Published in final edited form as:
Mech Dev. 2007 ; 124(9-10): 715–728. doi:10.1016/j.mod.2007.07.002.

Cloning and characterization of a novel *MyoD* enhancer-binding factor

Masakazu Yamamoto[1,+], **Christopher D. Watt**[2,+], **Ryan J. Schmidt**[2,#], **Unsal Kuscuoglu**[3,^], **Roger L. Miesfeld**[3], and **David J. Goldhamer**[1,*]

1 *Center for Regenerative Biology, Department of Molecular and Cell Biology, University of Connecticut, Storrs, Connecticut 06269*

2 *Department of Cell and Developmental Biology, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104*

3 *Department of Biochemistry and Molecular Biophysics, University of Arizona, Tucson, Arizona 85721*

Abstract

Glucocorticoid induced gene-1(*Gig1*) was identified in a yeast one-hybrid screen for factors that interact with the *MyoD* core enhancer. The *Gig1* gene encodes a novel C2H2 zinc finger protein that shares a high degree of sequence similarity with two known DNA binding proteins in humans, Glut4 Enhancer Factor and Papillomavirus Binding Factor (PBF). The mouse ortholog of PBF was also isolated in the screen. The DNA binding domain of Gig1, which contains TCF E-tail CR1 and CR2 motifs shown to mediate promoter specificity of TCF E-tail isoforms, was mapped to a C-terminal domain that is highly conserved in Glut4 Enhancer Factor and PBF. In mouse embryos, in situ hybridization revealed a restricted pattern of expression of *Gig1* that overlaps with *MyoD* expression. A nuclear-localized *lacZ* knockin null allele of *Gig1* was produced to study *Gig1* expression with greater resolution and to assess *Gig1* functions. X-gal staining of $Gig1^{nlacZ}$ heterozygous embryos revealed *Gig1* expression in myotomal myocytes, skeletal muscle precursors in the limb, and in nascent muscle fibers of the body wall, head and neck, and limbs through E14.5 (latest stage examined). *Gig1* was also expressed in a subset of *Scleraxis*-positive tendon precursors/rudiments of the limbs, but not in the earliest tendon precursors of the somite (syndetome) defined by *Scleraxis* expression. Additional regions of *Gig1* expression included the apical ectodermal ridge, neural tube roof plate and floor plate, apparent motor neurons in the ventral neural tube, otic vesicles, notochord, and several other tissues representing all three germ layers. *Gig1* expression was particularly well represented in epithelial tissues and in a number of cells/tissues of neural crest origin. Expression of both the endogenous *MyoD* gene and a reporter gene driven by *MyoD* regulatory elements was similar in wild-type and homozygous null $Gig1^{nlacZ}$ embryos, and mutant mice were viable and fertile, indicating that the functions of *Gig1* are redundant with other factors.

*Address Correspondence to: David J. Goldhamer, Center for Regenerative Biology and Department of Molecular and Cell Biology, University of Connecticut, Storrs, Connecticut 06269 Tel: (860) 486-8337; Fax: (860) 486-8809. E-mail: E-mail: david.goldhamer@uconn.edu.
[+]Should be considered co-first authors
[#]Present Address, Neuroscience Program, University of California Los Angeles School of Medicine
[^]Present Address, Department of Surgery, Stanford University School of Medicine

Keywords

MyoD; myogenesis; tendon precursor; lacZ knockin; gene targeting; core enhancer; transcription; Gig1; papillomavirus binding factor; Pbf; mouse

Introduction

The MyoD family of basic-helix-loop transcription factors regulates skeletal muscle determination and differentiation. *MyoD* and *Myf-5* serve partially redundant functions in establishing the skeletal muscle lineage, and embryos carrying mutations in both factors are devoid of skeletal muscle (Rudnicki et al., 1993). Considerable progress has been made in identifying the *cis* regulatory elements that control the expression of *MyoD* and *Myf-5*. In mammals, *MyoD* is regulated by two muscle-specific enhancers, the core enhancer positioned at −20 kb and the distal regulatory region (DRR) at −5 kb (Goldhamer et al., 1992; Tapscott et al., 1992; Asakura et al., 1995; Goldhamer et al., 1995). These enhancers have largely complementary functions that together regulate the wild-type pattern of *MyoD* expression throughout embryogenesis. Transgenic and knockout data have shown that the core enhancer regulates the initial expression of *MyoD* in skeletal muscle precursor cells, suggesting that this element is a direct or an indirect target of upstream signaling events that control early myogenesis (Kablar et al., 1999; Chen et al., 2001; Chen and Goldhamer, 2004). Very little is known, however, about the transcriptional pathways that directly regulate core enhancer activity.

We previously identified two adjacent linker-scanner mutations, LS4 and LS5, which define a 30-bp sequence (referred to here as Required Element 2; RE2) that is essential for core enhancer activity in all skeletal muscle lineages (Kucharczuk et al., 1999). DNase I protection and electromobility shift assays have demonstrated that RE2 is bound by nuclear proteins in vitro, although their identity is not known (Goldhamer et al., 1995; unpublished observations). Based on Transcription Element Search System analysis (TESS; Schug and Overton, 1997), RE2 does not contain consensus binding sites for the effectors of signaling pathways known to be important in early myogenesis, including the Shh, Wnt, Notch and BMP pathways (Munsterberg et al., 1995; Pourquie et al., 1996; Tajbakhsh et al., 1998; Borycki et al., 1999; Delfini et al., 2000), suggesting that novel factors bind RE2 and regulate core enhancer activity.

Using a yeast one-hybrid screen for factors that interact with RE2, we report the isolation of two related genes, an uncharacterized gene referred to as *Glucocorticoid induced gene-1* (*Gig1*), and *Pbf*, the mouse ortholog of the transcriptional regulator, Papillomavirus Binding Factor (PBF; Boeckle et al., 2002). *Gig1* was independently isolated in a screen for glucocorticoid-induced genes in thymocytes (Chapman et al., 1995). Gig1 and Pbf proteins share a high degree of amino acid identity with the human transcription factor, Glut4 Enhancer Factor (Oshel et al., 2000) and, together, represent a new family of zinc finger DNA binding proteins (Tanaka et al., 2004). Recently, Glut4 Enhancer Factor and PBF were independently isolated (referred to as HDBP1 and HDBP2, respectively) in a screen for transcriptional regulators of the human *Huntington's disease* gene (Tanaka et al., 2004).

In situ hybridization and analysis of *lacZ* expression in heterozygous $Gig1^{nlacZ}$ knockin embryos showed that *Gig1* is expressed in *MyoD*-expressing muscle precursor cells and differentiating myocytes/fibers as well as in several other tissues, with prominent expression in subset of tendon precursors and definitive tendon rudiments. Additionally, *Gig1* is expressed in the notochord, neural tube floor plate and apical ectodermal ridge, key signaling centers that control axial patterning and limb outgrowth. *MyoD* expression was not affected in $Gig1^{nlacZ}$

homozygous null embryos, suggesting that functions of *Gig1* in myogenesis are redundant, a possibility we are currently investigating.

Results

Isolation of RE2-interacting proteins

The RE2 sequence of the *MyoD* core enhancer, which corresponds to an essential region defined by linker-scanner mutants 4 and 5 (Kucharczuk et al., 1999), was used as bait in a yeast one-hybrid screen in order to identify core enhancer-interacting factors. An E11.0 mouse embryo cDNA library was used, as the *MyoD* core enhancer is transcriptionally active at this stage (Goldhamer et al., 1995; Kablar et al., 1999; Kucharczuk et al., 1999). Of 1.07×10^7 clones screened, 41 positive clones were isolated, 18 of which represented four highly related cDNA inserts. These four inserts conferred robust growth on selective media when re-transformed into the RE2 yeast reporter strain, but failed to confer growth to a yeast reporter strain based on the p53 binding site (data not shown). Preliminary analyses indicated that the remaining 23 clones were non-specific.

Sequence analysis showed that the four isolated cDNA inserts correspond to two related murine genes. Two of the inserts, isolated a total of ten times, correspond to the predicted murine ortholog of human *PBF* (*Pbf*, accession number XM354825) (Boeckle et al., 2002). The remaining two inserts, isolated a total of eight times, were derived from *Gig1* (Chapman et al., 1995). *Gig1* cDNA was initially identified in WEHI 7.2 thymocytes in a screen for glucocorticoid-induced genes using mRNA differential display, and a full-length cDNA was cloned by screening a WEHI 7.2 cDNA library (Genbank accession number AF292939). The overall organization of the Gig1 protein is similar to that of PBF/pbf (42% identity/53% similarity with human PBF) and human Glut4 enhancer factor (44% identity/52% similarity) (Oshel et al., 2000). Interestingly, mouse Glut4 enhancer factor was not isolated in the present screen, despite having a highly similar C-terminal DNA binding domain (see Discussion).

