**INTRODUCTION**

The *Drosophila* mesoderm forms from the ventral-most cells of the early embryo that invaginate during gastrulation. The expression and function of two transcription factors, the basic helix-loop-helix (bHLH) protein Twist (Twi) and the zinc-finger protein Snail (Sna), in ventral cells located between ~15 and 85% egg length are essential for their invagination and for subsequent mesodermal tissue development. Although the absence of either *twi* or *sna* activity results in similar phenotypes, the molecular roles of the two genes in this pathway differ. *twi* functions in activating a variety of mesoderm-specific target genes, including several that are known to regulate the invagination, patterning and differentiation of the mesoderm. By contrast, *sna* is thought to act largely, if not exclusively, as a repressor of a number of neuroectodermal targets, permitting mesoderm formation by restricting the expression of these genes to areas outside the presumptive mesoderm (Kosman et al., 1991; Leptin, 1991; Reuter and Leptin, 1994; Hemavathy et al., 1997; Wakabayashi-Ito and Ip, 2005; Sandmann et al., 2007; Zeitlinger et al., 2007).

Notably, however, there is at least one group of mesodermal cells that requires *sna*, but not *twi*, for its initial phase of development. This cell group is located ventrally within the domain of early *twi* and *sna* expression, but is restricted to the posterior tip of the mesoderm between ~7.5 and 15% egg length. Because it is fated to develop into the longitudinal muscles of the midgut, it has been termed the caudal visceral mesoderm (CVM), and migrate onto the trunk visceral mesoderm prior to undergoing myoblast fusion and muscle fiber formation. We show that *HLH54F* expression in the CVM is regulated by a combination of terminal patterning genes and *snail*. We generated *HLH54F* mutations and show that this gene is crucial for the specification, migration and survival of the CVM cells and the longitudinal midgut muscle founders. *HLH54F* mutant embryos, larvae, and adults lack all longitudinal midgut muscles, which causes defects in gut morphology and integrity. The function of *HLH54F* as a direct activator of gene expression is exemplified by our analysis of a CVM-specific enhancer from the Dorsocross locus, which requires combined inputs from *HLH54F* and Biniou in a feed-forward fashion. We conclude that *HLH54F* is the earliest specific regulator of CVM development and that it plays a pivotal role in all major aspects of development and differentiation of this largely *twist*-independent population of mesodermal cells.

**KEY WORDS:** bHLH proteins, Cell migration, Mesoderm, Specification, Visceral muscles, *Drosophila*

**SUMMARY**

*HLH54F*, the *Drosophila* ortholog of the vertebrate basic helix-loop-helix domain-encoding genes capsulin and musculin, is expressed in the founder cells and developing muscle fibers of the longitudinal midgut muscles. These cells descend from the posterior-most portion of the mesoderm, termed the caudal visceral mesoderm (CVM), and migrate onto the trunk visceral mesoderm prior to undergoing myoblast fusion and muscle fiber formation. We show that *HLH54F* expression in the CVM is regulated by a combination of terminal patterning genes and *snail*. We generated *HLH54F* mutations and show that this gene is crucial for the specification, migration and survival of the CVM cells and the longitudinal midgut muscle founders. *HLH54F* mutant embryos, larvae, and adults lack all longitudinal midgut muscles, which causes defects in gut morphology and integrity. The function of *HLH54F* as a direct activator of gene expression is exemplified by our analysis of a CVM-specific enhancer from the Dorsocross locus, which requires combined inputs from *HLH54F* and Biniou in a feed-forward fashion. We conclude that *HLH54F* is the earliest specific regulator of CVM development and that it plays a pivotal role in all major aspects of development and differentiation of this largely *twist*-independent population of mesodermal cells.

**KEY WORDS:** bHLH proteins, Cell migration, Mesoderm, Specification, Visceral muscles, *Drosophila*
Within the TVM, the founder cells of the circular midgut muscles are induced by Jelly belly (Jeb) signals acting through the receptor tyrosine kinase Alk (Englund et al., 2003; Lee et al., 2003). The founder myoblasts from the TVM fuse one-to-one with adjacent fusion-competent myoblasts to binucleated syncytia that form the circular midgut muscles (San Martin et al., 2001; Klapper et al., 2002). Subsequently, after the migrating CVM-derived founder cells have arrived at their destinations, each fuses with multiple fusion-competent cells from the TVM left over from the TVM founder cell fusions (San Martin et al., 2001). It is these multinucleated syncytia that will then differentiate into the longitudinal visceral muscles, which run perpendicularly to the circular muscles along the entire length of the midgut (Klapper et al., 2001). The longitudinal gut muscle fibers from the outer layer of the developing visceral musculature are tightly interwoven with the circular gut muscle fibers from the inner layer (Schröter et al., 2006).

