including tracheal-fusion, leading-edge, foregut atrium cells, nervous system, hindgut, and anal pad cells. RNAi experiments indicate that *dysfusion* is required for dorsal branch, lateral trunk, and ganglionic branch fusion but not for fusion of the dorsal trunk. The *escargot* gene, which is also expressed in fusion cells and is required for tracheal fusion, precedes *dysfusion* expression. Analysis of *escargot* mutants indicates a complex pattern of *dysfusion* regulation, such that *dysfusion* expression is dependent on *escargot* in the dorsal and ganglionic branches but not the dorsal trunk. Early in tracheal development, the Trachealess bHLH-PAS protein is present at uniformly high levels in all tracheal cells, but since the levels of Dysfusion rise in wild-type fusion cells, the levels of Trachealess in fusion cells decline. The downregulation of Trachealess is dependent on *dysfusion* function. These results suggest the possibility that competitive interactions between basic helix-loop-helix–PAS proteins (Dysfusion, Trachealess, and possibly Similar) may be important for the proper development of the trachea.

The insect tracheal system consists of an intricately branched system of tubules that provide oxygen throughout the animal. The formation of the trachea consists of a series of developmental events, and its analysis provides an excellent model system for studying the morphogenesis of other branched structures, such as the vertebrate lung airways, circulatory system, kidney ducts, and excretory epithelia (9, 27, 38). The trachea is derived from an array of segmentally repeated clusters of precursor cells. After the tracheal precursor cells divide and invaginate, they extend branches, and the branches from neighboring segments fuse to form the mature tracheal tree. The fusion process is mediated by a distinct fusion cell residing on each branch (41). Branching and fusion are complex cellular processes and pose a number of developmental questions. How are fusion cells specified during tracheal development? What are the short-range and long-range factors that guide tracheal branches to their fusion partners? What is the nature of the adhesive and contact-guidance interactions that mediate fusion and allow the formation of adherens junctions that seal intercellular junctions? How is the cytoskeleton rearranged to allow the tracheal lumen to extend throughout the branch?

The tracheal primordia extend branches in six directions under the guidance of the *branchless* gene (44). All of these branches except the visceral branch will fuse with tracheal branches derived from other primordia. The dorsal trunk is formed by fusion of anterior and posterior branches from adjacent segments, as is the lateral trunk, which is formed by fusion of lateral trunk anterior and posterior branches. The dorsal branches travel over the dorsal side of the embryo and fuse along the dorsal midline to their partner from the identical hemisegment. The ganglionic branches migrate ventrally and join at the ventral midline, although only the three anteriormost branches fuse. There is a single fusion cell for each branch, and the fusion cells are characterized by patterns of gene expression distinct from other tracheal cells (41). The Escargot (Esg) zinc finger transcription factor is prominently expressed during tracheal development in fusion cells and no other tracheal cell (41, 42, 46). *esg* mutants show fusion cell defects in the lateral trunk, dorsal branch, and ganglionic branch, but dorsal trunk fusion is relatively normal. Examination of *esg* mutants indicates that expression of several, but not all, fusion cell-specific genes and markers is absent in the dorsal branch and the ganglionic branch (42), and there is an excess of branching and gene expression associated with branching. In addition, *esg* is required for *DE-cadherin* expression and the ability of fusion cells to form adhesive contacts and adherens junctions (46). The loss of *esg* in lateral trunk is the most extreme, resulting in death of the fusion cells. Important issues regarding the tracheal function of *esg* deal with the identity of genes regulated by *esg* and why *esg* is required for tracheal fusion in some branches but not others.

Three proteins that function prominently in tracheal devel-

* Corresponding author. Mailing address: Department of Biochemistry, CB#3280 Fordham Hall, The University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-3280. Phone: (919) 962-4380. Fax: (919) 962-8472. E-mail: steve_crews@unc.edu.

opment are the Tracheless (Trh), Similar (Sima), and Tango (Tgo) basic helix-loop-helix (bHLH)–PAS proteins. Trh and Tgo form a heterodimer that controls transcription and initial formation of the tracheal primordia along with the Drifter (Dfr) POU-homeobox coactivator (4, 16, 32, 43, 50). Sima and Tgo form a protein dimer that controls the transcriptional response to hypoxia (22, 29). Since low cellular oxygen conditions induce additional tracheal branching (18), Sima may autonomously or nonautonomously be required for terminal tracheal branching. Tgo is found in all embryonic cells. In the absence of a bHLH-PAS partner protein, Tgo is found in the cytoplasm, but in the presence of a partner protein, they dimerize, translocate into the nucleus, bind DNA, and activate transcription (49). Since there may exist multiple bHLH-PAS partners of Tgo in the same cell, it has been proposed that the function and levels of bHLH-PAS proteins may be regulated by competitive interactions (13, 52). In part for this reason, it is important to identify all *Drosophila melanogaster* bHLH-PAS proteins and determine where they are expressed.

In the present study, we describe a novel bHLH-PAS gene, *dysfusion* (*dys*), which is expressed in all tracheal fusion cells, as well as the epidermal leading edge cells and several other cell types. The *dys* gene is a member of a discrete subfamily of bHLH-PAS proteins conserved between nematodes, insects, and mammals. Tgo accumulates in nuclei of *dys*-expressing cells, suggesting that it is a partner of Dys in vivo. *dys*-RNAi experiments reveal tracheal fusion defects in the lateral trunk, dorsal branch, and ganglionic branch but not in the dorsal trunk. The *esg* gene is expressed in all fusion cells before *dys*, and *esg* expression is normal in *dys* RNAi-injected embryos. However, *esg* mutant embryos show an absence of *dys* expression in tracheal fusion cells in most branches, but not the dorsal trunk, further indicating branch-specific function of *esg*. The appearance of Dys in tracheal fusion cells coincides with a steep drop in Trh levels in fusion cells, and this reduction is *dys* dependent. This indicates that one function of *dys* in tracheal fusion cell development is to downregulate Trh protein levels. This provides the first in vivo evidence that bHLH-PAS proteins regulate levels of other bHLH-PAS proteins during development and possibly influence cell fate and morphogenetic decisions.

MATERIALS AND METHODS

***Drosophila* strains.** *Drosophila* strains utilized include *esg*[05730] (lethal P[*lacZ*] insertion in *esg*), *esg*[G66] (null P[*esg-lacZ*] mutation), *Df(2L)osp*[29] (*esg* deficiency), *puc*[A251.1F3] (P[*lacZ*] insertion in *puckered*), *trh*[1] (severe mutation), and *Df(3R)Espl3* (deficiency [96F1; 97B1] that removes *dys*). Homozygous mutant embryos were identified by balancing mutations over either CyO *Krüppel-Gal4* UAS-*GFP*, TM3 *Krüppel-Gal4* UAS-*GFP*, or CyO P[*ftz-lacZ*] (5), and selecting for embryos without *GFP* or *ftz-lacZ* expression.

Identification of *dys* gene and mRNA structure. The human hypoxia-inducible factor 1α (HIF-1α) PAS-1 region was screened versus *Drosophila* genomic sequence data (1) by using tBLASTN (2), and a novel bHLH-PAS-containing gene was identified at 96F on the third chromosome. *Drosophila dys* cDNA was cloned by using an reverse transcription-PCR (RT-PCR) strategy based on the genomic sequence. Initially, four pairs of primers predicted to be in *dys* coding sequences were utilized: (i) exons 2 and 3, S1 (ACGTGCGATCGACGATATGT) and A1 (CGCATTTATGAGATCGCGTC); (ii) exons 3 to 6, S2 (GATGCAAACAAA TCGACGAA) and A2 (ATCGTAAACGGAGTCACCGT); (iii) exons 5 to 8, S3 (GCTTCTTGATGCATGCTCACA) and A3 (TTGCGCCTCCGTCAGATTAT); and (iv) exons 8 to 10, S4 (ATAATCTGACGGAGGCGCAA) and A4 (TCTG GCTTAAGTGCTGCGAA). Total RNA from an overnight collection of *Drosophila* embryos was extracted by using an RNeasy Mini Kit (Qiagen). Poly(A)⁺

mRNA was extracted from total RNA by using Oligotex (Qiagen). First-strand cDNA was synthesized by using SuperScript reverse transcriptase (Gibco-BRL) and used as a template for PCRs. PCR amplification of primer pairs 1 to 4 resulted in the generation of *dys* cDNA fragments of 515, 388, 665, and 1,141 bp. The amplified fragments were cloned into pCR2.1-TOPO (Invitrogen) and sequenced at the UNC Automated Sequencing Facility.

5′-RACE. The 5′-most *dys* sequences were obtained by performing 5′ RACE (rapid amplification of cDNA ends) with a SMART RACE cDNA amplification kit (BD Biosciences). First-strand cDNA was synthesized from poly(A)⁺ RNA at 42°C for 1 h with PowerScript reverse transcriptase by using *dys* internal primer A3 (TTGCGCCTCCGTCAGATTAT). The SMART II A oligonucleotide (AA GCAGTGGTATCAACGCAGAGTACGCGGG) was included. This oligonucleotide anneals to the C-rich cDNA 3′ tail and serves as a primer for second-strand DNA synthesis. Primer S5 (TCGTCTGCGGCTGCTGCTGTTGC), derived from the *dys* sequence, was used to perform PCRs for 5′ RACE, along with a universal primer (CTAATACGACTCACTATAGGGCAAGCAGTGGT ATCAACGCAGAGT) that can anneal to the SMART II A oligonucleotide sequence. The amplified fragments were cloned and sequenced.

3′ RACE. The 3′-most *dys* sequences were obtained by performing 3′ RACE with a SMART RACE cDNA amplification kit. First-strand cDNA was synthesized from poly(A)⁺ RNA at 42°C for 1 h with reverse transcriptase by using the oligo(dT)-based 3′-RACE CDS primer A [AAGCAGTGGTATCAACGCAGA GTAC(T)₃₀N₋₁N, where N is A, C, G, or T and N₋₁ is A, C, or G]. The 3′ RACE forward primer (GATGTCAGATCCTTTGGTACCGGTCACGTA) was derived from the *dys* sequence and used in PCRs along with the universal primer that can anneal to primer A. All amplified fragments were cloned and sequenced.

Sequence comparisons and phylogenetic analysis. The Dys bHLH sequence was aligned to bHLH sequences from representatives of all subfamilies of bHLH-PAS proteins by using CLUSTALX (48). The alignment was then used to create a phylogenetic tree by using the neighbor-joining method (40) and PAUP*4.0 software (45). Clades were assessed by bootstrap analysis of 1,000 repetitions. The tree was visualized with TreeView (33). The Dys sequence from *Anopheles gambiae* was determined by searching the *Anopheles* genome sequence (15) for Dys-related sequences. Significant homology was observed for all exons except exons 1, 2, and 4. The *Caenorhabditis elegans* cDNA clone sequence of C15C8.2 was provided by JoAnne Powell-Coffman (Iowa State University), and the human NXF (H-NXF) sequence was provided by Cam Patterson (University of North Carolina at Chapel Hill).