The *Gig1* cDNA represents a 13.9-kb mRNA transcript that includes a 1.7-kb open reading frame and an 11.8-kb 3′ untranslated region. The *Gig1* gene, which is located on chromosome 3 A1, is predicted to have 9 exons spanning 182.1 kb (see Fig. 4), based on sequence comparison to the cDNA. An additional mRNA species may be generated by alternative splicing, based on a recent cDNA entry in the EMBL database (BC083106). The mature transcript encodes a 566 amino acid protein with the following motifs/features (Fig. 1; also see Tanaka et al., 2004): a centrally located C2H2 zinc finger, a putative nuclear localization signal (PIPRKRK), a putative nuclear export signal (**M**DK**V**TAA**M**V**L**) that is conserved at key hydrophobic residues with a nuclear export signal in PBF (**M**DE**M**MAA**M**V**L**), a serine-rich region (SRR), two proline-rich regions (PRR), and a highly conserved C-terminal basic domain that has been shown to mediate DNA binding of Glut4 Enhancer Factor and PBF (Boeckle et al., 2002;Tanaka et al., 2004). The DNA binding domain includes TCF-E tail CR1 and CR2 motifs (KKCRARFGΩ+-X_4-WC-X_2-CRRKKKC-X-RΩΩ), which influence promoter specificity of TCF-E isoforms (Atcha et al., 2003).

Analysis of genomic and EST databases revealed predicted genes related to *Gig1* in a variety of other vertebrate species, including rat, chicken, frog, zebrafish, and pufferfish (data not shown). A single predicted gene related to *Gig1* is also present in *Drosophila* (CG11676; accession number NM141707).

Functional identification of the Gig1 DNA binding domain

Glut4 Enhancer Factor and PBF have been shown to bind to the core sequence, CCGG (Oshel et al., 2000; Boeckle et al., 2002; Tanaka et al., 2004), whereas RE2 of the *MyoD* core enhancer

contains the related sequence CCTG. In electromobility shift assays (EMSA), bacterially-purified GST-Gig1 fusion protein bound to the canonical CCGG sequence of the human papillomavirus P2 element (Boeckle et al., 2002), but not to the RE2 element (data not shown). This may reflect a lower affinity for the CCTG site, or the need for additional cellular co-factors for DNA binding. The smallest cDNA isolated in the one-hybrid screen included the last 285 bp of *Gig1* coding sequence (data not shown), suggesting that this region contains an RE2 recognition motif. As a functional test for Gig1 binding to RE2, and to map the RE2 recognition domain, a series of VP16 activation domain fusions were tested in 10T1/2 cells for their ability to activate a reporter system based on RE2 (Fig. 2A, B). All VP16-Gig1 fusion proteins were expressed and nuclear localized (Fig. 2D). The largest fragment examined, residues 394 to 566 of Gig1, exhibited sequence-specific activation of the RE2 reporter when fused to VP16 (Fig. 2C). Three smaller C-terminal fragments were tested to narrow the location of the putative DNA binding domain (Fig. 2B). No activation was observed with either pVP16 394–470 or pVP16 514–566, while pVP16 471–566 retained activity in this assay (Fig. 2C). Also, no significant activity was observed when any of the VP16 activation domain fusions were tested against a control reporter that lacks RE2 sites (UASR) or has linker sequences (4xLSM UASR) in place of RE2 (Fig. 2C), mutations that destroy enhancer activity in transgenic mice (Kucharczuk et al., 1999). Thus, the smallest identified RE2 recognition motif spans the last 96 residues of Gig1, which contains Conserved Region 3 (CR3; Tanaka et al., 2004) and its associated TCF-E CR1/CR2 motifs (Atcha et al., 2003). Although the CR3 region of Glut4 Enhancer Factor and PBF is sufficient to mediate sequence-specific DNA binding in EMSAs (Tanaka et al., 2004), the present data demonstrates that pVP16 514–566, which includes CR3, is not sufficient to mediate RE2-specific activation (Fig 2C). To determine whether the TCF-E CR1/CR2 motif was necessary for sequence-specific recognition, two mutations were introduced into Gig1 471–566 (Fig 2B; pVP16 TECR1mut and pVP16 TECR2mut) that were shown to inactivate TCF-E CR1/CR2 (Atcha et al., 2003). Neither of the Gig1 mutants was able to activate a RE2-based reporter (Fig. 2C), demonstrating that an intact TCF-E CR1/CR2 motif is necessary for RE2 sequence recognition.

Overview of *Gig1* expression

Gig1 was expressed at all stages of embryogenesis examined, from E9.5 through E14.5, as revealed by RT-PCR (data not shown) and whole mount in situ hybridization (Fig 3). Prominent staining was observed in somites, particularly in epaxial domains (Fig. 3A–C), in limb bud mesenchyme (Fig. 3B,C) and in the otic vesicle (Fig. 3A–C). Because of a poor signal-to-noise ratio with *Gig1* in situ probes in some tissues (Fig. 3; data not shown), and relatively low resolution of whole mount in situ methods, positive identification of a number of *Gig1*-expressing cell/tissue types was not possible. To characterize *Gig1* expression with greater resolution and reduced background, we produced an *nlacZ* knockin allele of *Gig1* ($Gig1^{nlacZ}$) in which the *nlacZ* gene encoding nuclear-localized β-gal followed by 3-tandem SV40 polyadenylation sequences replaced most of exon 1 coding sequence, creating a fusion with the first 5 amino acids of *Gig1* (Fig. 4A, B). Of 24 ES cell clones screened, 9 clones were correctly targeted; three clones were chosen for chimera formation, and two exhibited germ-line transmission. *Gig1* transcripts were not detected in $Gig1^{nlacZ/nlacZ}$ embryos by RT-PCR (Fig. 4C) using primers anchored in exon 2, or in exons 7 and 9, indicating the absence of read-through transcription past the *nlacZ* cassette. As all mouse *Gig1* EST database entries (ENSEMBL) include exons 7–9, it is unlikely that 5′-truncated *Gig1* transcripts are produced by internal promoter utilization. Collectively, these data indicate that $Gig1^{nlacZ}$ is a true null allele and that *nlacZ* expression provides an accurate readout of *Gig1* transcription and translation.

X-gal staining of $Gig1^{nlacZ/+}$ embryos revealed *Gig1* expression in derivatives of all three germ layers (Fig. 5; Table 1), with prominent representation in mesodermal derivatives, including

developing skeletal muscles and tendons (described in detail below), and in the notochord in the posterior region of the embryo (Fig. 5B, C, I). Ectodermal derivatives include key signaling centers in the embryo, such as the neural tube roof plate and floor plate (Fig. 5E), and the apical ectodermal ridge (AER) of the limb bud (Fig. 5B, C). Other ectodermal derivatives exhibiting staining include the otic vesicle epithelium (Fig. 5A–C, H), apparent motor neurons in the ventral neural tube (Fig. 5E), and skin epithelium, which is visible in the tail by E10.5 (Fig. 5I), and becomes prominent throughout the embryo by E14.5 (see Fig. 8D–H). Although marker studies are required for definitive identification, staining in the dorsal neural tube and mesenchymal cells subjacent to the ectoderm in the dorsal embryo (not shown), in sympathetic chain ganglia (Fig. 5G), in the walls of vascular elements in the trunk (Fig. 5F,G), and in mesenchymal cells of the developing heart (not shown), is consistent with *Gig1* expression in neural crest and a number of neural crest derivatives. Among endodermal derivatives, the thyroid rudiment (Fig. 5F) represents the most prominent *Gig1*-expressing tissue at the stages examined. Although a comprehensive study of *Gig1* expression in the adult has not been undertaken, we have observed prominent expression in several regions of the brain, including the cerebral cortex, dentate gyrus and olfactory bulbs (data not shown). Interestingly, despite robust expression in developing musculature of the embryo, we have not detected X-gal staining in mature or regenerating skeletal muscle in adult $Gig1^{nlacZ/+}$ mice (unpublished observations).

Gig1 expression in developing skeletal muscles and tendons

At E9.5–9.75 (Fig. 5A), X-gal staining was observed in the first 8–12 occipito-cervical somites, and followed an anterior to posterior sequence of activation (Fig. 5B, C), paralleling the gradient of somite formation and maturation. Expression of *Gig1* in somite myotomes was confirmed by co-expression of a MyoDGFP transgene (−24GFP) that faithfully recapitulates *MyoD* expression and marks all skeletal muscle regions (Fig. 6; data not shown). In this transgene (Fig 6A), GFP expression is directed by 24 kb of *MyoD* 5′ flanking sequences, which include the two known *MyoD* enhancers, the core enhancer (Goldhamer et al., 1992;Goldhamer et al., 1995) and distal regulatory region (Asakura et al., 1995). *lacZ* expression directed by identical sequences has been described in detail (Chen et al., 2001). At E10.5, X-gal staining was most prominent in the central myotomes of interlimb somites and was absent from the lateral hypaxial domain (Fig. 5B; Fig. 6C–E). At E11.5, staining remained most intense in the central myotomes, although weak expression was now observed in a portion of the hypaxial myotome (Fig. 5C; Fig. 6F–H). In situ hybridization for *Gig1* (Fig. 3) and *lacZ* (Fig. 5D) mRNA also showed predominantly epaxial expression. At later stages, intercostal muscles and epaxial deep back muscles were X-gal-stained (not shown). Of note, *Gig1* expression in skeletal muscle appears to be restricted to embryonic stages, as X-gal staining of mature and regenerating adult skeletal muscles has not been observed (data not shown).