**HLH54F** is the earliest known marker of the CVM primordia and its expression is maintained throughout the development and differentiation of the longitudinal midgut muscles (Georgias et al., 1997). To test whether **HLH54F** plays an important role in the development of the CVM and longitudinal midgut musculature, we generated loss-of-function mutations for this gene by imprecise P-excision and EMS mutagenesis screens. We demonstrate that in the absence of **HLH54F** activity, no longitudinal gut muscle founder cells are formed. The absence of all tested CVM markers and the observed apoptotic death of the cells that would normally be destined to form CVM in **HLH54F** mutants, show that **HLH54F** has an essential role in determining the CVM and in specifying the founder cells of the longitudinal gut musculature. This function includes feed-forward regulation and direct binding to target enhancers (e.g. from the Dorsocross genes). We also show that ectopic expression of **HLH54F** can interfere with normal somatic muscle, cardiac and TVM development. Further, we extend the known pathway of CVM development by showing that the initiation of **HLH54F** expression is largely independent of **twi**, but depends critically on the combined activities of **sna** and terminal patterning genes, particularly the synergistic activities of fork head (**fkh**) and brachynteron (**byn**). Hence, the CVM primordia are determined at the intersection of the domains of these mesodermal and terminal regulators.

**MATERIALS AND METHODS**

*Drosophila* strains

The following *Drosophila* strains were used: **twi**¹⁰, **hlh**¹⁰, **byn**⁴, **hkb**¹⁴¹⁸, **sna**¹⁴, **EY06760** (all from Bloomington Stock Center, Indiana University, USA), **fkb**¹⁷⁶, **byn**¹ fkb1⁷⁷⁵, **zfh**² (from R. Reuter, Tübingen University, Germany), **Df(3R)Exel5020**, **Df(2R)14H10W-21** and **Df(2R)02B10V-12** (Mohr and Gelbart, 2002) (from the authors), **cro-cac** (Hacker et al., 1995) and **sln m6-lacZ** #5 (Lee and Frasch, 2000). For UAS/GAL4-induced ectopic expression, embryos were grown at 28°C using the lines **UAS-p35** (Zhou et al., 1997), **UAS-bin #35** (Zaffran et al., 2001), **byn-GAL4** UAS-GFP (Johansen et al., 2003) (from J. Lengyel, UCLA, USA), **bap3-GAL4** (Lee et al., 2003), **2XP-E-twi-GAL4**: **24B-GAL4** (Reim and Frasch, 2005), **UAS-FLP-LSI** (P(Kk103428/v103965) and UAS-da-IR (**P(Ds3440)**/v51300) (both from VDRC, Vienna, Austria) (Dietzl et al., 2007).

**P**-excision and EMS mutagenesis screens

The viable P insertion EY06760 was used in an excision screen with A2-J-derived transposase. Balanced w males in F2 were crossed individually to Df(2R)14H10W-21 females (see Fig. 1) (Mohr and Gelbart, 2002). A total of 598 single excision lines were screened over Df(2R)14H10W-21 for semi-lethality. Semi-lethality was confirmed over Df(2R)02B10Y-12. For two semi-lethal lines, **HLH54F**¹²⁷ and **HLH54F**¹³⁸, the molecular lesions were sequenced using genomic PCR fragments from DNA of homozygous escaper flies.

In addition, we performed an EMS mutagenesis screen for mutants with disrupted specification or migration of the CVM. **HLH54F**-RFP reporter males were mutagenized with 30 mEM EMS (Lewis and Bacher, 1968) and balanced with Cyo, twi>EGFP. Embryos homozygous for the mutated chromosome were scored for the absence, reduction or mislocation of the RFP-labeled CVM cells. Two lines with absent RFP in the CVM failed to complement the semi-lethality of **HLH54F**¹²⁷ and were semi-lethal or lethal, respectively, in trans to larger deficiencies. PCR analysis identified **HLH54F**¹⁰²¹ as a deficiency of **HLH54F** and flanking regions (lethal, with yet undefined breakpoints). Sequencing of genomic DNA amplified from **HLH54F**¹²⁷ identified a C-to-T transition at position 121 of the open reading frame (ORF) as compared with the parental chromosome.

**Construction of HLH54F-lacZ and HLH54Fb-GAL4**

Genomic **HLH54F** sequences were cloned into pPhelin (upstream sequences, **HLH54Fa-lacZ** or P-helical (introns, **HLH54Fb-lacZ**); downstream sequences, **HLH54Fc-lacZ**) (Barolo et al., 2000). PCR fragments from yw DNA were obtained with the following primers (5′ to 3′): for the upstream region, primers HLH-UPS-5′ CCGAT-1ACCAATGTTTCAGATGTCCCGGAATCCCT; for the introns with exon 2, primers HLH-INTRON-5′ TCTCAAGGATGATGAAATCATATTTATAG and HLH-INTRON-3′ TAATGGTATGTGGATGATCCAAGGACCAT; for the downstream region, primers HLH-DS-5′ CTAGCCATGC-CCATGTCACCATAGCGAATA and HLH-DS-3′ TACATGGGGCATT- TCTCCTCAGAGTTTGTTCT. Only **HLH54Fb-lacZ** expressed lacZ.

**HLH54Fd-GAL4** was generated from the second site into another fragment amplified with primers Gal4-2 EcoRI (ATTGTGCAAACAGGAGGCTAATACTTTCC) and HLH-INTRON-3′ into the EcoRI site of the p221-GAL4 vector (from C. Klämbt, Münster University, Germany).