Dys antibody production. Polyclonal antibodies against Dys were generated by using a Dys fusion protein containing an N-terminal His₆ tag. A *dys* cDNA fragment corresponding to amino acids (aa) 584 to 803 was cloned into the *Xho*I site of pET-15b (Novagen). After transformation into *Escherichia coli* BL21(DE3), cells were grown and His-Dys fusion protein synthesis induced by the addition of IPTG (isopropyl-β-D-thiogalactopyranoside). Inclusion bodies were prepared, solubilized in 10% sodium dodecyl sulfate (SDS), dialyzed in 0.05% SDS–1 mM phenylmethylsulfonyl fluoride–1× phosphate-buffered saline, and purified by using Talon metal affinity resins (Clontech). The eluted His-Dys protein was then dialyzed in 0.01% SDS–1× phosphate-buffered saline. Two rats and two rabbits were immunized by subcutaneous injection of protein (Pocono Rabbit Farm & Laboratory).

***Drosophila* immunostaining.** Whole-mount embryos were immunostained as described previously (34). The following primary antibodies and dilutions were used for immunostaining: rat anti-Dys (1:200), rabbit anti-Dys at 1:800, mouse monoclonal antibody (MAb) 2A12 at 1:10 (24), MAb anti-Tgo at 1:1 (49), MAb anti-β-galactosidase (anti-β-Gal) at 1:1,000, rabbit anti-β-Gal (Cappell) at 1:2,000, and rat anti-Trh at 1:200 (49). The following secondary antibodies were used: fluorescein isothiocyanate-labeled anti-mouse and anti-rat immunoglobulin G (IgG; Molecular Probes), Texas red-labeled anti-rabbit and anti-mouse IgG (Jackson Laboratories), Alexafluor 594-labeled anti-mouse IgM, Cy5-labeled anti-mouse IgG (Molecular Probes), alkaline phosphatase (AP)-labeled anti-rat IgG and IgM, and biotin-SP-labeled anti-mouse IgG (Jackson Laboratories) used with streptavidin-horseradish peroxidase. All secondary antibodies were used at 1:200 except for Alexafluor 594-labeled anti-mouse IgM, which was used at 1:1,000. Nonfluorescence staining was developed by using either AP color development buffer (Bio-Rad) or diaminobenzidine with horseradish peroxidase. Photomicrographs were taken with either Zeiss LSM-510 confocal or Axiophot microscopes.

***Drosophila* in situ hybridization.** *Drosophila* whole-mount embryos were hybridized in situ to a *Drosophila dys* antisense RNA probe. The probe was 1.1 kb and was generated by PCR amplification of a *dys* cDNA clone. The primers A4 and S4 were used, and the region amplified included exons 8 to 10. The PCR product was cloned into PCRII-TOPO, and a digoxigenin (DIG)-labeled RNA

probe was synthesized by using SP6 RNA polymerase (Promega) at 37°C for 2 h. The probe was added to the hybridization buffer (50% deionized formamide; 5× SSC [1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate]; 100 μg of sonicated, denatured salmon sperm DNA/ml; 100 μg of *E. coli* tRNA/ml; 50 μg of heparin/ml; 0.1% Tween 20; pH 4.5). Hybridization was carried out for 16 h at 70°C. After a washing step, the presence of *dys* mRNA was detected by incubating embryos with AP-conjugated anti-DIG antibody (1:1,000) and reaction with nitroblue tetrazolium plus X-phosphate. Embryos were mounted in 70% glycerol and examined under bright-field optics.

RNAi. Double-stranded RNA (dsRNA) for RNA interference (RNAi) experiments was prepared for *dys* and for *green fluorescent protein* (*GFP*) by using modifications of existing protocols (20, 28). The *GFP*-dsRNA acted as a negative control. The templates used for synthesis of dsRNA were the products of two-step PCR. The first step amplified a fragment of cDNA that contains a partial T7 RNA polymerase promoter site. The second PCR step amplified the fragment and completed the addition of a T7 RNA polymerase site adjacent to both ends of the cDNA sequence. The first PCR used to amplify *dys* cDNA utilized the following primer pair: 5′-CGACTCACTATAGGGCGACTCCACGCAACAACCTG-3′ and 5′-CGACTCACTATAGGGTCTGGCTTAAGTGCTGCGAA-3′ (the T7 sequences are underlined). This fragment includes 1,027 bp (residues 1399 to 2426; exons 6 to 10) of the *dys* sequence. The *GFP*-dsRNA template was generated by amplifying a 729-bp fragment of pNEGFPX.1 that contains the *GFP* gene. The primer pair used was 5′-CGACTCACTATAGGGATGGTGAGCAAGGGCGAGGAG-3′ and 5′-CGACTCACTATAGGGGTACAGCTCGTCCATGCCGAG-3′. The second step of the PCR used the T7 promoter sequence (ATAGAATTCTCTAGAAGCTTAATACGACTCACTATAGGG) as a primer. This resulted in an amplified fragment with T7 RNA polymerase sites at both ends, allowing the simultaneous synthesis of sense and antisense RNA. DNA templates were purified by using QIAquick PCR purification kit (Qiagen). dsRNAs were synthesized by using T7 RNA polymerase and DNA templates were removed with RNase-free DNase. After phenol-chloroform extraction and ethanol precipitation, the dsRNAs were dissolved in Tris-EDTA (TE) at a final concentration of 5 μM. Analysis of dsRNAs by agarose gel electrophoresis indicated that bands of the expected sizes were present.

Injection needles were baked to remove RNase for injection. Embryos from w^{1118} or esg^{05730} flies were collected over a 60-min period at 25°C and attached to a coverslip with double-stick tape. They were covered in no. 700 halocarbon oil and dsRNA solution injected into the ventral and/or posterior side of the embryo along ca. 30 to 50% of the egg length by using a Picospritzer III Picopump. The injected embryos were incubated at 18°C for 26 h and at stages 15 and 16 were collected, dechorionated, fixed, devitellinized, and immunostained. Other embryos were allowed to hatch into first-instar larvae and then transferred to vials. After further incubation at 25°C, the tracheal phenotypes of second-instar larvae were examined by bright-field microscopy.

***dys* viability assays.** *dys*-dsRNA and *GFP*-dsRNA-injected larvae were transferred to vials with fly food at a density of 40 larvae/vial and then incubated at 25°C. The number of pupae and emerging adults were counted for each vial.
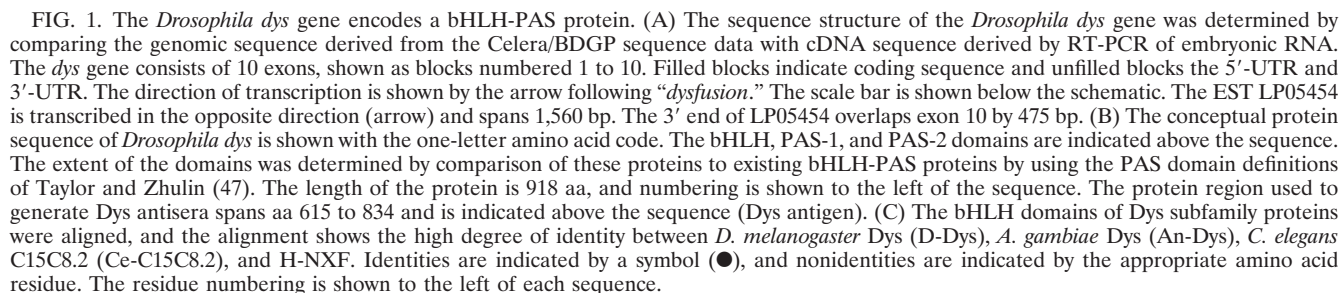
RESULTS

Identification and sequence of the *Drosophila dys* gene. Bioinformatic searches for additional *Drosophila* bHLH-PAS genes utilized bHLH and PAS sequences screened against the genomic sequence generated by the Celera/Berkeley *Drosophila* Genome Project (BDGP) (1). When a human HIF-1α PAS-1 region was screened against *Drosophila* genomic DNA by using the tBLASTN algorithm (3), a novel bHLH-PAS gene was identified in the BAC clones BACR13F13 and BACR08G22, which map to a cytological position, 96F-97A, on chromosome 3. Later, the gene was also identified by the BDGP-Celera Consortium by using the Genie algorithm (37) and in global searches for bHLH genes (23, 35). Genie initially predicted four distinct genes in the region, three of which contain multiple predicted exons: CG14554 (one exon), CG12561 (three exons), CG14553 (three exons), and CG14552 (four exons). Using RT-PCR and DNA sequence analysis, we showed that these four predicted genes all constitute a single gene, *dys*, now

referred to as CG32474 (Fig. 1A). *dys* maps to 96F9-10 according to FlyBase (11).

Preliminary experiments indicated that *dys* is a rare embryonic mRNA, an observation consistent with its absence from the BDGP embryonic expressed sequence tag (EST) collection of 8,809 distinct mRNAs. RNA was isolated from *Drosophila* embryos, and RT-PCR was carried out with primers designed from genomic sequences. The 5′ and 3′ ends of the mRNA were identified by RACE. These experiments provide the sequence of the *dys* mRNA. The genomic sequence comprises 10 exons and spans 21,697 bp (Fig. 1A). The 5′-untranslated region (UTR) determined by RACE is predicted to be 521 nucleotides (nt), and the 3′-UTR is 500 nt. Of particular interest, the 3′-UTR of *dys*, which resides in exon 10, overlaps by 475 bp the sequence of EST LP05454 (39), which is transcribed in the other direction (Fig. 1A).

The coding sequence of *dys* (Fig. 1B) shows a clear relationship to other bHLH-PAS proteins (8, 47). It has a bHLH domain near the N terminus, followed by PAS-1 and PAS-2 domains. The C-terminal regions of bHLH-PAS proteins generally contain transcriptional activation domains (see, for example, reference 12), although the protein sequences are poorly conserved. Consistent with that observation, the C-terminal region of Dys shows little homology with other proteins but contains histidine-rich and proline-rich regions similar to the C terminus of other bHLH-PAS proteins (8). The region N terminal to the bHLH domain encoded by exons 1 and 2 is 152 aa, which is uncharacteristically long for bHLH-PAS proteins. This region contains a large number of glutamine residues, suggesting that at it may act as an N-terminal transcriptional activation domain. There is also an unusually long spacer region in the PAS-1 domain that contains a large number of serines and repeats of Gly-Gly-Ala. The length of the predicted protein is 918 aa with a predicted molecular mass of 102 kDa.