In the limbs, expression of *lacZ* was first detected at E10–10.5 in mesenchymal cells located centro-medially that include portions of the pre-muscle masses (Fig. 5B; data not shown). Muscle forming regions located more proximally in the limb bud did not show X-gal staining (Fig. 5B). By E11.5, *lacZ* expression had extended to more proximal regions of the limb (Fig. 5C,D; Fig. 7A–C), and was observed in definitive muscle beds at E14.5 (Fig. 8G–I), although muscle-to-muscle variation in X-gal staining was observed, and not all muscle groups exhibited X-gal staining (Fig. 7A–C, G–I). In developing muscles, X-gal staining was observed in a broader domain than *MyoD* expression, and included both GFP-positive and GFP-negative cells within the muscle beds (Fig. 7D–I, data not shown). In this regard, the pattern of X-gal staining in the early limb bud was similar to the pattern of expression of *Scleraxis*, which encodes a bHLH transcription factor (Cserjesi et al., 1995) and is a marker for tendon and ligament precursors (Schweitzer et al., 2001;Brent et al., 2003;Fig. 8A–C). That the muscle-associated *Gig1* expression at E10.5 and E11.5 represents expression in tendon precursors is

suggested by the robust *lacZ* expression in definitive tendon rudiments at E14.5 (Fig. 8D–G). Interestingly, X-gal staining was observed in apparent tendon precursors and tendon rudiments in the distal limb, whereas proximal tendon elements were weakly stained or unstained (Fig. 8E–H). This contrasts with *Scleraxis* expression in apparently all tendon precursor populations and tendon rudiments along the proximo-distal axis of the limb (Fig. 8C;Schweitzer et al., 2001). In addition to *Gig1* expression in muscle and tendon lineages in the limb, *Gig1* expression was also observed in the skin ectoderm, with a "cap" of strongly stained ectoderm covering the tips of the digits, and was particularly intense in the developing footpad mesenchyme and overlying ectoderm (Fig. 8D, F, G). Although absent from pre-cartilage mesenchyme and definitive condensing cartilage of the limb skeleton, *lacZ* expression was localized to sites of joint formation in E14.5 embryos (Fig. 8D, E).

Scleraxis also is expressed in early tendon precursors of the axial skeleton, marking a newly defined somite sub-compartment, the syndetome (Brent et al., 2003). In contrast, somitic expression of *Gig1* was restricted to the myotomes (Fig. 5A–D; Fig. 6), and X-gal staining was not detected in axial tendons or their precursors, as identified by *Scleraxis* expression, in $Gig1^{nlacZ/+}$ embryos (Fig. 8A).

MyoD expression in $Gig1^{nlacZ/nlacZ}$ embryos

$Gig1^{nlacZ/nlacZ}$ mice are viable and fertile and preliminary analyses did not detect obvious developmental defects. Targeted deletion of the *MyoD* core enhancer results in delayed *MyoD* expression in branchial arch and limb myogenic populations (Chen and Goldhamer, 2004). To test whether *Gig1* is essential for core enhancer activity, we performed in situ hybridization for *MyoD* on $Gig1^{nlacZ/nlacZ}$ embryos. Comparison of heterozygous and homozygous null embryos revealed no differences in the pattern or intensity of in situ hybridization signals (data not shown). Expression of −24GFP also was not affected in $Gig1^{nlacZ/nlacZ}$ embryos. This lack of phenotype could reflect functional compensation between *Gig1* and *Pbf*, a possibility we are currently exploring.

Discussion

In this study, two highly related cDNAs, Gig1 and Pbf, were isolated in a yeast one-hybrid screen for factors that interact with a required element (referred to here as RE2) of the *MyoD* core enhancer (Goldhamer et al., 1995; Kucharczuk et al., 1999). *Pbf* is the mouse ortholog of the human gene PBF, which encodes a DNA binding factor that regulates papillomavirus gene expression (Boeckle et al., 2002), whereas *Gig1* is a heretofore uncharacterized gene that is induced in thymocytes following glucocorticoid treatment (Chapman et al., 1995). Gig1 and Pbf proteins share a high degree of sequence similarity with the human transcription factor, Glut4 Enhancer Factor, which has been implicated in the insulin-dependent regulation of the *GLUT4* gene (Oshel et al., 2000; Knight et al., 2003). Interestingly, Glut4 enhancer factor and PBF also interact with an essential neuronal specific element located in the promoter of the *Huntington's disease* gene (Tanaka et al., 2004). Collectively, Gig1, PBF/Pbf and Glut4 enhancer factor represent a new family of DNA binding proteins that are probably components of transcriptional pathways that regulate a number of developmental and physiologic processes.

The region of Gig1 responsible for DNA binding resides in the C-terminal region spanning residues 473–566, as these residues comprise the smallest Gig1 cDNA clone isolated in the one-hybrid screen, and sequence-specific DNA binding domains have been identified within the corresponding C-terminal portions of Glut4 enhancer factor and PBF (Tanaka et al., 2004). Further, Gig1 471–566 was sufficient to mediate RE2-specific reporter activation when fused to the VP16 activation domain. Although Gig1, like Glut4 enhancer factor (Oshel et al., 2000; Tanaka et al., 2004) and PBF (Boeckle et al., 2002; Tanaka et al., 2004) binds a core CCGG motif (data not shown), we have been unable to demonstrate *in vitro* binding of Gig1

to RE2, which contains the related sequence, CCTG. This may reflect the need for co-factors for high affinity binding of Gig1 to RE2, although we cannot rule out the formal possibility that Gig1 activates transcription indirectly, by heterodimerization with an unknown RE2 binding protein expressed in yeast and mammalian cells. In this regard, the single C_2H_2 zinc finger, which is not required for DNA binding, appears to function as a protein-protein interaction domain (unpublished observations). Additionally, the TCF-E CR1 and CR2 motifs, which are necessary but not sufficient for activation of the RE2 reporter, may influence DNA target specificity, either by mediating direct DNA binding or as protein interaction domains, as shown for specific TCF-E tail isoforms (Atcha et al., 2003; Hecht and Stemmler, 2003).

Gig1 expression was investigated by whole mount in situ hybridization for endogenous *Gig1* mRNA, and X-gal staining of $Gig1^{nlacZ/+}$ embryos. X-gal staining provided greater resolution and signal-to-noise ratio compared to *Gig1* in situ hybridization, and was the primary method used to evaluate *Gig1* expression. For areas of prominent *Gig1* expression in which comparisons can be made, such as somites, otic vesicles, limb bud mesenchyme, and nephrogenic mesoderm, in situ hybridization and X-gal staining showed good agreement, although staining patterns were not identical (compare Figs. 3 and 5). The pattern of expression revealed by X-gal staining is expected to reflect *Gig1* transcriptional and translational activity, as the vector design did not alter the *Gig1* initiation codon or 5′ untranslated sequences. Therefore, region-specific translational control could contribute to differences observed, although the close agreement of X-gal staining and in situ hybridization for *lacZ* mRNA (Fig. 5C, D) suggests that this is not a major factor. Alternatively, differential stabilities of *Gig1* and *lacZ* mRNAs could account for observed differences.

Gig1 is expressed in various tissues/embryonic regions representing all three germ layers. In addition to expression in developing skeletal muscle and tendons (see below), several features of *Gig1* expression are of particular note. First, *Gig1* expression was observed in key embryonic signaling centers, including the AER of the limb buds, the neural tube floor plate and roof plate, and the notochord in posterior regions of the embryo. Second, *Gig1* is expressed in a number of epithelial cell types with representatives from all three germ layers, including restricted expression in neural ectoderm, surface ectoderm, otic vesicle epithelium, lymphatic duct, pharyngeal epithelium, nephrogenic mesoderm and others (Table 1). Interestingly, the epithelial dermomyotome, from which the *Gig1*-positive myotome is derived, did not express *Gig1*. Third, neural crest derivatives are prominently represented among *Gig1*-expressing cell types (Table 1). Given the diversity of cell types that express *Gig1*, it is premature to speculate as to its functional significance in these disparate cells/tissues. We note, however, that preliminary data indicate that Gig1 has potent transcriptional repression activity and interacts with transcriptional co-repressors in yeast two-hybrid assays (unpublished observations).

Gig1 expression marks skeletal muscle precursors and differentiating skeletal muscles as shown by the temporal and spatial overlap with *MyoD* expression. However, distinct differences in *Gig1* and *MyoD* expression were noted. For example, while both genes were expressed in the myotomes at E9.5 to E9.75, *Gig1* was expressed in the anterior, older myotomes, whereas *MyoD* is first expressed in interlimb somites (Tajbakhsh et al., 1997; Chen et al., 2001). Further, at E10.5 and E11.5, *Gig1* expression is most pronounced in the epaxial myotomal domain, whereas *MyoD* is expressed throughout the myotomes, and exhibits prominent expression in the hypaxial domain of interlimb somites (present study; Goldhamer et al., 1995; Tajbakhsh et al., 1997; Chen et al., 2001). Finally, *Gig1* expression in the limbs shows pronounced regional variation, with expression most prominent in muscle precursor populations and early differentiating muscle masses located in the distal myogenic populations of the limb bud; *Gig1* expression was low or absent in several of the most proximal muscles of the limbs and associated girdles.