**Genomic rescue experiment**

**HLH54F**-genres, containing 4.5 kb of sequences from the **HLH54F** locus, including the intergenic regions up to the 5′ ends of both neighboring genes, was generated by genomic PCR and cloning into pCSpaeR2 (NotI/BamHI). Sequencing confirmed the correct ORF. Rescue of CVM formation was tested in the genetic background of w; **HLH54F**¹²⁷ and **HLH54Fb-lacZ/SM6 eve-lacZ** and **HLH54F**-genres. Upon crossing this line with **HLH54F**¹⁰²¹ Cyo twi>EGFP, one copy rescued the semi-lethality.

**Doc enhancer constructs**

The 566 bp DocF4-s1 enhancers (3L:9,022,917–9,023,482, vR5.27) were cloned into EcoRlHoxh of a modified pH-Stinger vector, which has attB sequences inserted into its Avrl site (for nGFP reporters) (H. Jin and M.F., unpublished) and pH-Pelican (for DocF4-s1-lacZ) (Barolo et al., 2000). For transformation, the landing site of M{3XP3-RFP.attP'}ZH-22A was used (Bischof et al., 2007) (Bestgene, Chino Hills, CA, USA). In DocF4-s1Hmut, the E-boxes CANNTG (positions 227, 293, 417, 436, 508) were mutated to AGNNTG via de novo DNA synthesis (Mr. Gene, Regensburg, Germany). In DocF4-s1Bmut, the Bin binding core motifs T/CAAA (positions 162, 265, 289, 444, 458) were mutated to T/CAGTCA and the CAACA at position 486 to CGGCCCA.

**Staining procedures**

Antibody stainings and fluorescent double stainings for proteins and mRNAs were performed as described (Azpiazu and Frasch, 1993; Knirr et al., 1999). The following antibodies were used: rabbit anti-Twi (1:2000, from S. Roth, Cologne University, Germany), guinea pig anti-Hb9 (Exelixis, FlyBase) (1:500, from J. Sneath, Washington University School of Medicine, USA), guinea pig anti-Eve (1:300) (Kosman et al., 1998), rabbit anti-Eve (1:3000, our own production), rabbit anti-Mef2 (1:750, from H. Jin and M.F., unpublished), and rabbit anti-Vasa (1:1000, our own production), guinea pig anti-Hb9 (Exelixis, FlyBase) (1:500, from J. Sneath, Washington University School of Medicine, USA), guinea pig anti-Eve (1:300) (Kosman et al., 1998), rabbit anti-Eve (1:3000, our own production), rabbit anti-Mef2 (1:750, from H. Jin and M.F., unpublished), and rabbit anti-Vasa (1:1000, from A. Ephrussi, EMBL Heidelberg, Germany and Ruth Lehmann, NYU)
RESULTS

HLH54F encodes a conserved bHLH transcription factor and is driven by an intronic enhancer in the CVM

Drosophila HLH54F encodes a bHLH transcription factor and sequence alignments of the bHLH domain show that it is orthologous to a single gene in Tribolium and to the vertebrate paralogs capsulin (also known as Tcf21, Pod-1, epicardin) and musculin (also known as Myor) (Fig. 1A,B). Within Drosophila, the closest bHLH domains belong to Hand, Fer1 and Twi (Fig. 1A,B). Notably, all of these Drosophila and vertebrate genes, with the exception of neuronally expressed Fer1, are prominently expressed in mesodermal tissues (Georgias et al., 1997; Hidai et al., 1998; Lu et al., 1998; Quaggin et al., 1998; Robb et al., 1998a; Robb et al., 1998b; Lu et al., 1999; Quaggin et al., 1999; Tomancak et al., 2002). To obtain a useful tool for examining the genetic function of HLH54F in the developing CVM of Drosophila embryos, which expresses HLH54F (Georgias et al., 1997), as well as for analyzing the development of this tissue in general, we aimed to generate lacZ reporter constructs that reflect this expression pattern.

HLH54Fb-lacZ, which spans the two introns (Fig. 1C), was found to recapitulate the endogenous HLH54F expression pattern in the CVM. Reporter gene activity first appeared during gastrulation in a small domain abutting the posterior-most border of the gastrulation furrow, which reflects the early expression of HLH54F mRNA at the posterior tip of the mesoderm anlage during late blastoderm stages (Fig. 2A,B). This group of cells is known as the CVM primordium (Campos-Ortega and Hartenstein, 1997; Kusch and Reuter, 1999). During germ band elongation, the CVM cells marked by HLH54Fb-lacZ and HLH54F mRNA were internalized and occupied a position within the posterior bend formed by the posterior midgut (PMG) rudiment (stage 9; Fig. 2C,D). Subsequently, the CVM cells from this cluster migrated anteriorly along the dorsal and ventral margins of the TVM and spread along its entire length (Fig. 2E,F; see also below). Eventually the CVM cells fused with fusion-competent cells from the TVM to form multinucleated longitudinal midgut muscles (Fig. 2H; compare with Fig. 2G).