Dys-related proteins constitute a discrete subset of evolutionarily conserved bHLH-PAS proteins. Sequence comparisons of all existing bHLH-PAS proteins to Dys reveal that it is a member of a discrete subfamily of bHLH-PAS proteins (Fig. 1C and 2). The complete sequences of representative bHLH-PAS proteins were compared by CLUSTALX and displayed by a phylogenetic tree generated by the neighbor-joining method algorithm (40). *C. elegans* C15C8.2 (7) (J. Powell-Coffman, unpublished data), *Drosophila* Dys (D-Dys), *A. gambiae* Dys (An-Dys) (15), and H-NXF (6, 30) (C. Patterson, unpublished data) cluster as a discrete branch of the phylogenetic tree consisting of representative bHLH-PAS proteins (Fig. 2). This result indicates that a *dys* ancestral gene existed before the divergence of nematodes, insects, and vertebrates. Quantitatively, the bHLH region of D-Dys shows the following percent identities to the other Dys subfamily proteins: An-Dys (96%), H-NXF (58%), and C15C8.2 (57%) (Fig. 1C). The percent identity was only 30% within the bHLH region to the next closest bHLH-PAS protein, *Drosophila* Spineless (Ss). Similar relationships were observed by comparing PAS-1 domains and PAS-2 domains, although the percent identity was less than for the bHLH region, as generally observed with bHLH-PAS protein comparisons. The PAS-1 domain of Dys is characterized by a long insertion containing a number of Gly-Gly-Ala repeats. This insertion is absent in C15C8.2 and H-NXF, but

A

1   2   3   4   5   6   7 8 9 10

1 kb   *dysfusion* →   LP05454

B

```
Dys     (1)  MPNAIGASSITKWISQIESFEQRELVTNICKRAIDDMYADNKRGLDKVTAATPPTHALQQQQQQQHPHQQQQHLQHVLQL

                                                                          |-------
Dys    (81)  QQQQQQHLQQQQHHLQHQMHTHHNFQQHETSPTVSQQQHMQQQQPQTTQQQHATQQQQQQQQQTAGMFTRFDANKSTKGA

             ---- basic-helix-loop-helix ----------------|                          |-
Dys   (161)  SKMRRDLINAEIANLRDLLPLPQSTRQRLSQLQLMALVCVYVRKANYFQQVFKRHNAMHHLHHQTATPNIDFSKALSGFL

             ----------------------------------- PAS-1 -----------------------------------|
Dys   (241)  MMLTQNGKLLYISDNAAEYLGHSMEDLLIHGDSVYDIIDKQDHANIQAELNRNVPPQPGGSSQASSGGAGGAGGAGDSTS

             --------- PAS-1 -------------|                          |---- PAS-2 ----
Dys   (321)  SMGSAAGAGGAGGASSLEGEHRNFLCRMNVSRNARRQNRFGDQKVVLVQGHYLSFLPLCSRNEPVFLATCTPIAMPETRE

             ------------------------------------ PAS-2 ------------------------------------
Dys   (401)  CVVQGATNVFTTIHSMDMKIAHIDKNGEFHLGYDKSTLQGTSWYNLIHSDNLTEAQKKHSLIIQSEQDRSCILLVRMQRS

             -------------------------------|
Dys   (481)  SGEYTWVHVVLQVRDSPDSTQQQVIVCTNQVLNDREASIMLANSWLYHYYTVQSKIQFGNPLDSGIRVPPGTPSSGYYSP

                                                          |------ Dys antigen ------
Dys   (561)  HSSHPPATPGGATPNLGIASSFGPHHHPHHSHHPHHHHHHHHPATAYHHHPHMGTYGGVYHAKAAIELKEDQPLFCFQEP

             ------------------------------------------------------------------------------
Dys   (641)  VDYSQVNSHVNPPNATSTPNPPISKSSTPPPPKKRARKLEPLYLHAERSPAAAITPEPDCYAIHTHPAAAYATSSVVSSD

             ------------------------------ Dys antigen ---------------------------------
Dys   (721)  ASYIYATIVPTRPRLPNKSLPPDPADFIEQWNPSPPWSESAQKASLDSFGSSHELSPCITTTPPTPTSATPHQSGLGTTQ

             -------------------------------|
Dys   (801)  TIGPAFSFEWMSDPLVPVTYQQWPPTGDHHQHLSQHLSQSQNHQNPLQQQQNEQLLLGQQPHLHQQHQHLALHHGRGEHS

Dys   (881)  LEGESQGVVVVPPIPISIPIPTATGHGDAGEPVEDRG
```

C

```
                    Basic          Helix 1        Loop        Helix 2

D-Dys    (153)  ANKSTKGASKMRR  DLINAEIANLRDLLP  LPQSTRQRLS  QLQVMALVCVYVRKA
An-Dys     (3)  •••••••••L••   ••••S••••••••••  ••••••••••  •••••••••••••••
Ce-C15C8.2 (77) QQR••R••••Q••  •Q••V••QK••••••  •SDLIKD••F  •••V•S•G•IFI••H
H-NXF      (1)  MYR••••••A••   •Q•••••R••KE•••  •AEADKV•••  Y•HI•S•A•I•T••G
```

FIG. 1. The *Drosophila dys* gene encodes a bHLH-PAS protein. (A) The sequence structure of the *Drosophila dys* gene was determined by comparing the genomic sequence derived from the Celera/BDGP sequence data with cDNA sequence derived by RT-PCR of embryonic RNA. The *dys* gene consists of 10 exons, shown as blocks numbered 1 to 10. Filled blocks indicate coding sequence and unfilled blocks the 5′-UTR and 3′-UTR. The direction of transcription is shown by the arrow following "*dysfusion*." The scale bar is shown below the schematic. The EST LP05454 is transcribed in the opposite direction (arrow) and spans 1,560 bp. The 3′ end of LP05454 overlaps exon 10 by 475 bp. (B) The conceptual protein sequence of *Drosophila dys* is shown with the one-letter amino acid code. The bHLH, PAS-1, and PAS-2 domains are indicated above the sequence. The extent of the domains was determined by comparison of these proteins to existing bHLH-PAS proteins by using the PAS domain definitions of Taylor and Zhulin (47). The length of the protein is 918 aa, and numbering is shown to the left of the sequence. The protein region used to generate Dys antisera spans aa 615 to 834 and is indicated above the sequence (Dys antigen). (C) The bHLH domains of Dys subfamily proteins were aligned, and the alignment shows the high degree of identity between *D. melanogaster* Dys (D-Dys), *A. gambiae* Dys (An-Dys), *C. elegans* C15C8.2 (Ce-C15C8.2), and H-NXF. Identities are indicated by a symbol (●), and nonidentities are indicated by the appropriate amino acid residue. The residue numbering is shown to the left of each sequence.

An-Dys has an even larger insertion with a different sequence content consisting of stretches of His, Ser, and Gly. In contrast, the PAS-2 domain is similar in length and colinear between the subfamily members. Sequences N-terminal to the bHLH domain, the spacer between bHLH and PAS-1 domains, and the regions C-terminal to the PAS-2 domain are poorly conserved among subfamily members, whereas the spacer between PAS-1 and PAS-2 shows significant sequence identity.

***dys* is expressed in diverse embryonic cell types.** Antisera were prepared against bacterially synthesized Dys protein derived from the unconserved C-terminal region of the protein, and whole-mount embryos were immunostained. Dys protein was found in a number of embryonic cell types (Fig. 3A to C), and these sites of expression were identical to those observed for *dys* RNA, as determined by *dys* in situ hybridization (Fig. 3D and E) (32). The specificity of Dys antisera was further
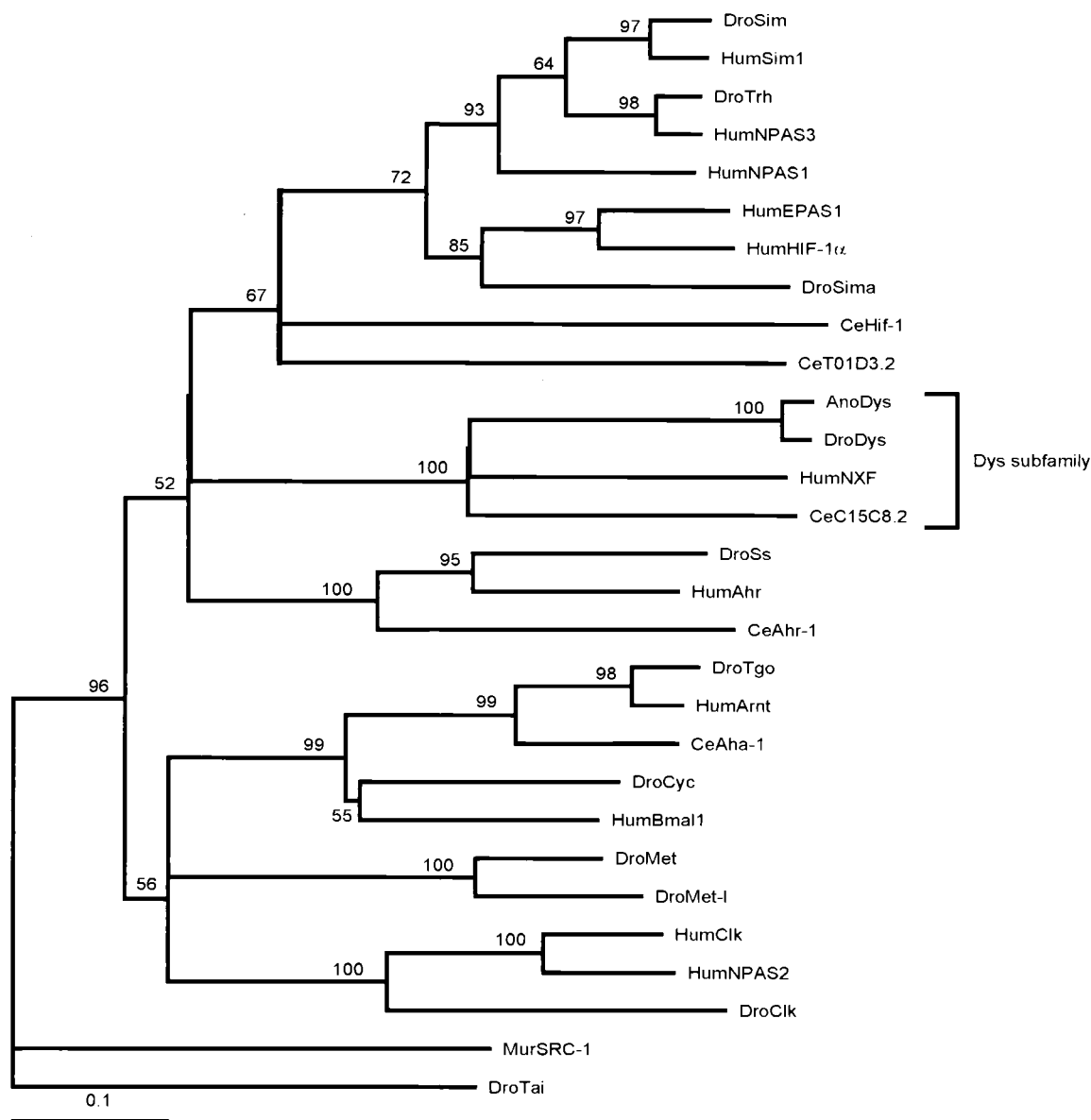
FIG. 2. Dys belongs to a distinct subfamily of bHLH-PAS proteins. Representatives of all families of bHLH-PAS proteins were compared by using the bHLH domain protein sequences and aligned by using CLUSTALX. The alignment is represented by a phylogram generated by using the neighbor-joining method. The tree demonstrates that DroDys, AnoDys, CeC15C8.2, and H-NXF (HumNXF) represent a distinct, evolutionarily conserved subfamily of bHLH-PAS proteins (bracket). The scale represents the fraction of nonidentical amino acids residues along each branch. Numbers along each branch are the bootstrap confidence limits with 1,000 repetitions. The identities of the other proteins can be found in Taylor and Zhulin (47). The species designation precedes each protein acronym (Ano, *A. gambiae*; Ce, *C. elegans*; Dro, *Drosophila*; Mur, murine; Hum, human).

demonstrated by immunostaining embryos derived from *Df (3R)Espl3* flies (Fig. 3F). The breakpoints of the *Df(3R)Espl3* deficiency are 96F1 and 97B1 and are predicted to delete the *dys* gene, which resides at 96F9-10. Embryos homozygous mutant for *Df(3R)Espl3* showed no staining with the Dys antisera. Additionally, in experiments described below (see Fig. 8 and 9), injection of *dys* RNAi into embryos abolished Dys immunoreactivity. The absence of Dys from the mutant embryos was not due completely to the loss of cells that are Dys positive in wild-type embryos, since the use of appropriate markers indi-

cated their presence (Fig. 9). These experiments indicate that the Dys antisera specifically recognizes Dys protein.