Gig1 expression in the limb mesenchyme showed significant overlap with *Scleraxis* expression, which marks tendon and ligament precursors in the chick and mouse (Schweitzer et al., 2001; Brent et al., 2003). Both genes are expressed in the proximo-medial and sub-ectodermal limb mesenchyme at early limb bud stages, show extensive overlap with *MyoD* expression but are expressed in a broader domain associated with the developing muscle masses, and are expressed in definitive tendon rudiments at later stages. Unlike *Scleraxis*, however, *Gig1* expression was most prominent in central and distal limb mesenchyme and definitive tendon rudiments and, in this regard, more closely resembles the expression of the distal tendon markers EphA4 and follistatin, which specifically mark autopod tendons in the chick (Patel et al., 1996; D'Souza and Patel, 1999). Further, *Scleraxis* appears to be an earlier tendon marker than *Gig1*, based on the presence of *Scleraxis* mRNA in tendon precursors associated with condensing digit cartilage at E12.5, prior to the onset of *Gig1* expression in these distal tendon precursors (Fig. 8). Also, *Gig1* expression was not detected in the syndetome, the newly defined somite sub-compartment comprised of tendon precursors that are marked by *Scleraxis* expression and which form the tendons associated with the axial skeleton (Brent et al., 2003). Collectively, these data suggest that *Gig1* expression marks a sub-population of tendon precursors restricted to the limb, particularly the distal limb. It will be important to determine whether these regional differences in *Gig1* expression in limb muscle and tendon precursors is a response to positional signals, or reflects intrinsic differences in these precursor populations (see Schafer and Braun, 1999).

Targeted deletion of the entire 258 bp core enhancer results in a delay in *MyoD* expression in branchial arches and limb buds (Chen and Goldhamer, 2004), yet the absence of Gig1, which binds a required region of the core enhancer, did not phenocopy this expression defect. Similarly, expression of the −24GFP transgene, which includes the core enhancer, was unaffected in $Gig1^{nlacZ/nlacZ}$ embryos (not shown). That $Gig1^{nlacZ}$ is a true null allele is indicated by RT-PCR analysis, which did not detect *Gig1* transcripts in $Gig1^{nlacZ/nlacZ}$ embryos using either exon 2 primers, or primers anchored in exons 7 and 9, which span sequences that encode the highly conserved C-terminal region of the protein. One possibility is that *Gig1* regulates subtle aspects of *MyoD* expression that were not detectable by the present expression analysis. Alternatively, functions of Gig1 may overlap with those of other factors, or such factors may functionally compensate when Gig1 is absent. Two obvious candidates are Pbf and Glut4 enhancer factor, given their high degree of structural relatedness with Gig1, particularly in the C-terminal DNA binding domain. Interestingly, while a probable mouse ortholog of human *Glut4 enhancer factor* gene has been identified in EST and genomic databases (not shown), the mouse gene does not appear to encode a full length functional protein due to the presence of numerous termination codons that interrupt the open reading frame (unpublished observations). The results of the yeast one-hybrid screen are consistent with this notion, as Glut4 enhancer factor was not represented among the 18 specific cDNA clones that were isolated, all of which represented either *Gig1* or *Pbf*. Gene targeted will address overlapping functions of *Gig1* and *Pbf*.

Experimental Procedures

Yeast one-hybrid screen

The MATCHMAKER one-hybrid screen system (BD Biosciences Clontech) was used to identify proteins that bind to RE2. RE2 consists of a 30-bp sequence defined by linker-scanner mutagenesis (LS4 and LS5; Kucharczuk et al., 1999) with an additional 5 bp of upstream and downstream flanking sequence. Eight tandem copies of the 40-bp element (5′-TCTGAGAGGGTAACTTTATCCTGCTTCTTTCAGCCAAGTA-3′), each separated by a 6-bp linker sequence (GGATCT; used for concatemerization) were cloned into the EcoRI and XbaI sites of pHis-1. Using homologous recombination, 8xRE2 pHis-1 was integrated into the

genome of YM4271 to generate a stable reporter strain. His^- media supplemented with 7.5 mM 3-amino-1, 2, 4-triazole (Sigma) was sufficient to suppress background growth of the 8xRE2 pHis-1 reporter strain. Using a LiCl/PEG transformation protocol (Gietz and Woods, 1998), the 8xRE2 pHis-1 reporter strain was used to screen an 11.0 dpc mouse embryo cDNA library (BD Biosciences Clontech) according to manufacturer's recommendations. Plasmid DNA was recovered from transformants (YEASTMAKER yeast plasmid isolation kit, BD Biosciences Clontech) that generated colonies greater than 2 mm in diameter on His^-/Leu^- media supplemented with 7.5 mM 3-amino-1, 2, 4-triazole after 6 days of incubation at 30°C. Isolated cDNAs that supported growth when retransformed into the 8xRE2 pHis-1 reporter strain were selected for further study.

Sequence analysis

cDNA inserts were sequenced at the University of Pennsylvania Department of Genetics DNA Sequencing Facility. Web-based BLAST searches were performed against either the Genbank or Ensembl database. PSORTII (http://psort.ims.u-tokyo.ac.jp/) was used to locate potential nuclear localization signals. TESS (http://www.cbil.upenn.edu/tess/) was used to analyze RE2 for potential transcription factor binding sites.

Cloning of full length *Gig-1*

WEHI 7.2 cells were treated with dexamethasone for 8 h, after which total RNA was prepared by a single-step guanidine thiocyanate lysis method and was enriched for polyadenylated RNA using an oligo(dT) cellulose column (GIBCO BRL). Universal RiboClone cDNA synthesis system (Promega) and oligo(dT) or *Gig1* specific internal 3′ primers were used to prepare cDNA. Size-fractionated and adaptor-ligated cDNA fragments were cloned into the Lambda Zap II vector (Stratagene) and packaged using Gigapack III Gold packaging extracts (Stratagene). cDNA library screening and plasmid excision were performed according to the manufacturer's suggestions. Identified cDNA inserts were then sequenced and assembled.

Generation of expression constructs

Full length *Gig1* coding sequence was amplified using the GC-RICH PCR system (Roche) with the following PCR primer pair: 5′-GTGAATTCCAGGCATGCAGG C-3′ (forward) and 5′-CGTGTCGACGCTCAGTCAATGAA-3′ (reverse). The PCR product was cloned into the EcoRI and SalI sites of pBluescript KS+, and then subcloned into pVP16 (BD Biosciences Clontech). Deletions were made using either available restriction enzyme sites or standard PCR methods. Point mutations were introduced into pVP16 471–566 by PCR-based site-directed mutagenesis with PfuTurbo (Stratagene) and the following PCR primer pairs: 5′-GGAGAAGGCGTAGCGCTCGCACTAGTTTATGGGATGGAAAACAGG-3′ (TECR1mut, forward); 5′-CCTGTTTTCCATCCCATAAACTAGTGCGAGCGCTACGCCTTCTCC-3′ (TECR1mut, reverse); 5′-GGTGCACCGCCGTAGCCCTGGCACTAGCCTGCCAGCGGTTCATTG-3′ (TECR2mut, forward); 5′-CAATGAACCGCTGGCAGGCTAGTGCCAGGGCTACGGCGGTGCACC-3′ (TECR2mut, reverse). Point mutations and correct reading frames were confirmed by DNA sequencing.

The wild-type reporter (4xLSM UASR) was produced by inserting four tandem copies of RE2 upstream of the *Goosecoid* minimal promoter (−104 *Gsc*; Watabe et al., 1995) in the luciferase reporter plasmid, pGL3 (Promega). The mutant reporter (4xLSM UASR) was constructed by replacing four copies of RE2 with four copies of an unrelated, transcriptionally inert, sequence (Kucharczuk et al., 1999). Three tandem GAL4 Upstream Activating Sequence elements (UAS, 5′-CGGAGGACAGTACTCCG-3′) were inserted between wild-type or mutant RE2 elements and the minimal promoter, to allow for the use of GAL4-VP16 as a positive control,

which consistently generated >100 fold activation with each of the reporters tested (data not shown).

Luciferase reporter assays

10T1/2 cells were plated at a density of 1.7×10^4 cells/well in 24 well plates 20 hours prior to transfection. Using FuGENE 6 (Roche), luciferase reporter plasmids (180 ng) were co-transfected with an expression plasmid (20 ng) and a SV40pRL internal control (2 ng). Competition between experimental and control plasmids was minimal using these quantities of DNA. Cells were harvested 48 hours after the transfection and reporter activity was measured with a Turner luminometer using the Dual Luciferase Reporter Assay System (Promega). Fold activation was calculated relative to the activity of the UASR control reporter co-transfected with an empty expression vector (pVP16). Each experimental condition was performed in duplicate and experiments were repeated three times. Calculations and standard statistical analysis were performed using Microsoft Excel X. Error bars represent two standard errors.