The early steps of cell migration were examined in more detail by stainings of HLH54Fb-lacZ embryos in the context of the neighboring tissue primordia, namely the Twi-labeled trunk mesoderm and the Hb9-labeled endoderm of the PMG rudiment (Thisse et al., 1988; Broihier and Skeath, 2002). At stage 9, the...
HLH54F mutations generated by imprecise P excisions and EMS mutagenesis

To create null alleles of HLH54F a P element, EY06760, located downstream of HLH54F and just inside the 5’ UTR of CG5009 (Fig. 1C) was excised. Data from Mohr and Gelbart on deficiencies with defined breakpoints at 54EF indicated that loss of HLH54F may cause semi-lethality (Mohr and Gelbart, 2002). Indeed, two of our excision chromosomes that caused semi-lethality in trans with two of these deficiencies, Df(2R)14H10W-21 and Df(2R)02B10Y-12, carried deletions of HLH54F coding sequences (Fig. 1C, HLH54F^S1750 and HLH54F^S1750^598). Because the deletions in both alleles also remove a small portion of the 5’ UTR of CG5009, we generated a rescue construct encompassing the entire HLH54F locus (Fig. 1C). Our observation that this genomic construct can rescue the semi-lethality to full viability and also rescues the observed loss of CVM (see below) in a HLH54F^S508 background demonstrate that the phenotypes described herein with this allele are solely due to the loss of the HLH54F gene.

We also obtained two ethyl methane sulfite (EMS)-induced mutations in HLH54F from an ongoing forward genetic screen for mutations affecting the specification and/or migration of the CVM (see Materials and Methods). Both alleles, HLH54F^S0323 and HLH54F^S0323^598, are semi-lethal in trans to the P-excision allele HLH54F^S598 and show the same CVM phenotype. Whereas HLH54F^S0323 is a deletion of HLH54F and flanking genes, a point mutation in HLH54F^S0323^598 causes a change in a conserved Arg to a Trp within the DNA-binding basic domain (Fig. 1A). The equivalent residue (Arg33) of another bHLH protein, Max, is known to contact the phosphate backbone of the DNA (Ferre-D’Amare et al., 1993). Thus, based upon the molecular and genetic data, HLH54F^S0323 and the P-excision alleles appear to be functional nulls. Transheterozygotes of the P-excision alleles with the EMS alleles or with larger deficiencies at the HLH54F locus, in which the effects of any second site mutations are eliminated, occur as adult escapers at less than 10% of the expected frequency and cannot be grown as a stock.

HLH54F is required for proper specification and migration of the CVM

Based on its expression pattern, it was anticipated that HLH54F would play an important role in the development of the CVM and longitudinal midgut muscles. The genes for the Dorsocross T-box factors are normally expressed in the CVM between stages 10 and 12 (Fig. 4A,C,C‘) (Reim et al., 2003), but were not turned on in the caudal mesoderm in HLH54F mutant backgrounds (Fig. 4B,D,D’;
note that $\text{HLH54F}^{\text{S1170}}$ mRNA and $\text{HLH54Fb-lacZ}$ were still expressed, which rules out autoregulation and allows identification of the cells that normally become CVM. Likewise, $\text{bin}$, a FoxF gene that is normally expressed in all three types of visceral mesoderm (Fig. 4E) (Zaffran et al., 2001), failed to be expressed in the CVM in $\text{HLH54F}$ mutants, whereas it was still expressed in the TVM and hindgut visceral mesoderm (HVM) (Fig. 4F). Other markers of the migrating CVM, such as $\text{croc-lacZ}$, which reflects the CVM expression of the forkhead domain gene $\text{crocodile}$ (croc) (Häcker et al., 1995), the Ig domain-encoding gene $\text{beaten path IIa}$ (beat-IIa) (Tomancak et al., 2002), the enhancer trap expression of a sloppy paired (slp) reporter ($\text{slp-m4-lacZ}$) (Lee and Frasch, 2000), and expression of $\text{byn}$ after stage 10, were also absent in the corresponding areas of $\text{HLH54F}$ mutants (Fig. 4G-I; data not shown). The only marker of CVM primordia not affected in $\text{HLH54F}$ mutants was the high-level expression of the zinc-finger homeodomain transcription factor $\text{Zfh1}$ (see Fig. S1 in the supplementary material), although, as for $\text{byn}$, $\text{zfh1}$ expression also disappeared prematurely after stage 10. The loss of all tested CVM markers at all stages in the absence of $\text{HLH54F}$ (with the exception of early $\text{Zfh1}$) and the observation that the cells that would normally form CVM failed to undergo any rearrangements and movements (Fig. 4B,D,D’; compare with Fig. 4A,C,C’) suggest that $\text{HLH54F}$ is required for the specification of the CVM and show that development of the longitudinal visceral muscles is critically dependent on $\text{HLH54F}$ function.