Staining with Dys antisera is highly restricted temporally and spatially. Little, if any, expression is observed before stage 12, but at stage 12 expression begins in many of the Dys-positive embryonic cell types. At stage 14, Dys protein is observed in (i) precursors to the foregut atrium, (ii) a subset of nervous system cells that are either part of the medial brain or the frontal ganglion, (iii) four rows of tracheal fusion cells, (iv) cells at the position of the epidermal leading edge, (v) the hindgut, and
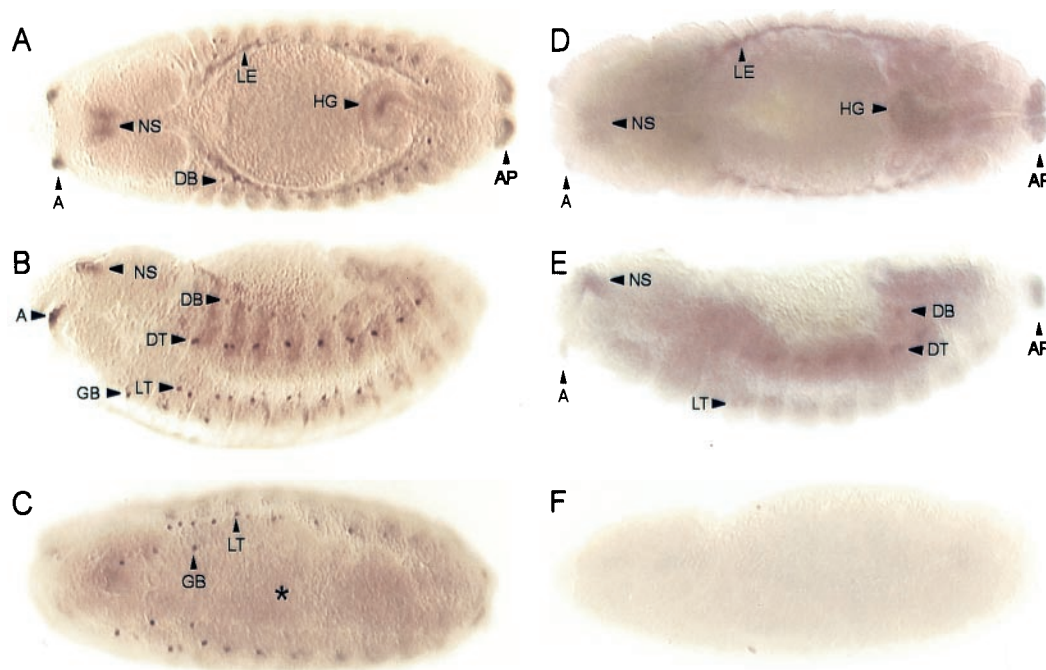
FIG. 3. The *dys* gene is expressed in multiple cell types. Dys localization was analyzed by immunostaining whole-mount embryos with α-Dys antisera, followed by AP histochemistry. Visualization is by Nomarski optics. (A to C) The embryo is at stage 14; anterior is to the left. (A) Dorsal view. Dys protein is present in anterior cells corresponding to foregut atrial precursors (arrowhead A), nervous system cells in part of either the medial brain or the frontal ganglion (arrowhead NS), the leading edge of the dorsal epidermis (arrowhead LE), tracheal fusion cells of the dorsal branch (arrowhead DB), hindgut (arrowhead HG), and anal pad (arrowhead AP). (B) Sagittal view illustrating tracheal fusion cell staining of dorsal branch (arrowhead DB), dorsal trunk (arrowhead DT), lateral trunk (arrowhead LT), and ganglionic branch (arrowhead GB). Atrial cells and nervous system staining are indicated as in panel A. (C) Ventral view showing *dys* expression in the ganglionic branch fusion cells (arrowhead GB) that later fuse to form the ventral anastomoses, and lateral trunk fusion cells (arrowhead LT). The ventral midline is indicated by an asterisk. (D) Dorsal view of a stage 13 embryo hybridized in situ with *dys* antisense DIG RNA probe. Hybridization (arrowheads) was detected in the same cell types as observed in panel A for Dys antibody staining. (E) Sagittal view of a stage 13 embryo hybridized to a *dys* antisense probe showing hybridization to the same cell types that stained with anti-Dys in panel B. (F) Sagittal view of a stage 14 *Df(3R)Espl3* homozygous mutant embryo stained with α-Dys. *Df(3R)Espl3* is deleted of the *dys* gene. There is a complete absence of α-Dys staining in all cell types, which confirms the specificity of the α-Dys staining. The parental strain contained a *Kr-Gal4* UAS-*GFP* balancer chromosome, and *Df(3R)Espl3* homozygous mutants were selected based on their absence of *GFP* expression.

(vi) the anal pad (Fig. 3A to C). The frontal ganglion consists of cells of the stomagastric nervous system located between the brain and roof of the pharynx (14). Expression remains in all of these cell types through the remainder of embryonic development. All Dys-positive cells showed nuclear localization (Fig. 4 to 6). The leading-edge and tracheal cells are conclusively identified in the present study, and more-detailed analyses of the other Dys-positive cells will be provided elsewhere.

dys is expressed in the leading edge. During the process of dorsal closure, the dorsalmost rows of cells on each side are referred to as the leading edge. These cells (i) are required for the process of dorsal closure, (ii) respond to cellular signals regulating closure, and (iii) form the dorsal zipper after closure (21). These cells are characterized by changes in cell shape during the closure process (17), and they interdigitate and adhere to each other after meeting at the dorsal midline. *dys* was examined for expression in the leading-edge cells by examining for colocalization with *Puckered* (*Puc*). The *puc* gene is expressed in the leading edge, and *puc*[A251.1F3], an enhancer trap line with a P[*lacZ*] insertion in the *puc* gene expresses *lacZ* in the leading edge cells (25). Double staining embryos with anti-Dys and anti-β-Gal revealed colocalization of the two proteins (Fig. 4).

dys is expressed in tracheal fusion cells. Dys protein was observed in four distinct longitudinal, segmentally repeated sets of cells along the ectoderm (Fig. 3A to C). The locations of these cells resembles the pattern of genes expressed exclusively in tracheal fusion cells (41). Embryos were double stained with MAb 2A12, which stains the tracheal lumen, and anti-Dys to determine the relationship between Dys-positive cells and the trachea (Fig. 5). Before fusion, each branch contained a single Dys-positive cell, whereas after fusion there were two Dys-positive cells at the fusion site of each branch. This was observed at all sites of tracheal fusion, including the dorsal trunk, lateral trunk, dorsal branch, and ganglionic branch, but in no other tracheal cells. Only the three most anterior (G0, G1, and G2) of the 10 ganglionic branches fuse at the ventral midline. These three branches possess Dys-positive cells, but the others do not. Confirmation that the Dys-positive cells are tracheal fusion cells was shown by double staining an *escargot-lacZ* (*esg-lacZ*) enhancer trap line with anti-β-Gal and anti-Dys (Fig. 6). Colocalization was observed in all Esg-positive tracheal fusion cells.

dys expression precedes tracheal fusion. Time course experiments were carried-out by double staining P[*esg-lacZ*] embryos with anti-Dys, anti-β-Gal, and anti-Trh. Anti-Trh stain-
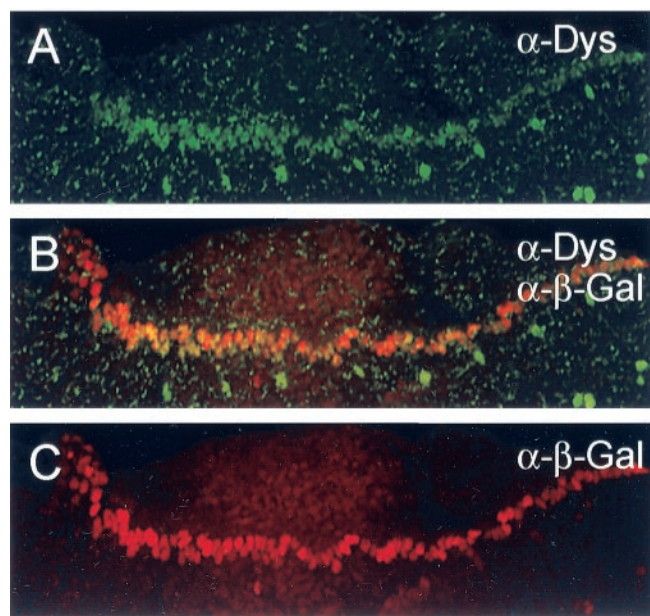
FIG. 4. *dys* is expressed in the leading-edge cells. A stage 14 *puc*[A251.1F3] embryo that expresses P[*lacZ*] in the leading-edge cells is shown. The embryo was double stained with anti-Dys (A; green) and anti-β-Gal (C; red). A merged image is shown in panel B. Anterior is to the left and dorsal is at the top. Colocalization of both Dys and β-Gal was observed in leading-edge cells.

ing visualizes tracheal development at all stages of development (Fig. 6). Primary branches originate from tracheal precursor cells at stage 11. At this time, Dys protein is undetectable. At stage 12, before the tracheae fuse in the dorsal trunk, *dys* expression is first detected. It remains on at high levels throughout the remainder of embryonic development (to stage 17) in the dorsal trunk. Lateral trunk cells fuse later (stage 15) than the dorsal trunk, and *dys* expression correspondingly also appears later in the lateral trunk (stage 13) than in the dorsal trunk (stage 12) (Fig. 6). However, *dys* expression in the lateral trunk still precedes its fusion. The dorsal branches and ganglionic branches fuse later (stages 16 to 17) than the lateral trunk and dorsal trunk (data not shown), and *dys* expression also appears later (stage 14) than in the dorsal trunk and lateral trunk. Thus, Dys protein appears before fusion of each branch, and its appearance in each branch type correlates with the timing of branch fusion.

Dys colocalizes with nuclear Tgo in vivo. Since Tgo is nuclear only in the presence of a partner bHLH-PAS protein, if Dys and Tgo dimerize in vivo, Tgo should be found in nuclei of Dys-positive cells. Embryos were double stained for Dys and Tgo (Fig. 7). Most relevant are the leading-edge cells, since no other bHLH-PAS proteins are known to be expressed in these cells. Strong Tgo nuclear localization was observed that overlaps exactly with Dys (Fig. 7A and B). Tracheal fusion cells were also examined for nuclear Tgo levels. Examination of fusion cells of all four types of branches indicated high levels of nuclear Tgo accumulation colocalizing with Dys nuclear localization (Fig. 7C to H). These results are consistent with Tgo being a partner for Dys in vivo, although the presence of Trh in tracheal fusion cells could also contribute to Tgo nuclear

localization (but see below). Of particular note is the finding that, qualitatively, the levels of nuclear Tgo appear constant in fusion cells throughout development, both before and after the appearance of Dys.