Transgenic and knockout constructs

The MyoDGFP (−24GFP) transgene was constructed by excising 24 kb of human *MyoD* 5′ flanking sequence from −24lacZ (Chen et al., 2001) by digestion with SalI and AscI (AscI linkers were added to the unique XhoI site at the 3′ end of human genomic sequences) and inserting into SalI and AscI (produced by linker addition to the unique Sma I site) sites in the polylinker of pEGFP-1 (Clontech). Prior to subcloning, the NotI site just 3′ of EGFP was destroyed by blunt-ending with Klenow, and a new NotI site was introduced by linker addition to the AflII site, which lies 3′ of SV40 poly sequences. The plasmid was double-banded in cesium chloride equilibrium gradients, excised from vector sequences with Not1 (a 5′ NotI site lies just 3′ of the SalI site used for subcloning), and purified as previously described (Goldhamer et al., 1995).

The *Gig1* targeting vector was designed to replace exon 1 sequences encoding amino acids 6–150 and the first 169 bp of intron 1, with NLS-*lacZ* (*nlacZ*) followed by three tandem SV40 termination/polyadenylation sequences (3pA) and a floxed PGKNeo cassette. pCITE-pA was generated by inserting a SV40 polyadenylation sequence (blunted BamHI-MluI fragment of pEGFP-C3 (Clontech)) into the blunted BamHI site of pCITE-2a (Novagen). The same blunted SV40 polyadenylation sequence was sequentially introduced into the blunted BamHI site of pCITE-pA to generate pCITE-2pA. An *nlacZ* cassette with three SV40 polyadenylation sequences was generated by inserting the two-tandem SV40 polyadenylation sequence from pCITE-2pA into the NotI site of pPD46.21 (Fire et al., 1990), to produce pPDnLacZ3pA. pPD5′ Gig1nLacZ3pA was constructed by inserting the *Gig1* 5′ UTR (366 bp) and the first 15 nucleotides of exon 1 (a blunted BanI fragment excised from the *Gig1* cDNA clone) into the blunted XbaI site of pPDnLacZ3pA. To generate PL452-EF, *Gig1* intron 1 sequences from nucleotides 170 to 550, were amplified by PCR with the following primers; 5′-TATGCTAGCCTGTCGGTGCTCAGAGG-3′ (for) and 5′-TATGTCGACTAAGTGCTTAGGATGTCCAG-3′ (rev), digested with NheI and HincII, blunted with Klenow, and inserted into the blunted BamHI site of PL452 (Liu et al., 2003). The 5′Gig1nLacZ3pA fragment from pPD5′Gig1nLacZ3pA was excised by digestion with SalI, and introduced into the SalI site of PL452-EF to generate the *Gig1* mini-targeting vector containing an nLacZ3pA-floxed PGKNeo cassette flanked by *Gig1* fragments that are homologous to the targeting site. Subcloning *Gig1* genomic DNA and insertion of the nLacZ3pA-floxed PGKNeo cassette into the retrieved *Gig1* genomic DNA were performed by a recombineering-based method (Liu et al., 2003). Two PCR products which correspond to the ends of 5′ and 3′ homology arms were amplified with the following primer sets; 5′ arm: 5′-ATAAGCGGCCGCAGGTTACATGTCATCTGTGAG -3′ (for) 5′-GTCAAGCTTCTTGCTCAGATTGCACAGGTC -3′ (rev); 3′ arm: 5′-

GTCAAGCTTGAGATATCCTGACACTGTATC -3′ (for) 5′-TCTACTAGTGATATCTGCTCACTGATAATC -3′ (rev), and digested with NotI/HindIII and HindIII/SpeI, respectively. A gap repair plasmid (PL253-ABYZ) was generated by inserting those digested PCR products together into the NotI/SpeI of PL253. To generate PL253-AtoZ plasmid, subcloning of 8.2kb *Gig1* genomic DNA was performed via gap repair with HindIII digested PL253-ABYZ, RP24–229P11 BAC DNA (Children Hospital Oakland Research Institute) and EL350 host bacteria as described previously (Liu et al., 2003). The final *Gig1* targeting vector was generated by recombineering using the NotI fragment of the *Gig1* mini-targeting vector, PL253-AtoH and EL350 host bacteria as described previously (Liu et al., 2003).

Production of transgenic and gene-targeted mice

MyoDGFP (−24GFP) transgenic mice were produced by the University of Pennsylvania Transgenic and Chimeric Mouse Facility by pronuclear injection of BL6SJLF2/J single-cell embryos. Stable lines were produced by crossing to FVB mice and homozygous lines established by interbreeding. Two lines were characterized, each of which faithfully recapitulated *MyoD* expression in skeletal muscle between E10.5 and E14.5, the developmental stages used in the present study (see Fig. 6A; data not shown). The pattern of GFP expression closely matched that of *lac*Z in −24lacZ embryos (Chen et al., 2001), which utilized identical regulatory elements.

ES cell electroporation and production of chimeras was performed by the University of Connecticut Gene Targeting and Transgenic Facility (GTTF). The Gig1nLacZ3pA targeting vector was linearized with NotI and electroporated into 129S6/C57BL6 hybrid ES cells (D1: established by GTTF). Screening of ES cell clones was performed by southern blot hybridization with the PCR-generated probes corresponding to sequences outside of the 5′ and 3′ homology arms. The following primers were used: 5′ probe, 5′-CCATGCGTACTAAATAAGTTCTT-3′ (for), 5′-AACTTCTCCACGGATCCCAG-3′ (rev); 3′ probe, 5′-CATTGCTGGTGAAGGTGTGG -3′ (for) 5′-TTGTCTCCGGCACAGAATGG-3′ (rev). Chimeric mice were produced from targeted ES cell clones by aggregation with CD1 embryos. Chimeric mice were crossed to $Hprt^{Cre/+}$ mice (Jackson Labs; Tang et al., 2002) to remove the floxed PGKNeo cassette, removal of which was confirmed by PCR using the following primers: 5′-CCTGCAGCCCAATTCCGATCATATTC-3′ (for); 5′-GACCTCGTCCAGCTCCCAAGTTCC-3′ (rev). Germ line transmission of the targeted allele was assessed by PCR for *lacZ* with a forward primer that lies within the *nlacZ* cassette (5′-CCGAAATCCCGAATCTCTATC-3′) and a reverse primer in intron 1 of Gig1 (5′-TTGGCTTCATCCACCACATAC-3′). Lines were maintained by breeding to FVB mice.

Embryos from at least two litters were analyzed for each developmental time point. For staging, noon on the day of the vaginal plug was considered E0.5. Embryo staging was confirmed by somite counts.

RT-PCR

Total RNA was extracted from appropriately staged mouse embryos using the RNeasy mini kit (QIAGEN), and treated with RQ1 DNase I (Promega). cDNA was generated using the ProtoScript First Strand cDNA Synthesis Kit according to the manufacturer's protocol (NEB). *Gig1* transcripts were detected using exon 2 primers, 5′-TTAAGCGGGAGATGACCTTC-3′ (for) and 5′-CTTTCCGCTTCTGCTCAAGA-3 (rev), generating a 230-bp product, and primers anchored in exon 7 (5′-AGTCACTTTCACTGGCGTTC-3′) and 9 (5′-CAGGCGGTGCACCACATGTC-3′), generating a 195-bp product. The *β-actin* PCR (280-

bp) was performed with primers 5′-TAGGCACCAGGGTGTGATGG-3′ (for) and 5′-GTACATGGCTGGGGTGTTGAA-3′ (rev).

Whole-mount in situ hybridization

Whole mount in situ hybridization was performed as described previously (Henrique et al., 1995) with the following modifications. Hybridization was performed in 50% formamide/5X SSC pH 4.5/2% SDS/2% Boehringer's Blocking Reagent/250 μg/ml yeast tRNA/100 μg/ml heparin overnight at 70°C. The hybridized embryos were incubated in Solution X (50% formamide/2X SSC pH 4.5/1% SDS) for 20 minutes at 70°C four times and in a 1:1 mixture of Solution X and MABT (100 mM maleic acid/150 mM NaCl/0.1% Tween-20, pH7.5) for 20 minutes at 70°C. For detection of *Scleraxis* mRNA after X-gal staining, rehydrated embryos were incubated in 1% NP-40/1% SDS/0.5% deoxycholate/50mM TrisHCl pH8.0/1mM EDTA/150mM NaCl two times for 20 minutes each, rinsed in PTW (PBS/0.1% Tween-20), incubated in 10 μg/ml Proteinase K/PTW for 20 minutes, and post-fixed in 4% paraformaldehyde (PFA)/0.2% glutaraldehyde/PTW for 20 minutes.

Digoxigenin-labeled riboprobes were synthesized according to the manufacturer's protocol (Roche). The *MyoD* probe corresponded to nucleotides 782 – 1747 of its cDNA (NM_010866). The *Gig1* probe corresponded to nucleotides 2088 – 3065 of 3′ UTR sequences (AF292939). The *Scleraxis* template represented its entire coding sequence.

Histology and X-gal histochemistry

For detection of both GFP and β-gal in whole-mounts, embryos were fixed with 2% PFA/0.25% glutaraldehyde in PBS (pH 7.4) for 30 to 60 minutes at 4°C, and rinsed in three changes of PBS over 30 minutes. Following incubated in X-gal staining solution (2 mM $MgCl_2$, 4 mM $K_3Fe(CN)_6$, 4 mM $K_4Fe(CN)_6$, 0.1% Triton X-100 and 0.1% X-gal) in PBS at 37°C overnight, embryos were rinsed in PBS and re-photographed with bright-field illumination.