It has been observed that the loss of cell fate specification can cause either cell fate transformation or cell death (reviewed by Bonini and Fortini, 1999; Mann and Morata, 2000; Werz et al., 2005). Hence, we used an $\text{HLH54Fb-lacZ}$ reporter to address the fate of the caudal mesodermal cells that are no longer identified as CVM in the absence of $\text{HLH54F}$. Following stage 10, we observed a dramatic decrease in the number of $\text{HLH54Fb-lacZ}$-positive cells and a total absence of migrating cells in the $\text{HLH54F}$ mutants (compare Fig. 4L with 4K). Apoptosis of these particular cells was demonstrated in a TUNEL assay, which showed TUNEL-positive cells containing traces of lacZ signals in the vicinity of the remaining $\text{HLH54Fb-lacZ}$-positive cells in stage 12 $\text{HLH54F}$ mutants (Fig. 4N). This is in contrast to the virtual lack of TUNEL signals in $\text{HLH54Fb-lacZ}$-positive cells in control embryos (Fig. 4M).
To examine whether any possible cell fate transformations or migratory abilities of the HLH54Fb-lacZ-positive cells in HLH54F mutant backgrounds are obscured by their death, we forced the expression of a pan-caspase inhibitor, p35 (BacA\textsuperscript{p35}), via UAS-p35 and a byn-GAL4 driver in these cells to prevent their apoptosis (Brand and Perrimon, 1993; Hay et al., 1994). Blocking apoptosis specifically in the CVM (along with the presumptive hindgut ectoderm) of HLH54F mutants resulted in a nearly normal-sized HLH54F-lacZ-positive domain (Fig. 4O; compare with Fig. 4C). However, these cells still did not migrate or differentiate into any kind of muscles and, instead, remained as unfused cells randomly distributed in the vicinity of the hindgut (Fig. 4P; data not shown). Thus, we infer that HLH54F is required not only for the survival of the CVM, but also for the initial specification and, in this context, for the correct migratory properties of these cells. In the absence of HLH54F activity, the failure of these cells to be specified apparently causes their entry into apoptosis.

The CVM primordia are thought to form the founder cells for all longitudinal visceral muscles. Apparently as a consequence of the absence of these muscles (see below), midgut constrictions were not formed efficiently and were present to a variable extent in the mutants, with a small fraction of mutant embryos lacking midgut constrictions altogether (Fig. 5A-C).

The larval longitudinal midgut muscles persist during metamorphosis and contribute to the longitudinal gut musculature of adult flies (Klapper, 2000). To address the question of whether the CVM is the sole source of founder cells for the longitudinal visceral muscles in larvae and adults, midguts from HLH54F mutant third-instar larvae and adult escapers, as well as from wild-type controls, were stained for F-actin and Troponymosin. In wild-type animals, these stainings highlight the transverse circular midgut muscles as well as the longitudinal visceral muscles (Fig. 5D,F,H). By contrast, longitudinal visceral muscles were present neither in HLH54F mutant larvae nor in adult flies (Fig. 5E,G,I). At this level of resolution, we did not detect any clear alterations of the larval circular gut muscles and the midgut that would be caused by the absence of the longitudinal gut muscles (Fig. 5E). However, midguts isolated from mutant adults tended to be more curled into a spiral shape and smaller in diameter than their straighter wild-type isolated from mutant adults tended to be more curled into a spiral hand.

As HLH54F was isolated originally as a binding partner of the E-protein Daughterless (Da) (Georgias et al., 1997), we sought to obtain indications as to whether these two proteins might functionally interact in vivo during longitudinal muscle development. Indeed, RNAi knockdowns of HLH54F and da via HLH54F-GAL4-driven expression of the respective inverted repeat sequences caused similar phenotypes in the adult longitudinal gut muscles, consisting of a variable reduction in the number and thickness of fibers as well as their uneven distribution around the midgut. In general, the effects appeared stronger upon knocking down both genes together (see Fig. S2 in the supplementary material). These data are indicative of functional interactions between these two bHLH proteins during longitudinal gut muscle development. Future studies need to address these interactions in more detail and compare them with the repressive effects on myogenesis reported for Da-Twi heterodimers (Castañon et al., 2001).

Based upon alterations in the mesodermal expression pattern of Even-skipped (Eve) and a disrupted Met2 pattern in myocardial cells in byn mutant embryos, it has been proposed that the CVM plays a non-autonomous role in the normal development of adjacent dorsal muscle and cardiac progenitors during its migration (Kusch and Reuter, 1999). However, with our HLH54F mutants, which cause the specific loss of the CVM, we did not see any of the reported effects on the specification and morphogenesis of the heart and dorsal muscles, which seems to rule out a role of the CVM in the development of these tissues (see Fig. S1 in the supplementary material). It has also been proposed that germ cells require the CVM to direct their early migration from the basal surface of the PMG in anterior directions along the mesoderm (Broihier et al., 1998). Our observations in HLH54F mutants confirmed that the CVM does make the guidance of germ cells towards the gonadal mesoderm more reliable, although it is not absolutely required for this migration (see Fig. S3 in the supplementary material).
HLH54F acts upon a CVM-specific Doc enhancer in combination with its downstream factor Bin

An exhaustive analysis of the Dorsocross (Doc) locus for enhancers (D.S., H. Jin, M.F. and I.R., unpublished) identified a 566 bp element, DocF4s1, that recapitulates the activity of the Doc genes (Fig. 6A,E,I). DocF4s1 contains five E-boxes and six forkhead domain binding motifs that are expected to bind HLH54F and Bin, respectively (Fig. 6I). Ectopic expression of HLH54F via twi-GAL4 caused an expansion of DocF4s1 enhancer activity within the caudal mesoderm as well as into the head mesoderm, and analogous expression of bin caused similar, albeit less pronounced, effects (Fig. 6B,C). Notably, co-expression of HLH54F and bin caused much stronger effects, indicating that the products of the two genes cooperate in activating DocF4s1. The absence of ectopic expression in the trunk mesoderm points to the requirement for additional, as yet unknown activating or repressive factors.