Injection of *dys*-dsRNA results in tracheal fusion defects. The relationship between *dys* expression and tracheal fusion branches suggests a role of *dys* in tracheal fusion events. Mutants of *dys* are not available, so *dys*-RNAi was utilized to test function. *dys*-dsRNA was injected into embryos, and tracheal development was assessed. Negative controls included embryos injected with TE and *GFP*-dsRNA. Staining of *dys*-dsRNA-injected embryos with anti-Dys indicated that Dys protein was undetectable (Fig. 8A and B and 9). An indication that *dys* is required for normal development is demonstrated by the increased lethality of *dys*-dsRNA-injected embryos compared to those injected with negative controls. 98% of the embryos ($n = 200$) injected with *dys*-dsRNA died before adulthood compared to 48% injected with *GFP*-dsRNA ($n = 200$). Both samples had similar levels of embryos hatching into larvae: 66% of *dys*-dsRNA injected embryos hatched and 62% injected with *GFP*-dsRNA hatched. Thus, *dys* is probably not required for embryonic viability. The *dys* ds-RNA-injected embryos that survived embryonic development died as second- and third-instar larvae.

Since loss of *dys* function is lethal and *dys* is prominently expressed in tracheal fusion cells, we examined *dys*-dsRNA-injected embryos for defects in tracheal fusion. Injected embryos were analyzed by (i) immunostaining with MAb 2A12, which stains the tracheal lumen and outlines tracheal branches,
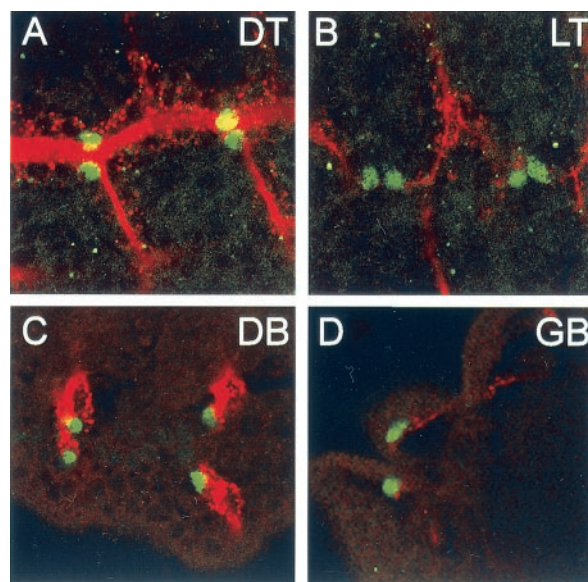

FIG. 5. *dys* is expressed at sites of tracheal fusion. Whole-mount stage 15 and 16 embryos were stained with anti-Dys (green) and MAb 2A12 (red), which stains the tracheal lumen. Two segments are shown in panels A, B, and C; one segment is shown in panel D. Two Dys-positive cells can be seen at sites of tracheal fusion (panel A, panel B, and one segment of panel C), and one cell stained in each unfused branch at sites that will soon undergo fusion (panel D and one segment of panel C). Cells stained in all four types of tracheal fusion branches, including the dorsal trunk (DT) (A), the lateral trunk (LT) (B), the dorsal branch (DB) (C), and the ganglionic branch (GB) (D).
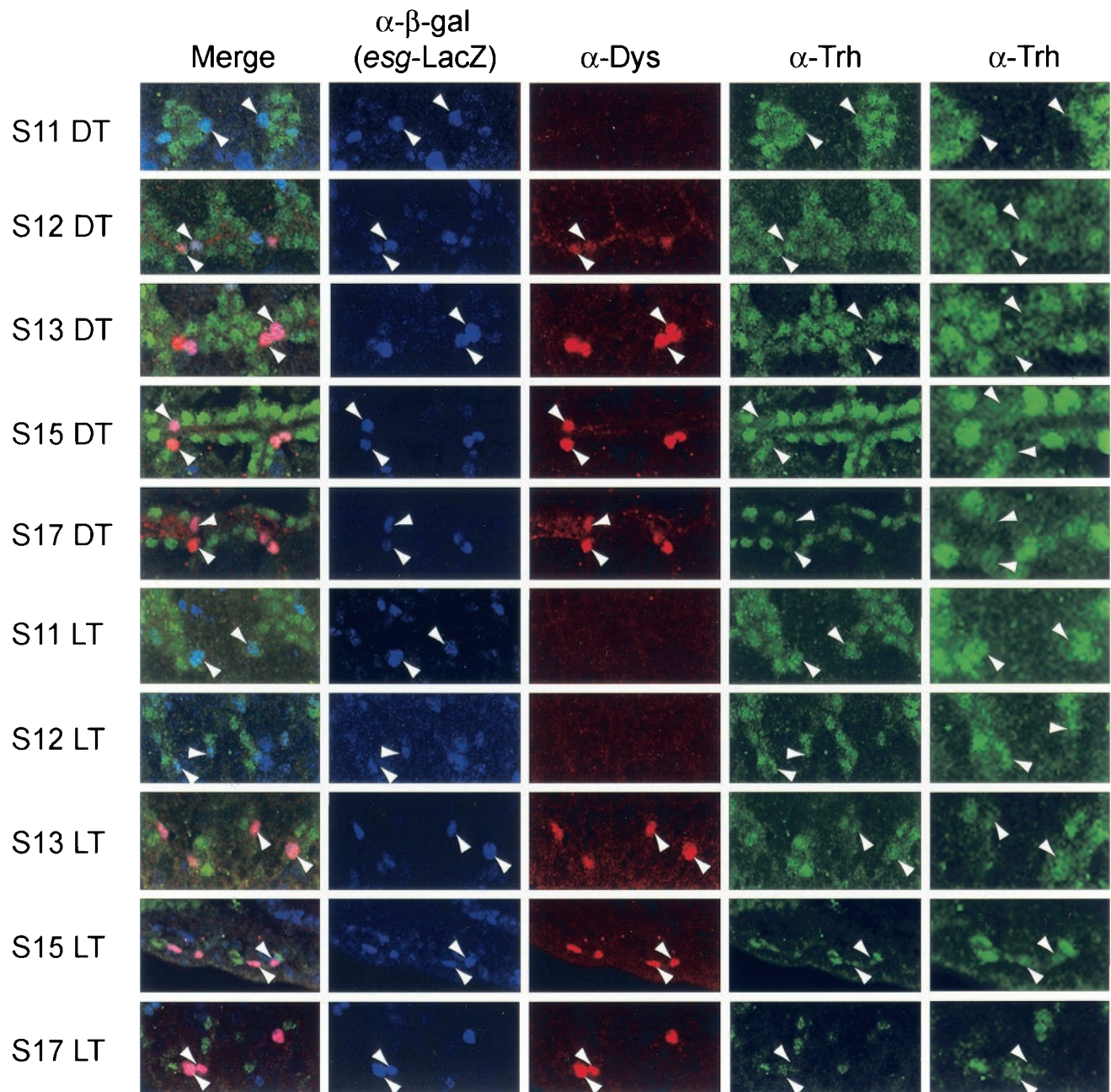
FIG. 6. Expression of *dys* during tracheal development. Images reveal the relative dynamics of Dys and Trh protein localization and *esg* expression. Embryos that express *lacZ* in an *esg* pattern (*esg*^{05730} P[*esg-lacZ*]/TM3) were triple stained with α-β-Gal (blue), α-Dys (red), and α-Trh (green). Shown are five successive stages of dorsal trunk (DT) development and five successive stages of lateral trunk (LT) development. The stages and tracheal branch type are indicated to the left of the images. The merge column shows triple staining, and the two α-Trh column are different magnifications of the same cells. Arrowheads indicate the location of selected tracheal fusion cells.

and (ii) examination of larval tracheae by using bright-field optics. The results showed a complete absence of lateral trunk, dorsal branch, and ganglionic branch fusion in embryos injected with *dys*-dsRNA. Embryos injected with TE or *GFP*-dsRNA showed normal tracheal fusion. In wild-type embryos, lateral trunk fusion occurs at stage 15 when the anterior lateral trunk buds and posterior lateral trunk buds from adjacent segments fuse (Fig. 8A and C) (24). Embryos at stage 15 or older and larvae from *dys*-dsRNA-injected embryos do not

have fused lateral trunks (Fig. 8B and D). The frequency of *dys*-dsRNA-injected embryos with lateral trunk defects was 92% ($n = 25$). Each embryo with lateral trunk defects showed a lack of lateral trunk tracheal fusion in all segments, indicating that expressivity was at 100%. In contrast, every embryo injected with TE or *GFP*-dsRNA ($n = 10$ each) showed fused lateral trunks in all segments. The specificity of the *dys*-RNAi treatment was validated for several reasons: (i) the overall embryonic morphology was normal, (ii) many embryos with
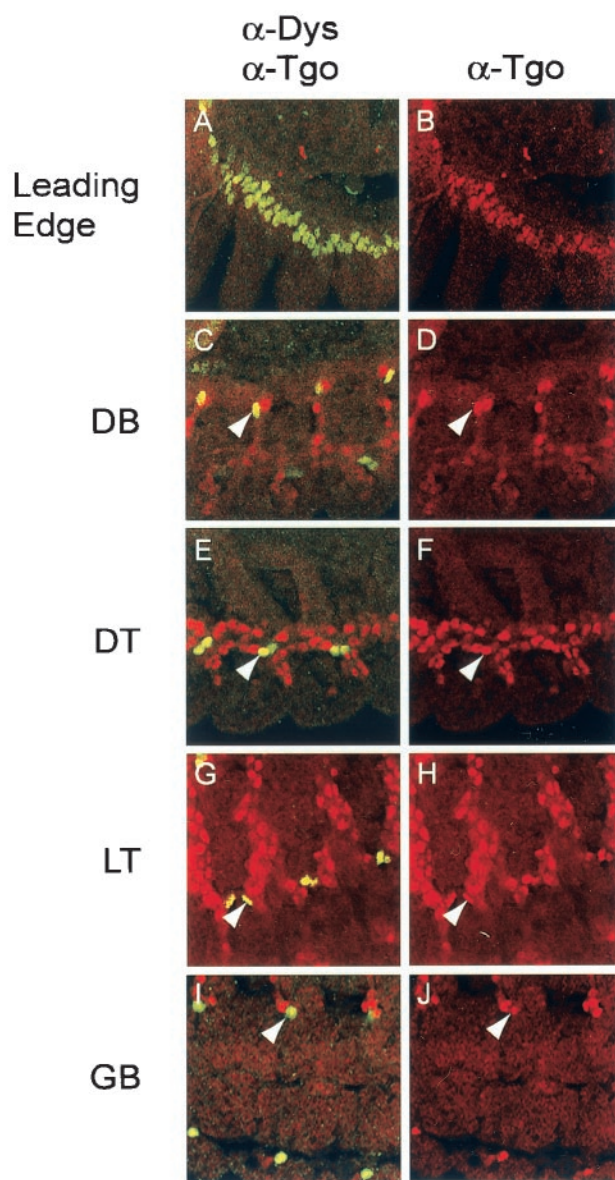
FIG. 7. Dys and Tgo show nuclear colocalization. Embryos were double stained with α-Dys (green) and α-Tgo (red). Sagittal views of stage 14 (A to F, I, and J) and stage 15 (G and H) embryos are shown. Images on the left show double-labeled merges (fusion cells in C, E, G, and I are yellow), and the images on the right show only the α-Tgo red channel. Arrowheads indicate a Dys-positive cell for comparison between left and right images. (A and B) Leading-edge cells show nuclear accumulation of both Dys and Tgo. (C to J) Tracheal fusion cells show nuclear colocalization with Tgo and that Tgo nuclear levels are comparable to those of other tracheal cells. The tracheal branch type is shown to the left of the images.

tracheal fusion defects were healthy enough to hatch and survive until the second- and third-instar stages, (iii) tracheal branching and morphology were generally normal (the only gross defects were in fusion), and (iv) control embryos had normal tracheal fusion, whereas *dys*-RNAi embryos showed a high degree of fusion defects.