For sequential detection of *lacZ* and *Scleraxis* mRNAs, embryos fixed in 1% PFA/PBS for 1hour at 4°C were processed for X-gal staining for 2 – 4 hours at 37°C, rinsed in PBS, postfixed in 4% PFA/PBS overnight at 4°C, and then processed for in situ hybridization for *Scleraxis* as above.

For detection of GFP and β-gal in cryosections, embryos were fixed in 2% PFA/0.25 % glutaraldehyde/PBS for 30 minutes or 1% PFA/PBS for 1 hour, rinsed with PBS and cryoprotected in 30% sucrose for 2 hours at 4°C. After embryos were embedded in O.C.T. (Tissue-Tek) and frozen in liquid nitrogen, 10 μm cryosections were collected using a Tape-Transfer System (CryoJane; Instrumedics), rinsed in PBS and incubated in 0.1 μg/ml 4′6′-diamidino-2-phenylindole dihydrochloride (DAPI)/PBS. Sections were photographed in 10% glycerol/PBS using epifluorescence, stained with X-gal as above, and re-photographed.

Standard methods were used for paraffin histology of X-gal-stained embryos.

Photography and images

Images were captured using a Hamamatsu C5810 color video camera and a Leica MZFLIII stereomicroscope (whole mounts) or Nikon E600 microscope (sections) with either bright-field or epifluorescence illumination. For co-detection of GFP and β-gal, GFP images were captured prior to X-gal staining to avoid quenching of the GFP fluorescence by the X-gal deposits. After re-photographing, GFP and X-gal images were aligned and merged using Adobe Photoshop. In some cases, the GFP signal was artificially colorized for improved visualization of co-localized GFP and β-gal signals.

Acknowledgements

We thank Shoko Yamamoto for technical assistance and members of the Goldhamer lab for critical review of the manuscript. Plasmids were kindly provided by Drs. Neal Copeland (PL253, PL452 and EL350) and Alan Rawls (Scleraxis). This work was supported in part by grants from the NIH to DJG (AR42644) and RLM (GM40738) and by the Jack Findlay Doyle II Charitable Fund to RLM.

References

Asakura A, Lyons GE, Tapscott SJ. The regulation of MyoD gene expression: conserved elements mediate expression in embryonic axial muscle. Dev Biol 1995;171:386–98. [PubMed: 7556922]

Atcha FA, Munguia JE, Li TW, Hovanes K, Waterman ML. A New beta -Catenin-dependent Activation Domain in T Cell Factor. J Biol Chem 2003;278:16169–75. [PubMed: 12582159]

Boeckle S, Pfister H, Steger G. A new cellular factor recognizes E2 binding sites of papillomaviruses which mediate transcriptional repression by E2. Virology 2002;293:103–17. [PubMed: 11853404]

Borycki AG, Brunk B, Tajbakhsh S, Buckingham M, Chiang C, Emerson CP Jr. Sonic hedgehog controls epaxial muscle determination through Myf5 activation. Development 1999;126:4053–63. [PubMed: 10457014]

Brent AE, Schweitzer R, Tabin CJ. A somitic compartment of tendon progenitors. Cell 2003;113:235–48. [PubMed: 12705871]

Chapman MS, Qu N, Pascoe S, Chen WX, Apostol C, Gordon D, Miesfeld RL. Isolation of differentially expressed sequence tags from steroid-responsive cells using mRNA differential display. Mol Cell Endocrinol 1995;108:R1–7. [PubMed: 7758820]

Chen JC, Goldhamer DJ. The core enhancer is essential for proper timing of MyoD activation in limb buds and branchial arches. Dev Biol 2004;265:502–12. [PubMed: 14732408]

Chen JC, Love CM, Goldhamer DJ. Two upstream enhancers collaborate to regulate the spatial patterning and timing of MyoD transcription during mouse development. Dev Dyn 2001;221:274–88. [PubMed: 11458388]

Cserjesi P, Brown D, Ligon KL, Lyons GE, Copeland NG, Gilbert DJ, Jenkins NA, Olson EN. Scleraxis: a basic helix-loop-helix protein that prefigures skeletal formation during mouse embryogenesis. Development 1995;121:1099–110. [PubMed: 7743923]

D'Souza D, Patel K. Involvement of long- and short-range signalling during early tendon development. Anat Embryol (Berl) 1999;200:367–75. [PubMed: 10460474]

Delfini M, Hirsinger E, Pourquie O, Duprez D. Delta 1-activated notch inhibits muscle differentiation without affecting Myf5 and Pax3 expression in chick limb myogenesis. Development 2000;127:5213–24. [PubMed: 11060246]

Fire A, White-Harrison S, Dixon D. A modular set of *lacZ* fusion vectors for studying gene expression in *Caenorhabditis elegans*. Gene 1990;93:189–98. [PubMed: 2121610]

Gietz RD, Woods RA. Transformation of Yeast by the Lithium Acetate/Single-Stranded Carrier DNA/PEG Method. Yeast Gene Analysis, Academic Press, Methods in Microbiology 1998;26:53–66.

Goldhamer DJ, Brunk BP, Faerman A, King A, Shani M, Emerson CP Jr. Embryonic activation of the myoD gene is regulated by a highly conserved distal control element. Development 1995;121:637–49. [PubMed: 7720572]

Goldhamer DJ, Faerman A, Shani M, Emerson CP Jr. Regulatory elements that control the lineage-specific expression of myoD. Science 1992;256:538–42. [PubMed: 1315077]

Hecht A, Stemmler MP. Identification of a Promoter-specific Transcriptional Activation Domain at the C Terminus of the Wnt Effector Protein T-cell Factor 4. J Biol Chem 2003;278:3776–85. [PubMed: 12446687]

Henrique D, Adam J, Myat A, Chitnis A, Lewis J, Ish-Horowicz D. Expression of a Delta homologue in prospective neurons in the chick. Nature 1995;375:787–90. [PubMed: 7596411]

Kablar B, Krastel K, Ying C, Tapscott SJ, Goldhamer DJ, Rudnicki MA. Myogenic Determination Occurs Independently in Somites and Limb Buds. Dev Biol 1999;206:219–231. [PubMed: 9986734]

Knight JB, Eyster CA, Griesel BA, Olson AL. Regulation of the human GLUT4 gene promoter: interaction between a transcriptional activator and myocyte enhancer factor 2A. Proc Natl Acad Sci U S A 2003;100:14725–30. [PubMed: 14630949]

Kucharczuk KL, Love CM, Dougherty NM, Goldhamer DJ. Fine-scale transgenic mapping of the MyoD core enhancer: MyoD is regulated by distinct but overlapping mechanisms in myotomal and non-myotomal muscle lineages. Development 1999;126:1957–1965. [PubMed: 10101129]

Liu P, Jenkins NA, Copeland NG. A highly efficient recombineering-based method for generating conditional knockout mutations. Genome Res 2003;13:476–84. [PubMed: 12618378]

Munsterberg AE, Kitajewski J, Bumcrot DA, McMahon AP, Lassar AB. Combinatorial signaling by Sonic hedgehog and Wnt family members induces myogenic bHLH gene expression in the somite. Genes Dev 1995;9:2911–2922. [PubMed: 7498788]

Oshel KM, Knight JB, Cao KT, Thai MV, Olson AL. Identification of a 30-base pair regulatory element and novel DNA binding protein that regulates the human GLUT4 promoter in transgenic mice. J Biol Chem 2000;275:23666–73. [PubMed: 10825161]

Patel K, Nittenberg R, D'Souza D, Irving C, Burt D, Wilkinson DG, Tickle C. Expression and regulation of Cek-8, a cell to cell signalling receptor in developing chick limb buds. Development 1996;122:1147–55. [PubMed: 8620841]

Pourquie O, Fan CM, Coltey M, Hirsinger E, Watanabe Y, Breant C, Francis-West P, Brickell P, Tessier-Lavigne M, Le Douarin NM. Lateral and axial signals involved in avian somite patterning: a role for BMP4. Cell 1996;84:461–71. [PubMed: 8608600]

Rudnicki MA, Schnegelsberg PN, Stead RH, Braun T, Arnold HH, Jaenisch R. MyoD or Myf-5 is required for the formation of skeletal muscle. Cell 1993;75:1351–9. [PubMed: 8269513]

Schafer K, Braun T. Early specification of limb muscle precursor cells by the homeobox gene Lbx1h. Nat Genet 1999;23:213–6. [PubMed: 10508520]

Schug, J.; Overton, G. TESS: Transcription Element Search Software on the WWW. Computational Biology and Informatics Laboratory, School of Medicine, University of Pennsylvania; Philadelphia: 1997. Technical Report CBIL-TR-1997-1001-v0.0

Schweitzer R, Chyung JH, Murtaugh LC, Brent AE, Rosen V, Olson EN, Lassar A, Tabin CJ. Analysis of the tendon cell fate using Scleraxis, a specific marker for tendons and ligaments. Development 2001;128:3855–66. [PubMed: 11585810]