To test whether the cooperative inputs of HLH54F and Bin upon DocF4s1 require the respective binding motifs for these proteins, we tested mutated DocF4s1 derivatives. Mutation of all five E-boxes caused a delay in the onset of enhancer activity until mid stage 11 (instead of late stage 10; data not shown) and a mild reduction in the number of CVM cells expressing the reporter (Fig. 6F). Likewise, upon mutation of all six Bin binding motifs, reporter gene activity in the CVM was delayed and occurred in even fewer cells and at lower levels (Fig. 6G). Importantly, simultaneous mutation of the binding motifs for both HLH54F and Bin led to a total loss of enhancer activity in the CVM, whereas the unrelated activity in bilateral clusters within the procephalic neurogenic ectoderm was still present (Fig. 6H). Altogether, these data provide strong evidence for a mechanism in which HLH54F and Bin cooperate in activating CVM-specific Doc expression via direct binding to this enhancer in a partially redundant fashion. Because Bin activation in these cells requires HLH54F (Fig. 4F), this mechanism represents a feed-forward mode of regulation that might be responsible for the postponed onset of Doc expression in the CVM as compared with HLH54F.

Ectopic expression of HLH54F is insufficient to drive CVM development

To test whether HLH54F is sufficient to turn other mesodermal derivatives into longitudinal midgut muscles or their progenitors, we expressed the gene ectopically under the control of the derivatives into longitudinal midgut muscles or their progenitors.

Several genes are expressed in overlapping domains in the posterior region of the blastoderm embryo and are essential for the specification of various terminal structures (reviewed by Bronner and Jäckle, 1991; Lengyel and Iwaki, 2002). At the same time, Twi and Sna are expressed in the ventral-most region of the embryo and are required for mesoderm development (Kosman et al., 1991; Leptin, 1991; Ip et al., 1992). There are also regulatory interactions between the posterior gap genes and these ventrally expressed genes; in particular, the repression of sna by the terminal gap gene huckebein (hkb) defines the posterior border of the future mesoderm (Reuter and Leptin, 1994). As all these genes are known to play roles in the development of the CVM (Kusch and Reuter, 1999), we examined whether they act in this pathway via regulating the expression of HLH54F.

**HLH54F expression is regulated through a combination of posterior and ventral patterning genes**

How is the expression of HLH54F in a small quadrant of cells established in the posterior portion of the presumptive mesoderm? Several genes are expressed in overlapping domains in the

**Fig. 6. HLH54F and Bin as co-regulators of a caudal visceral mesoderm-specific enhancer from the Doc locus. (A-H) lacZ (A-D) or nuclear GFP (E-H) reporter gene expression from DocF4s1 and derivatives is shown in early stage 12 Drosophila embryos as detected with anti-β-gal (diaminobenzidine) and anti-GFP (fluorescence) antibodies (dorsal views). (A) DocF4s1-lacZ in a wild-type background showing expression in bilateral CVM clusters at the beginning of cell migration (right). Additional activity is seen in bilateral clusters of the procephalic neuroectoderm (left). (B-D) DocF4s1-lacZ in embryos with UAS-HLH54F (B), UAS-bin (C) and UAS-HLH54F + UAS-bin (D), driven pan-mesodermally by twi-GAL4. (E) Control DocF4s1-GFP expression. (F) DocF4s1 Hmut-GFP expression (all E-boxes mutated). (G) DocF4s1 Bmut-GFP expression (all Bin binding motifs mutated). (H) DocF4s1 H+Bmut-GFP (all HLH54F and Bin binding motifs mutated). (I) Map of the DocF4s1 enhancer within the Doc locus and of HLH54F (H) and Bin (B) binding motifs within the enhancer.
Initially, at the blastoderm stage, twi extends to the poles of the embryo and, therefore, overlaps with the CVM primordium. As shown in Fig. 7C, HLH54F mRNA was still expressed in the absence of twi function, albeit at reduced levels (see Fig. 7A). Furthermore, HLH54F expression was still observed in the migrating CVM in the twi mutant embryos (Fig. 7D; compare with Fig. 7B). These results indicate that HLH54F expression is not directly dependent on twi and confirm that the CVM does not require twi for ingestion and initial migration [although the trunk mesoderm is required for proper pathfinding (Kusch and Reuter, 1999)]. HLH54F is also expressed in specific somatic muscle progenitors (where we have not detected any phenotype to date) (Fig. 7B) (Georgias et al., 1997). As expected, this expression was missing in twi mutants (Fig. 7D). In contrast to twi mutants, in the absence of sna there was no HLH54F mRNA expression in the CVM primordium at any time (Fig. 7E,F).