Examination of the larval dorsal and ganglionic branches in *dys*-dsRNA-injected individuals showed a complete absence of

fusion (Fig. 8E to H). No segments were observed with fused dorsal branch or ganglionic branches. Branching and migration of the dorsal branch and the ganglionic branch were essentially normal; only fusion defects were observed. Dorsal trunk fusion appeared normal (Fig. 8I and J), even though dorsal trunk fusion cells express *dys*, and Dys protein was absent in *dys*-dsRNA-injected embryos. The lack of dorsal branch, lateral trunk, and ganglionic branch fusion and the presence of dorsal trunk fusion in *dys*-dsRNA-injected embryos roughly mimics mutations in *esg*, which also affect dorsal branch, lateral trunk, and ganglionic branch fusion but not dorsal trunk fusion (42, 46). It was not possible to confirm the *dys*-RNAi results genetically, since analysis of the *Df(3R)Espl3 dys* deletion strain was uninformative. Staining of *Df(3R)Espl3* with anti-Trh and MAb 2A12 showed only the rudiments of antibody-positive staining, and these mutant embryos lack identifiable tracheae, presumably due to the deletion of genes other than *dys* (data not shown). Thus, fusion events could not be assayed.

Dys downregulates levels of Trh. *trh* is expressed in all tracheal cells at the beginning of tracheal development and acts as a master regulator of tracheal development. Trh proteins levels remain high during tracheal development in most tracheal cells (Fig. 6) (49). Since Tgo is the likely dimerization partner for both Dys and Trh, and these proteins could, in principle, compete for Tgo in the same cells, we examined whether Trh levels remained constant in tracheal fusion cells in the presence of high levels of Dys. Heterozygous *esg[05730]* embryos expressing *esg-lacZ* to mark fusion cells were stained for Dys, β-Gal, and Trh (Fig. 6). Before *dys* is expressed, Trh is at high levels in fusion cells at equivalent levels to adjacent tracheal cells. As Dys protein accumulates in the fusion cells, Trh levels decline and remain low through the end of tracheal development. This occurs in fusion cells in the dorsal trunk (Fig. 6), lateral trunk (Fig. 6), dorsal branch (not shown), and ganglionic branch (not shown). Examination of Dys-positive tracheal fusion cells ($n =$ 82) indicated that 93% showed a marked reduction in Trh levels compared to adjacent tracheal cells.

The requirement of *dys* in regulating Trh levels was examined by staining *esg[05730] esg-lacZ* embryos injected with *dys*-dsRNA for Trh in fusion cells (Fig. 9). Only heterozygous embryos were examined. Tracheal fusion cells were identified by staining for β-Gal. Levels of β-Gal in fusion cells were comparable to control-injected embryos, indicating that *esg* transcription is not substantially affected by the depletion of Dys. Trh levels in Dys-depleted fusion cells were at high levels comparable to the levels observed in adjacent tracheal cells. Quantitatively, 96% of fusion cells examined in *dys*-dsRNA-injected embryos showed high levels of Trh ($n = 23$ cells). This indicates that the reduction in Trh levels in fusion cells is dependent on *dys* function.

The *trh*, *tgo*, and *dfr* genes encode a transcription factor complex that is required for the development of tracheal cells from dorsal ectoderm and, directly or indirectly, the expression of most genes expressed in the trachea (4, 16, 50, 53). Appearance of Trh protein precedes Dys by several hours (50). The expression of *dys* was examined in a *trh* mutant background and found to be absent (data not shown). Thus, *trh* is required for *dys* expression, and Dys, in turn, modulates Trh levels.

Fusion cell levels of Dys are dependent on *esg* in all branches except the dorsal trunk. Previous analysis of *esg*

GFP-dsRNA   *dys*-dsRNA

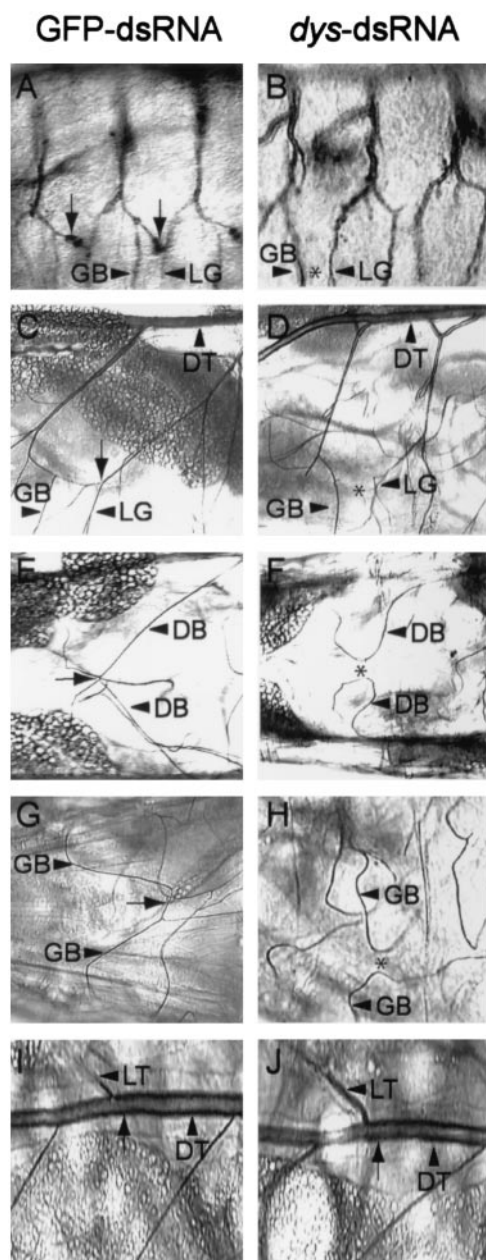FIG. 8. Removal of *dys* function by *dys*-RNAi reveals tracheal fusion defects. Blastoderm embryos (*w1118*) were injected with either *GFP*-dsRNA (negative control) or *dys*-dsRNA and assayed for tracheal defects. (A) Stage 16 embryo injected with *GFP*-dsRNA and stained with α-Dys and MAb 2A12 that stains the tracheal lumen. Prominently shown is the lateral trunk, which has fused. Arrows point to lateral trunk Dys-positive fusion cells. The ganglionic branch (arrowhead GB) and lateral branch G (arrowhead LG), which emanates from the lateral trunk, are also indicated. (B) Stage 16 embryo injected with *dys*-dsRNA and stained with α-Dys and MAb 2A12. The embryo showed a lack of lateral trunk fusion (the asterisk indicates the normal location of a lateral-trunk fusion in wild-type or control embryos). Ganglionic Branch (GB) and lateral branch a (LG) are noted by arrowheads. No Dys-positive cells were observed, indicating that the *dys*-RNA effectively abolished Dys protein. (C to J) Second-instar larvae injected with either *GFP*-dsRNA or *dys*-dsRNA were examined by bright-field microscopy for tracheal defects. Sites of fusion are indicated by an arrow in control embryos. (C and D) Sagittal views showing the lateral trunk with ganglionic branches and lateral branch G. The dorsal trunk is also shown. The *GFP*-RNA larvae showed

mutants revealed that expression of fusion cell genes was affected in some branches (dorsal branch and ganglionic branch) but not others (dorsal trunk) (42, 46). In the dorsal branch and the ganglionic branch, several fusion cell genes required *esg* function, but others did not. Expression of *dys* follows *esg* expression in all tracheal branches, raising the possibility that its expression requires *esg* function. Multiple *esg* mutants were analyzed, including the deletion strain *esgG66* (Fig. 10). Analysis of *esgG66* mutant embryos showed a complete absence of dorsal branch and ganglionic branch *dys* expression in *esg* mutant fusion cells. The lateral trunk also had no Dys-positive cells but had some *esg-lacZ*-positive cells, a result consistent with an earlier report that lateral trunk fusion cells die in *esg* mutant embryos (42). Thus, it is difficult to assess whether *esg* controls *dys* expression in the lateral trunk. In contrast, all dorsal trunk fusion cells were Dys positive. The P-element mutant, *esg05730*, showed a less severe phenotype. The dorsal branch fusion cells were all Dys negative, and dorsal trunk fusion cells were all Dys positive, as with the *esgG66* mutant embryos. However, there were occasionally *esg-lacZ*-positive cells present in the defective, unfused lateral trunk branches; some were Dys positive, and some were Dys negative. This was also observed in *esgG66*, but only rarely. The ganglionic fusion cells of *esg05730* were also occasionally Dys positive. In summary, *dys* fusion cell expression requires *esg* in fusion cells of the dorsal branch and ganglionic branch, and possibly the lateral trunk, but not in the dorsal trunk.

DISCUSSION

***Drosophila* Dys is a member of a distinct bHLH-PAS subfamily.** *Drosophila* Dys belongs to a novel conserved subfamily of bHLH-PAS proteins that includes *C. elegans* C15C8.2 and H-NXF. Dys also likely belongs to an extended family of *Drosophila* bHLH-PAS proteins, including Sim, Sima, Ss, and Trh, that dimerize with Tgo. The protein structure of Dys is conventional for bHLH-PAS proteins. The bHLH domain is near the N terminus of the protein and is followed by the PAS-1 and PAS-2 domains. One unusual feature compared to other bHLH-PAS proteins is the relatively long (152 aa) region N-terminal to the bHLH domain. This region has a large number of glutamine residues and may act as a transcriptional activation domain. The C-terminal residues after the PAS-2 domain are unconserved with C15C8.2, H-NXF, or any other protein but have histidine-rich, proline-rich, and glutamine-rich regions. These residues may also be transcriptional activation domains. The structure of the Dys protein suggests a DNA-

normal lateral trunk fusion (arrow) (C), but the *dys*-RNA larval lateral trunk failed to fuse (✻) (D). (E) Dorsal view showing the dorsal fusion site (arrow) of two dorsal branches in the *GFP*-RNA larva. (F) The *dys*-RNA dorsal branches migrate to the dorsal midline but fail to fuse (✻). (G and H) Ventral view showing the sites of ventral anastomosis of two pair of ganglionic branches. (G) In the *GFP*-RNA larva the arrow indicates the fusion of the two ganglionic branches indicated by arrowheads. (H) Fusion failed to occur (✻) in the *dys*-RNA larva. (I and J) Dorsal view showing that the dorsal trunks are fused in both *GFP*-RNA (I) and *dys*-RNA (J) larvae. The fusion site is indicated by arrows, and the lateral trunk branch is also labeled.
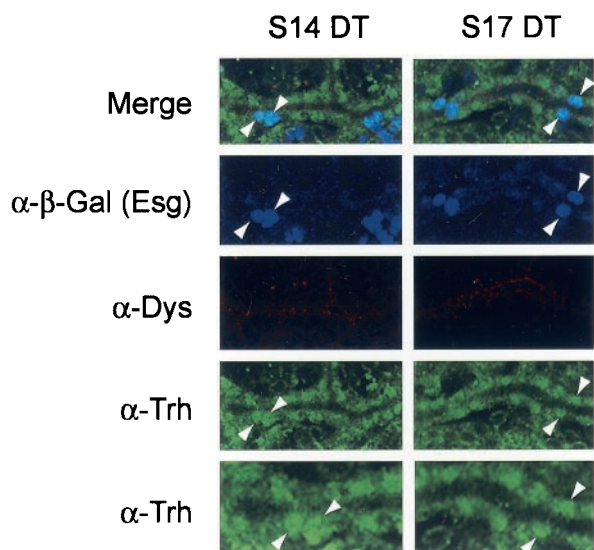
FIG. 9. *dys*-dsRNA-injected embryos show increased levels of Trh in fusion cells. Embryos from *esg^05730* P[*esg-lacZ*]/CyO P[*ftz-lacZ*] parents were injected with *dys*-dsRNA and immunostained with α-β-Gal (*esg-lacZ* and *ftz-lacZ*) (blue), α-Dys (red), and α-Trh (green). Only heterozygous embryos were analyzed. Dorsal trunks at stages 14 and 17 are shown. Arrowheads indicate selected fusion cells. *esg-lacZ* is expressed at comparable levels in *dys*-dsRNA-injected embryos compared to wild-type or *GFP*-dsRNA (control)-injected embryos (see Fig. 6). Staining with α-Dys shows an absence of Dys protein. The Trh levels in fusion cells are significantly higher than in wild-type or control-injected embryos (see Fig. 6) and are comparable to the levels in adjacent tracheal cells. The two α-Trh rows are higher magnifications of the same embryo.