Tajbakhsh S, Borello U, Vivarelli E, Kelly R, Papkoff J, Duprez D, Buckingham M, Cossu G. Differential activation of Myf5 and MyoD by different Wnts in explants of mouse paraxial mesoderm and the later activation of myogenesis in the absence of Myf5. Development 1998;125:4155–62. [PubMed: 9753670]

Tajbakhsh S, Rocancourt D, Cossu G, Buckingham M. Redefining the genetic hierarchies controlling skeletal myogenesis: Pax-3 and Myf-5 act upstream of MyoD. Cell 1997;89:127–38. [PubMed: 9094721]

Tanaka K, Shouguchi-Miyata J, Miyamoto N, Ikeda JE. Novel nuclear shuttle proteins, HDBP1 and HDBP2, bind to neuronal cell-specific cis-regulatory element in the promoter for the human Huntington's disease gene. J Biol Chem 2004;279:7275–86. [PubMed: 14625278]

Tang SH, Silva FJ, Tsark WM, Mann JR. A Cre/loxP-deleter transgenic line in mouse strain 129S1/SvImJ. Genesis 2002;32:199–202. [PubMed: 11892008]

Tapscott SJ, Lassar AB, Weintraub H. A novel myoblast enhancer element mediates MyoD transcription. Mol Cell Biol 1992;12:4994–5003. [PubMed: 1328870]

Watabe T, Kim S, Candia A, Rothbacher U, Hashimoto C, Inoue K, Cho KW. Molecular mechanisms of Spemann's organizer formation: conserved growth factor synergy between Xenopus and mouse. Genes Dev 1995;9:3038–50. [PubMed: 8543150]

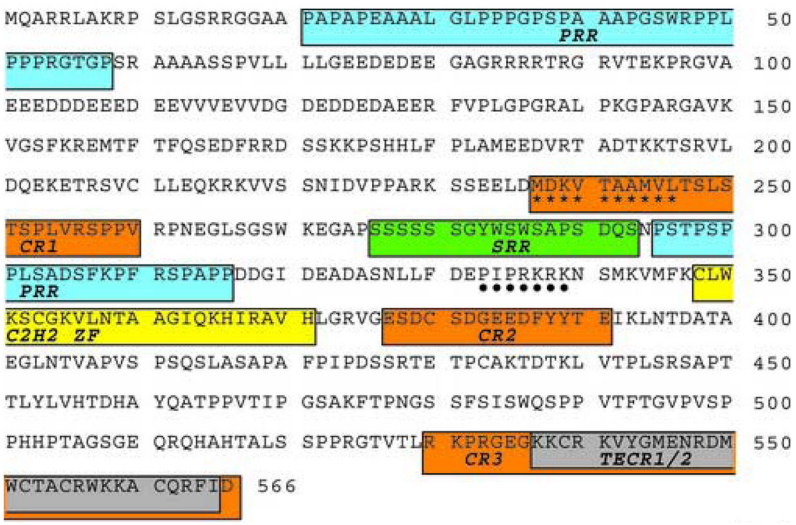

Fig. 1. Gig1 protein features

The predicted amino acid sequence of Gig1 is shown with the following protein features indicated: C2H2 zinc finger (*yellow*), putative NLS (*dots*), putative NES (*asterisks*), serine-rich region (*green*), two proline-rich regions (*blue*), TCF-E CR1 and CR2 motifs (Atcha et al., 2003; *gray*) and three previously defined conserved regions, CR1, CR2 and CR3 (Tanaka et al., 2004) shared between Gig1, Pbf and GEF (*orange*). The C-terminal DNA binding domain and the TCF-E CR1 and CR2 motifs are located within CR3.

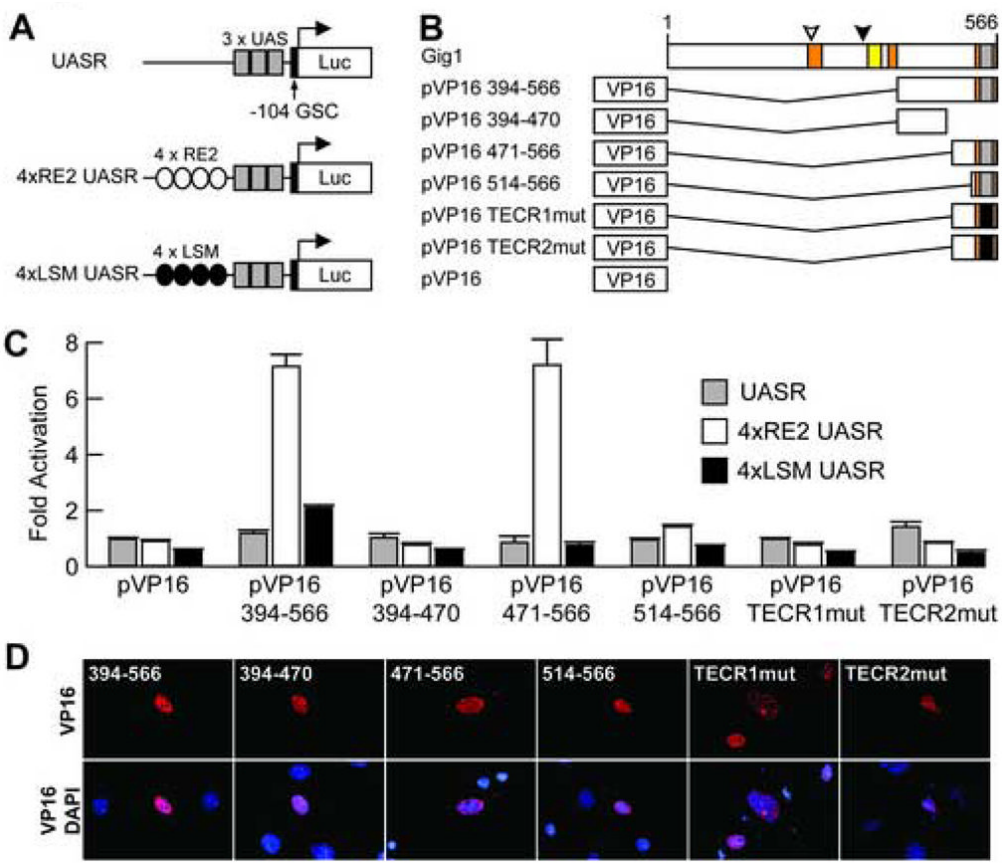

Fig. 2. The Gig1 DNA binding domain maps to a conserved C-terminal region that includes TCF-E CR1 and CR2 motifs

(**A**) Schematics of the luciferase reporters. UASR contains three tandem copies of the Gal4 UAS cloned upstream of a minimal *goosecoid* promoter (GSC). 4xRE UASR and 4xLSM UASR have four tandem copies of wild-type and mutant RE2, respectively, cloned upstream of UAS sequences. (**B**) Schematics of the VP16 AD-Gig1 fusions used to map the RE2 recognition motif. The relative locations of the C2H2 zinc finger (*yellow*), putative NLS (*closed arrowhead*), putative NES (*open arrowhead*), TCF-E CR1 and CR2 motifs (*gray*), and the three conserved regions, CR1, CR2 and CR3 (*orange*), are indicated. The mutant TCF-E CR1/CR2 motif is shown in black. The TCF-E CR1 motif (pVP16 TECR1mut) was changed from KKCRKVYGME to **VALAL**VYGME. The TCF-E CR2 motif (pVP16 TECR2mut) was changed from WCTACRWKKACQRFI to WCTA**VALAL**ACQRFI. (**C**) The average of three independent luciferase assays done in duplicate in 10T1/2 cells is shown. Fold activation corresponds to the ratio of the normalized activity of a test plasmid to the normalized activity of the UASR reporter challenged with pVP16. **(D)** An anti-VP16 antibody was used for immunofluorescence detection of VP16 AD-Gig1 fusions in transfected 10T1/2 cells. All fusions were expressed and nuclear localized (red; top and bottom panels). Cells stained only for DAPI in the bottom panel serve as untransfected controls.

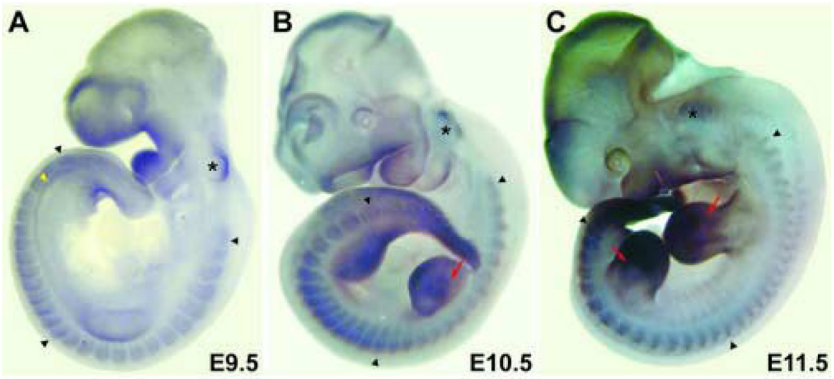

Fig. 3.