In hkb mutants, HLH54F expression was expanded towards the posterior (Fig. 7G), leading to an increased number of CVM cells as they started to migrate (Fig. 7H; compare with Fig. 7B). These data are in agreement with the known repressive activity of hkb on posterior sna expression and the involvement of sna in the activation of HLH54F expression, as shown above (see Fig. 7O for hkb and sna expression domains).

tailless (tll) is expressed in the posterior ~18% of the embryo (Fig. 7O) and is required for various structures derived from this expression domain, including the anal pads, hindgut, Malpighian tubules and the CVM (Pignoni et al., 1990; Kusch and Reuter, 1999; Lengyel and Iwaki, 2002). Accordingly, in tll10 (a strong hypomorph), only trace amounts of HLH54F mRNA expression were observed at blastoderm stage, and during elongated germ band stages caudal mesodermal HLH54F expression was lost completely (Fig. 7I,J). tll has been shown to act upstream of byn and fkh. byn encodes a T-box transcription factor that is expressed in a posterior stripe and is required for development of the hindgut, anal pads and Malpighian tubules (Fig. 7O) (Kispert et al., 1994; Singer et al., 1996). This domain of byn expression is established through activation by tll and repression by hkb at the posterior pole, which confines byn expression to within ~7.5-15% egg length (Pignoni et al., 1990; Singer et al., 1996). fkh encodes a winged helix transcription factor that is expressed in a slightly smaller, more posterior domain than tll (0 to ~15% egg length) that will form the future PMG and hindgut (Weigel et al., 1989). byn and fkh mutants showed a similar reduction in size of the domain of HLH54F expression during blastoderm and elongated germ band stages (Fig. 7K-N). In double mutants for byn and fkh, the reduction of HLH54F was much more severe than with the single mutants, indicating that these two genes synergize in activating HLH54F (Fig. 7P,Q).

The zinc-finger homeodomain-encoding gene zfh1 is expressed preferentially in the CVM (and hematopoietic mesoderm) during early mesoderm development (Fig. 7R) (Lai et al., 1991; Broihier et al., 1998). In embryos lacking zfh1, HLH54F was expressed normally, but the migration of the CVM cells was aberrant (Fig. 7S).
Fig. 8. Regulatory interactions during caudal visceral mesoderm development. There is high-level expression of \textit{zfh1} in presumptive caudal visceral mesoderm. Although shown separately, the developmental outputs downstream of \textit{HLH54F} are intimately connected. Black arrows, gene regulation; white arrows, developmental regulation. cvm, caudal visceral mesoderm; lvm, longitudinal visceral muscle.

Altogether, these data show that \textit{HLH54F} expression is positively regulated by ventrally active \textit{sna} (which itself is excluded from the posterior ~7.5% of the embryo by \textit{hkb}) as well as by the synergistic action of the terminal genes \textit{byn} and \textit{fkh}.

**DISCUSSION**

\textit{HLH54F} is a key regulator in the CVM, a population of cells in which the bHLH gene \textit{twi} appears to have only minor functions. Although \textit{twi} is initially co-expressed with \textit{HLH54F} in these cells, it makes only a small contribution to activating \textit{HLH54F} expression, and the expression of both bHLH genes rapidly becomes mutually exclusive. Instead of \textit{twi}, the activation of \textit{HLH54F} in the CVM primarily involves the combined activities of mesodermal \textit{sna} and the terminal genes \textit{fkh} and \textit{byn} (Fig. 8). As \textit{sna} is generally thought to act as a ventral repressor of non-mesodermal genes in early mesoderm development, it will be interesting to determine whether the positive requirement for \textit{sna} in the activation of \textit{HLH54F} expression is direct, which would be unique to date. Alternatively, \textit{HLH54F} might be activated by high levels of nuclear Dorsal and repressed by lateral genes that are repressed by \textit{sna} ventrally. Along the anteroposterior axis, the posterior border of \textit{HLH54F} expression is apparently defined by the posterior expression border of \textit{sna}, which is delineated by the repressive action from \textit{hkb}. We propose that the anterior border of \textit{HLH54F} is determined by near-maximal threshold levels of \textit{tll}, the expression of which declines steeply in the area anterior to the \textit{HLH54F} domain (Pisarev et al., 2009). However, \textit{tll} acts largely indirectly, through the combined activities of its downstream genes \textit{byn} and \textit{fkh}, in activating \textit{HLH54F} (Fig. 8). The low residual levels of \textit{HLH54F} mRNA in \textit{fkh byn} double mutants suggest the involvement of direct inputs from additional posterior activities, possibly \textit{tll} or maternal \textit{torso}. From the data shown herein and elsewhere (Hemavathy et al., 1997; Kusch and Reuter, 1999), it appears that high-level expression of \textit{zfh1} in the CVM largely depends on \textit{tll} and \textit{sna}, whereas \textit{HLH54F} and \textit{zfh1} do not depend on one another.

Notably, neither \textit{twi} nor \textit{HLH54F} is required individually for the internalization of the CVM cells during gastrulation (Kusch and Reuter, 1999) (this report), although we cannot exclude a redundant function. We have shown that the posterior portion of the mesoderm, which includes the CVM and portions of the presumptive HVM, bends around during gastrulation to form a second, internal mesodermal layer. It is conceivable that this movement is a passive process brought about by the invagination of the PMG rudiment. However, for subsequent migrations of CVM cells from these positions, the activity of \textit{HLH54F}, but not \textit{twi}, is crucial. In addition, \textit{byn}, \textit{zfh1} and \textit{fkh} are required for normal migration after stage 10 (Kusch and Reuter, 1999) (this report). Whereas their respective functions are likely to be cell-autonomous, the observed requirement of \textit{twi} for normal pathfinding of CVM cells is likely to be due to the absence of the migration substrate normally formed from the trunk mesoderm.