binding transcriptional activator, but this needs to be tested biochemically.

dys is expressed in a variety of embryonic cell types. *Drosophila dys* is expressed in a variety of embryonic cell types, including tracheal fusion, leading edge, foregut atrium, brain or stomagastric nervous system, hindgut, and anal pad cells. Initial expression was observed during mid embryogenesis at stage 12. The function of *dys* in these cell types is unknown, with the exception of the tracheal fusion cells, in which we showed that *dys* plays a developmental role. The *dys*-RNAi results showed that *dys* is an essential gene. However, *dys*-dsRNA-injected embryos did not die as embryos, but as second- and third-instar larvae. Thus, the role of *dys*, as assayed by RNAi, in the various embryonic cell types is not dramatic enough to cause embryonic lethality. Mutations in *esg* show fusion defects in the same tracheal branches as *dys* and are also lethal to larvae. The tracheal fusion defects and resulting putative respiratory deficiencies may be the cause of the larval lethality, since other observations have shown that animals with defective tracheae survive until late larval periods (24). However, both genes are expressed elsewhere, and defects in the other cell types may contribute to lethality.

dys is required for tracheal fusion. *dys* is expressed in tracheal fusion cells, and no other tracheal cells. This was shown by coexpression of *dys* with *esg*, a gene that is expressed in fusion cells and regulates tracheal fusion. *dys*-RNAi experiments were carried out to examine whether *dys* is involved in tracheal fusion. The results demonstrated that *dys* was re-

quired for fusion of the dorsal branch, lateral trunk, and ganglionic branch but not of the dorsal trunk. This phenotype was similar to the *esg* mutant phenotype that also affected the dorsal branch, lateral trunk, and ganglionic branch but not the dorsal trunk. Although branches differ in the details of the fusion process, tracheal fusion generally requires migration, recognition, and adhesion of fusion cells. *dys*-RNAi embryos showed relatively normal tracheal branches and migration. The occurrence of a single *esg-lacZ* cell in each *dys*-RNAi branch indicated that *esg*-positive tracheal fusion cells were present, and thus survival and gross cell fate is not controlled by *dys*. It is possible that *dys* controls aspects of fusion cell recognition, cell adhesion, or inhibition of nonfusion tracheal functions, such as branching.

Since *dys*, as well as *esg*, is expressed in dorsal trunk fusion cells, why is dorsal trunk fusion apparently unaffected in *dys*-RNAi-injected embryos? It is not likely due to incomplete expressivity of the *dys*-RNAi, since we were not able to detect Dys protein in *dys*-RNAi-injected embryos, including dorsal trunk fusion cells, and ~100% of lateral trunk, dorsal branch, and ganglionic branch branches in *dys*-dsRNA-injected embryos failed to fuse. As discussed previously (46), there are a number of differences between the dorsal trunk and the other branches that could contribute to differences in fusion behav-
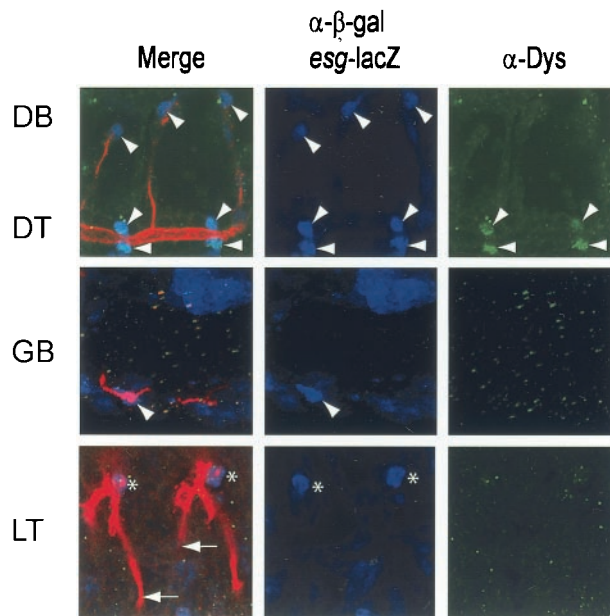

FIG. 10. *dys* fusion cell expression requires *esg* function in some tracheal branches. Anterior is to the left. Arrowheads indicate fusion cells. Fusion cell gene expression was examined in *esg^G66* mutant embryos, which express *esg-lacZ*. Stage 15 embryos were stained with α-β-Gal and α-Dys. Merge images are shown on the left. The dorsal branch (DB) has *esg-lacZ*-positive fusion cells, but there is no Dys protein present. The dorsal trunk (DT) fusion cells express both *esg-lacZ* and *dys*. Occasionally, *esg^G66* mutant embryos had a reduced number of dorsal trunk fusion cells. The ganglionic branch (GB) has *esg-lacZ*-positive fusion cells, but Dys protein is absent. Most lateral trunk (LT) fusion cells are absent due to cell death (42), and there are no *esg-lacZ* or Dys-positive cells present. The arrows indicate sites in which fusion cells are located in wild-type lateral trunks. *esg-lacZ*-positive cells that are not fusion cells are indicated by an asterisk.

ior. The larger diameter dorsal trunk has multiple cells comprising its circumference, unlike most of the other branches, which are thinner and have a single cell comprising the circumference. Dorsal trunk branches are in close proximity to their fusion partner and lack the filopodial extensions that help guide the other branches to their targets. The dorsal trunk also utilizes a mesodermal guidepost cell that mediates fusion (51). Similar guidepost cells have not been described for the other branches. Finally, *breathless* RNA levels begin to decline by stage 12 in the dorsal trunk due to *spalt* repression (31), which may eliminate the potential need to reduce *breathless* levels by decreasing Trh levels. These and other possible differences suggest why *dys* and *esg* may have different functions in different branches.

The *esg* gene is required in dorsal branch and ganglionic branch tracheal fusion cells for expression of several genes, including *shotgun* (*DE-cadherin*) and three late-expressing fusion cell genes (*fusion-4* to *fusion-6*), as well as repression of terminal branching genes (*DSRF* and *terminal-1*) (42, 46). Expression of two early-expressing fusion cell genes (*fusion-2* and *fusion-3*) are not dependent on *esg*. *dys* expression is also dependent on *esg*, in keeping with the role of *esg* in regulating early-expressing fusion cell gene expression. As with other genes expressed in fusion cells, *dys* expression was not dependent on *esg* in dorsal trunk cells. This implies that the ability of *esg* to activate transcription is fusion cell dependent and is due to the presence of different coregulatory proteins or modifier proteins in the different branches.

Cross-regulatory interactions among bHLH-PAS proteins. The *trh* gene is required for initiation of tracheal formation. *trh* expression is maintained throughout embryonic development in most tracheal cells, and this continued expression is due to autoregulation (50). However, the role of *trh* beyond its role in initiating tracheal formation is not well understood. We have shown that Trh protein levels fall specifically in all classes of tracheal fusion cells coincident with the rise in Dys levels. The nuclear levels of Tgo, the partner for both Dys and Trh, remain constant in fusion cells. The biological significance of the reduction in Trh remains to be investigated, but it is possible that fusion requires a reduction in Trh. One possibility is that Trh:Tgo is required for the expression or function of the *breathless* (*btl*) tyrosine kinase receptor that guides growing tracheal branches (31) and that *btl* function must be inhibited in fusion cells. Potentially, the only function of *dys* is to reduce Trh levels.

There are multiple mechanisms in which Dys could regulate Trh levels. These mechanisms include (i) competition between Dys and Trh for dimerization with Tgo, (ii) competitive Dys:Tgo binding to Trh:Tgo autoregulatory binding sites in the *trh* gene, (iii) activation of genes by Dys that encode proteins influencing *trh* RNA or protein stability, and (iv) inhibition of protein kinase B that is required for Trh nuclear transport (19). Conceptually, the first model is the simplest and most attractive. Trh autoregulates its own expression (50), and reduction in Trh:Tgo complexes by competition for Tgo by Dys would lead to a reduction in *trh* RNA and protein. In the second model, Dys:Tgo would function as a transcriptional repressor and extinguish *trh* RNA synthesis by binding Trh:Tgo autoregulatory sequences within the *trh* gene.

Evidence for the possible roles of these mechanisms has

emerged from studies on vertebrate and *Drosophila* bHLH-PAS proteins. In one study (13), it was demonstrated that HIF-1α outcompetes the Aryl hydrocarbon receptor (Ahr) bHLH-PAS protein for their common dimerization partner, Arnt, which is the vertebrate Tgo ortholog. In another study (52), Sim2 was shown to compete with HIF-1α for Arnt and partially block expression of a HIF-1α:Arnt responsive reporter gene. Sim2 can repress transcription and, by binding to HIF-1α:Arnt recognition sites on the reporter gene, it reduced its expression. The third model in which the presence of Dys reduces protein levels by activating the transcription of repressive or inhibitory factors is analogous to how Sim:Tgo represses the expression of genes in the central nervous system midline cells by activating transcription of a repressor (10). One additional issue is whether Dys reduces the levels of *Drosophila* bHLH-PAS proteins besides Trh. One possibility is Sim. Both Sim and Dys are expressed in anal pad cells (36), and *sim* mutants have anal pad defects (26, 36).

Evolution of bHLH-PAS proteins. *Drosophila* has four bHLH-PAS proteins that dimerize with Tgo: Sim, Sima, Ss, and Trh. Dys is also likely to dimerize with Tgo. One particularly interesting observation is that *Drosophila* Dys, Trh, and Sima are all involved in aspects of tracheal development. Mammals have closely related members of all of these proteins (Fig. 2). The mammalian proteins dimerize with either Arnt or the closely related Arnt2. *C. elegans* has four bHLH-PAS partners for AHA-1, the worm Tgo/Arnt ortholog. These partners include (i) C15C8.2, which is related to Dys and mammalian NXF; (ii) AHR-1, which is related to *Drosophila* Ss and vertebrate Ahr; (iii) HIF-1, which is related to *Drosophila* Sima and vertebrate HIF-1α; and (iv) T01D3.2, which is related to Sim and Trh. Since orthologs of Dys, Ss, Sima, and Sim/Trh are found in vertebrates, insects, and nematodes, these proteins had already diverged in the common ancestor of these species. The *C. elegans* C15C8.2 gene is expressed in the pharynx, a feeding organ (Powell-Coffman, unpublished). The mammalian *dys* gene, *NXF*, has been detected only in the brain (Patterson, unpublished). The evolutionary conservation of the Dys subfamily of proteins suggests a functional relationship, although the sites of expression in *Drosophila* and the other organisms, as studied to date, are diverse. One possibility based on the tracheal fusion phenotypes is that Dys regulates aspects of cell recognition or cell adhesion events. It will be interesting to determine in future studies what biochemical, developmental, or physiological features of the Dys proteins are conserved, as well as the evolutionary origins regarding the tracheal functions of Dys, Sima, and Trh.