Our genetic data show that, after the caudal mesodermal cells have ingressed in this manner, they do not develop any further in the absence of \textit{HLH54F} activity and undergo apoptosis. In the normal situation, \textit{HLH54F} is needed for the activation of several transcription factor-encoding genes at this stage, including \textit{bin}, \textit{croc} and the Doc genes. Although the functions of these genes in CVM development have not been defined, it is likely that they regulate specific aspects of CVM development downstream of, and perhaps in combination with, \textit{HLH54F}. The data from loss-of-function and ectopic expression analyses of \textit{HLH54F} show that this gene is essential, but not sufficient, for specification of longitudinal gut muscle founders. Parallel inputs, albeit less pervasive, appear to come from high-level \textit{zfh1}, which like \textit{HLH54F} is required for \textit{croc/croc-lacZ} expression (Kusch and Reuter, 1999). Altogether, we propose that \textit{HLH54F} is necessary for activating the vast majority of early CVM-specific genes, with one known exception being high-level \textit{zfh1}, and that \textit{zfh1}, \textit{byn} and \textit{fkh} in various combinations act together with \textit{HLH54F} to activate certain targets during the specification and early migration of CVM cells.

The continuous expression of \textit{HLH54F} in the CVM and longitudinal gut muscles suggests that this gene is not only required for specification, but is also directly involved in many other developmental processes, including the continued migration, myoblast fusion and differentiation of the CVM cells. Possible downstream targets of \textit{HLH54F} in the promotion of proper cell migration include \textit{beat-Ila}, which encodes an as yet uncharacterized membrane-anchored Ig domain protein, and the FGF receptor-encoding gene \textit{heartless} (\textit{htl}), which is known to be required for normal migration (Pipes et al., 2001; Mandal et al., 2004). In this context, it is interesting that the vertebrate orthologs of \textit{HLH54F} are expressed prominently in specific migrating populations of mesodermal cells as well. For example, \textit{musculin} is expressed in myoblasts at the myotomal lips that migrate into the developing limbs, and \textit{capsulin} is expressed in the migrating pro-epicardial cells (e.g. von Scheven et al., 2006). Therefore, it is possible that parts of the regulatory circuit in the control of cell migration have been conserved, even though they occur in different mesodermal cell types. \textit{Capsulin} is also expressed prominently in the splanchnopleura and tissues derived from it, including the developing smooth muscles of the stomach and gut (Hidai et al., 1998; Robb et al., 1998b; von Scheven et al., 2006). Therefore, \textit{HLH54F} and \textit{capsulin} might share some functions in the terminal differentiation of the respective gut musculatures in the different systems. Both Capsulin and Musculin have been characterized largely as repressors. However, their activity (and likewise that of \textit{HLH54F}) as repressors versus activators might well be context specific with respect to the particular enhancer, tissue or developmental stage in question and might depend on the relative concentrations of particular heterodimerization partners (Lu et al., 1999; Miyagishi et al., 2000; Castanon et al., 2001).
Based on the phenotype of bmy mutants, a role of the CVM in promoting midgut constrictions has also been proposed (Kusch and Reuter, 1999). The phenotype of HLH54F mutants confirms this effect, although we find that partial constrictions can frequently occur and that the effect is variable. We infer that the physical interactions between developing longitudinal and circular muscle fibers are necessary to provide the full force required for the efficient constriction of the midgut endoderm at the well-defined signaling centers. In the fully developed midgut, scanning electron microscopy image have revealed that the longitudinal fibers are tightly interwoven with the web-shaped circular fibers, which may explain the mechanical strength of this meshwork. Indeed, we find that the mechanical stability and integrity of the midgut, particularly in HLH54F mutant adults, are severely compromised.

In summary, HLH54F appears to sit at the top of the regulatory hierarchy of CVM development and is likely to fulfill additional key roles during the course of development of the CVM and the longitudinal gut muscles. Future efforts need to be directed towards dissecting additional downstream events that regulate the different steps of cell migration, myoblast fusion, morphogenesis and terminal differentiation of the longitudinal midgut musculature.

Acknowledgements
We thank the colleagues listed in the Materials and methods section as well as the Bloomington Stock Center (Indiana, USA), VDRC (Vienna, Austria) and DSHB (Iowa, USA) for providing fly stocks and antibody reagents. We gratefully acknowledge funding from the National Institutes of Health (NIDDK and NICHD) to M.F. and Deutsche Forschungsgemeinschaft (DFG) to M.F. and I.R. Deposited in PMC for release after 12 months.

Competing interests statement
The authors declare no competing financial interests.

Supplementary material
Supplementary material for this article is available at http://dev.biologists.orglookup/suppl?doi=10.1242/dev.046573/DC1

References
HLH54F specifies caudal mesoderm