REFERENCES

1. **Adams, M. D., S. E. Celniker, R. A. Holt, C. A. Evans, J. D. Gocayne, P. G. Amanatides, S. E. Scherer, P. W. Li, R. A. Hoskins, R. F. Galle, et al.** 2000. The genome sequence of *Drosophila melanogaster*. Science **287:**2185–2195.
2. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403–410.
3. **Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:**3389–3402.
4. **Boube, M., M. Llimargas, and J. Casanova.** 2000. Cross-regulatory interactions among tracheal genes support a co-operative model for the induction of tracheal fates in the *Drosophila* embryo. Mech. Dev. **91:**271–278.
5. **Casso, D., F. Ramirez-Weber, and T. B. Kornberg.** 2000. GFP-tagged balancer chromosomes for *Drosophila melanogaster*. Mech. Dev. **91:**451–454.
6. **Consortium, International Human Genome Sequencing.** 2001. Initial sequencing and analysis of the human genome. Nature **409:**860–921.
7. **Consortium, C. elegans Sequencing.** 1998. Genome sequence of the nematode *Caenorhabditis elegans*: a platform for investigating biology. Science **282:**2012–2018.
8. **Crews, S. T.** 1998. Control of cell lineage-specific development and transcription by bHLH-PAS proteins. Genes Dev. **12:**607–620.
9. **Davies, J. A.** 2002. Do different branching epithelia use a conserved developmental mechanism? Bioessays **24:**937–948.
10. **Estes, P., J. Mosher, and S. T. Crews.** 2001. *Drosophila single-minded* represses gene transcription by activating the expression of repressive factors. Dev. Biol. **232:**157–175.
11. **FlyBase Consortium.** 1999. The FlyBase database of the *Drosophila* genome projects and community literature. Nucleic Acids Res. **27:**85–88.
12. **Franks, R. G., and S. T. Crews.** 1994. Transcriptional activation domains of the Single-minded bHLH protein are required for CNS midline cell development. Mech. Dev. **45:**269–277.
13. **Gradin, K., J. McGuire, R. H. Wenger, I. Kvietikova, M. L. fhitelaw, R. Toftgard, L. Tora, M. Gassmann, and L. Poellinger.** 1996. Functional interference between hypoxia and dioxin signal transduction pathways: competition for recruitment of the Arnt transcription factor. Mol. Cell. Biol. **16:**5221–5231.
14. **Hartenstein, V.** 1997. Development of the insect stomatogastric nervous system. Trends Neurosci. **20:**421–427.
15. **Holt, R. A., G. M. Subramanian, A. Halpern, G. G. Sutton, R. Charlab, D. R. Nusskern, P. Wincker, A. G. Clark, J. M. Ribeiro, R. Wides, et al.** 2002. The genome sequence of the malaria mosquito *Anopheles gambiae*. Science **298:**129–149.
16. **Isaac, D. D., and D. J. Andrew.** 1996. Tubulogenesis in *Drosophila*: a requirement for the *tracheless* gene product. Genes Dev. **10:**103–117.
17. **Jacinto, A., S. Woolner, and P. Martin.** 2002. Dynamic analysis of dorsal closure in *Drosophila*: from genetics to cell biology. Dev. Cell **3:**9–19.
18. **Jarecki, J., E. Johnson, and M. A. Krasnow.** 1999. Oxygen regulation of airway branching in *Drosophila* is mediated by branchless FGF. Cell **99:**211–220.
19. **Jin, J., N. Anthopoulos, B. Wetsch, R. C. Binari, D. D. Isaac, D. J. Andrew, J. R. Woodgett, and A. S. Manoukian.** 2001. Regulation of *Drosophila* tracheal system development by protein kinase B. Dev. Cell **1:**817–827.
20. **Kennerdell, J. R., and R. W. Carthew.** 1998. Use of dsRNA-mediated genetic interference to demonstrate that *frizzled* and *frizzled 2* act in the Wingless pathway. Cell **95:**1017–1026.
21. **Knust, E.** 1997. *Drosophila* morphogenesis: movements behind the edge. Curr. Biol. **7:**R558–R561.
22. **Lavista-Llanos, S., L. Centanin, M. Irisarri, D. M. Russo, J. M. Gleadle, S. N. Bocca, M. Muzzopappa, P. J. Ratcliffe, and P. Wappner.** 2002. Control of the hypoxic response in *Drosophila melanogaster* by the basic helix-loop-helix PAS protein similar. Mol. Cell. Biol. **22:**6842–6853.
23. **Ledent, V., and M. Vervoort.** 2001. The basic helix-loop-helix protein family: comparative genomics and phylogenetic analysis. Genome Res. **11:**754–770.
24. **Manning, G., and M. A. Krasnow.** 1993. Development of the *Drosophila* tracheal system, p. 609–685. *In* M. Bate and A. Martinez Arias (ed.), The development of *Drosophila melanogaster*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
25. **Martin-Blanco, E., A. Gampel, J. Ring, K. Virdee, N. Kirov, A. M. Tolkovsky, and A. Martinez-Arias.** 1998. puckered encodes a phosphatase that mediates a feedback loop regulating JNK activity during dorsal closure in *Drosophila*. Genes Dev. **12:**557–570.
26. **Mayer, U., and C. Nüsslein-Volhard.** 1988. A group of genes required for pattern formation in the ventral ectoderm of the *Drosophila* embryo. Genes Dev. **2:**1496–1511.
27. **Metzger, R. J., and M. A. Krasnow.** 1999. Genetic control of branching morphogenesis. Science **284:**1635–1639.
28. **Misquitta, L., and B. M. Paterson.** 1999. Targeted disruption of gene func-

tion in *Drosophila* by RNA interference (RNA-i): a role for nautilus in embryonic somatic muscle formation. Proc. Natl. Acad. Sci. USA **96:**1451–1456.
29. **Nambu, J. R., J. O. Lewis, and S. T. Crews.** 1993. The development and function of the *Drosophila* CNS midline cells. Comp. Biochem. Physiol. **104A:**399–409.
30. **Ohe, N., K. Saito, and H. Kaneko.** 2002. The cDNA cloning of a novel bHLH PAS transcription factor NXF. GenBank, [Online.] www.ncbi.nlm.nih.gov.
31. **Ohshiro, T., Y. Emori, and K. Saigo.** 2002. Ligand-dependent activation of breathless FGF receptor gene in *Drosophila* developing trachea. Mech. Dev. **114:**3.
32. **Ohshiro, T., and K. Saigo.** 1997. Transcriptional regulation of *breathless* FGF receptor gene by binding of TRACHEALESS/dARNT heterodimers to three central midline elements in *Drosophila* developing trachea. Development **124:**3975–3986.
33. **Page, R. D.** 1996. TreeView: an application to display phylogenetic trees on personal computers. Comput. Appl. Biosci. **12:**357–358.
34. **Patel, N. H., P. M. Snow, and C. S. Goodman.** 1987. Characterization and cloning of fasciclin III: a glycoprotein expressed on a subset of neurons and axon pathways in *Drosophila*. Cell **48:**975–988.
35. **Peyrefitte, S., D. Kahn, and M. Haenlin.** 2001. New members of the *Drosophila* Myc transcription factor subfamily revealed by a genome-wide examination for basic helix-loop-helix genes. Mech. Dev. **104:**99–104.
36. **Pielage, J., G. Steffes, D. C. Lau, B. A. Parente, S. T. Crews, R. Strauss, and C. Klambt.** 2002. Novel behavioral and developmental defects associated with *Drosophila single-minded*. Dev. Biol. **249:**283–299.
37. **Reese, M. G., D. Kulp, H. Tammana, and D. Haussler.** 2000. Genie: gene finding in *Drosophila melanogaster*. Genome Res. **10:**529–538.
38. **Risau, W., and I. Flamme.** 1995. Vasculogenesis. Annu. Rev. Cell Dev. Biol. **11:**73–91.
39. **Rubin, G. M., L. Hong, P. Brokstein, M. Evans-Holm, E. Frise, M. Stapleton, and D. A. Harvey.** 2000. A *Drosophila* complementary DNA resource. Science **287:**2222–2224.
40. **Saitou, N., and M. Nei.** 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4:**406–425.
41. **Samakovlis, C., N. Hacohen, G. Manning, D. Sutherland, K. Guillemin, and M. A. Krasnow.** 1996. Development of the *Drosophila* tracheal system occurs by a series of morphologically distinct but genetically coupled branching events. Development **122:**1395–1407.
42. **Samakovlis, C., G. Manning, P. Steneberg, N. Hacohen, R. Cantera, and M. A. Krasnow.** 1996. Genetic control of epithelial tube fusion during *Drosophila* tracheal development. Development **122:**3531–3536.
43. **Sonnenfeld, M., M. Ward, G. Nystrom, J. Mosher, S. Stahl, and S. Crews.** 1997. The *Drosophila tango* gene encodes a bHLH-PAS protein that is orthologous to mammalian Arnt and controls CNS midline and tracheal development. Development **124:**4583–4594.
44. **Sutherland, D., C. Samakovlis, and M. A. Krasnow.** 1996. *branchless* encodes a *Drosophila* FGF homolog that controls tracheal cell migration and the pattern of branching. Cell **87:**1091–1101.
45. **Swofford, D. L.** 2002. PAUP*: phylogenetic analysis using parsimony (*and other methods), version 4. Sinauer Associates, Sunderland, Mass.
46. **Tanaka-Matakatsu, M., T. Uemura, H. Oda, M. Takeichi, and S. Hayashi.** 1996. Cadherin-mediated cell adhesion and cell motility in *Drosophila* trachea regulated by the transcription factor Escargot. Development **122:**3697–3705.
47. **Taylor, B. L., and I. B. Zhulin.** 1999. PAS domains: internal sensors of oxygen, redox potential, and light. Microbiol. Mol. Biol. Rev. **63:**479–506.
48. **Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins.** 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. **25:**4876–4882.
49. **Ward, M. P., J. T. Mosher, and S. T. Crews.** 1998. Regulation of *Drosophila* bHLH-PAS protein cellular localization during embryogenesis. Development **125:**1599–1608.
50. **Wilk, R., I. Weizman, L. Glazer, and B.-Z. Shilo.** 1996. *tracheless* encodes a bHLH-PAS protein and is a master regulator gene in the *Drosophila* tracheal system. Genes Dev. **10:**93–102.
51. **Wolf, C., and R. Schuh.** 2000. Single mesodermal cells guide outgrowth of ectodermal tubular structures in *Drosophila*. Genes Dev. **14:**2140–2145.
52. **Woods, S. L., and M. L. Whitelaw.** 2002. Differential activities of murine single minded 1 (SIM1) and SIM2 on a hypoxic response element: cross-talk between basic helix-loop-helix/per-Arnt-Sim homology transcription factors. J. Biol. Chem. **277:**10236–10243.
53. **Younossi-Hartenstein, A., and V. Hartenstein.** 1993. The role of the tracheae and musculature during pathfinding of *Drosophila* embryonic sensory axons. Dev. Biol. **158:**430–447.