In situ hybridization detection of Gig1 mRNA. At mid-gestational stages, Gig1 is prominently expressed in somites (arrowheads in A–C), particularly in epaxial domains. Expression is also observed in limb bud mesenchyme (red arrows, B and C) and the AER (also see Fig. 6), in the otic vesicle (asterisk, A–C) and in the nephrogenic mesoderm in the posterior of the embryo (yellow arrowhead, A). Other areas of expression were identified by *lacZ* expression in *Gig1*$^{lacZ/+}$ knockin embryos.

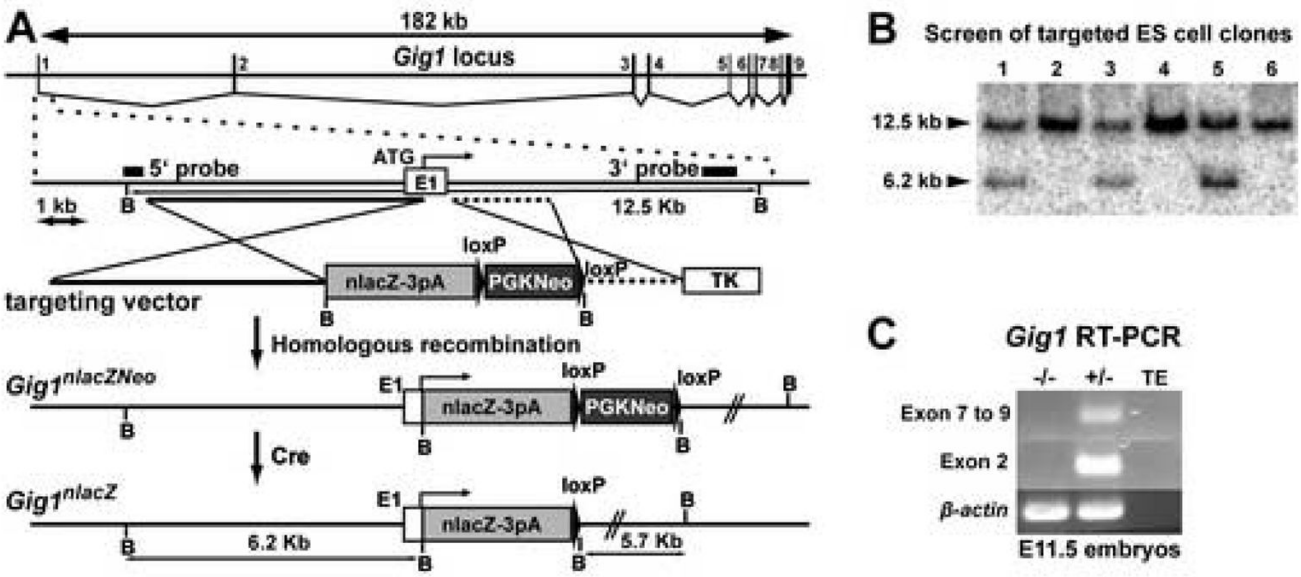

Fig. 4. *Gig1* locus and targeting strategy

(**A**) The *Gig1* gene spans approximately 182 kb and encodes a nine-exon 13.9-kb mRNA with a 1.7-kb open reading frame. An *nlacZ* cassette was targeted to exon 1, 15 bp downstream from the predicted start of translation, producing an in-frame fusion with *Gig1*. 5′ and 3′ homology arms are shown as a bold and dotted line, respectively. The floxed PGKNeo cassette was removed by cre-mediated recombination by crossing $Gig1^{nlacZNeo/+}$ mice with $HPRT^{cre}$ mice (Tang et al., 2002). **(B)** Representative southern blot with the 5′ probe showing targeted and non-targeted ES cell clones. The targeted allele generates a 6.2-kb band after digestion with BamH1 due to the introduction of a BamH1 site in the nlacZ-3pA cassette as shown in panel A (bottom). **(C)** RT-PCR analysis for *Gig1* mRNA. No mRNA was detected in $Gig^{nlacZnlacZ}$ embryos with primers anchored in exon 2, or in exons 7 and 9. *β-actin* was used as an internal standard.

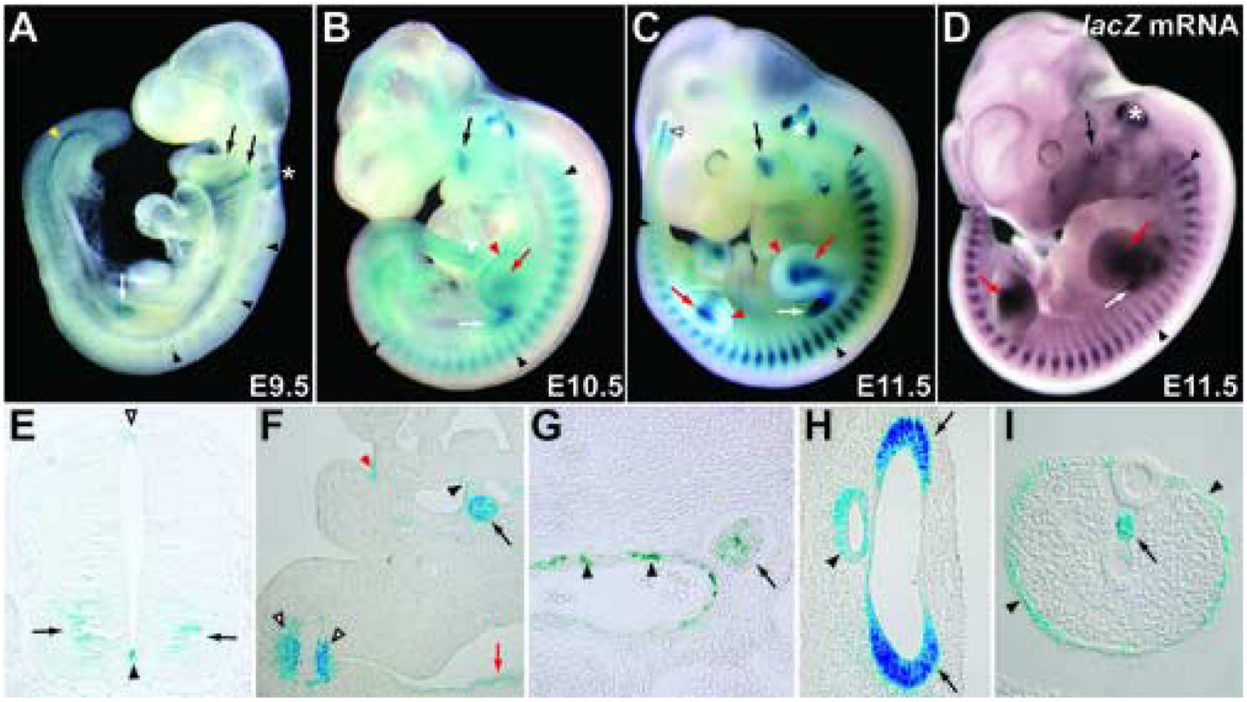

Fig. 5. Overview of *Gig1* expression at E9.5–11.5. (A–C. E–I)

X-gal stained $Gig1^{nlacZ/+}$ embryos. **(D)** $Gig1^{nlacZ/nlacZ}$ embryo processed for whole mount in situ hybridization for *lacZ* mRNA. **(A)** At E9.5, the anterior somites (black arrowheads), branchial arches (black arrows), otic vesicle (asterisk), and the flank mesenchyme (white arrow) just posterior to the emerging forelimb buds are X-gal stained. In the posterior of the embryo, the nephrogenic mesoderm is strongly positive (yellow arrowhead), and the notochord is weakly stained (not shown). **(B)** In addition to expression domains observed at E9.5 (labeling is the same as in A), X-gal staining of E10.5 embryos revealed *lacZ* expression in the centro-medial limb bud mesenchyme (red arrow), in the AER of the forelimb (red arrowhead) and hindlimb buds, and in the posterior notochord (white arrowhead). X-gal staining of somites extends to the hindlimb bud level and is most prominent in the epaxial domain. **(C)** At E11.5, limb bud staining includes more proximal regions of the limb buds. In somites, staining remains most prominent in the epaxial domain. Labeling is the same as in B. **(D)** The distribution of *lacZ* mRNA matched closely the X-gal staining pattern, with the exception that a greater signal was observed in the proximal forelimb and hypaxial myotomes. **(E–I)** Paraffin sections of X-gal stained E11.5 embryos. **(E)** X-gal staining in the ventral neural tube (arrows), roof plate (white arrowhead) and floor plate (black arrowhead). **(F)** Robust expression in the thyroid rudiment (arrow) and probable branchial arch-derived muscle precursors (white arrowheads). Cells of vascular wall (black arrowhead; also see panel G), the pharyngeal epithelium (red arrow), and surface ectoderm between branchial arches 2 and 3 (red arrowhead) are stained. **(G)** Higher magnification of the dorsal aorta showing X-gal-positive cells (arrowheads). Sympathetic chain ganglia are also X-gal-stained (arrow). **(H)** Otic vesicle epithelium (arrows) and associated endolymphatic duct (arrowhead) are X-gal stained. **(I)** X-gal staining of the notochord in the posterior embryo (arrow). The epidermis in the posterior embryos is also stained at this stage (arrowheads).

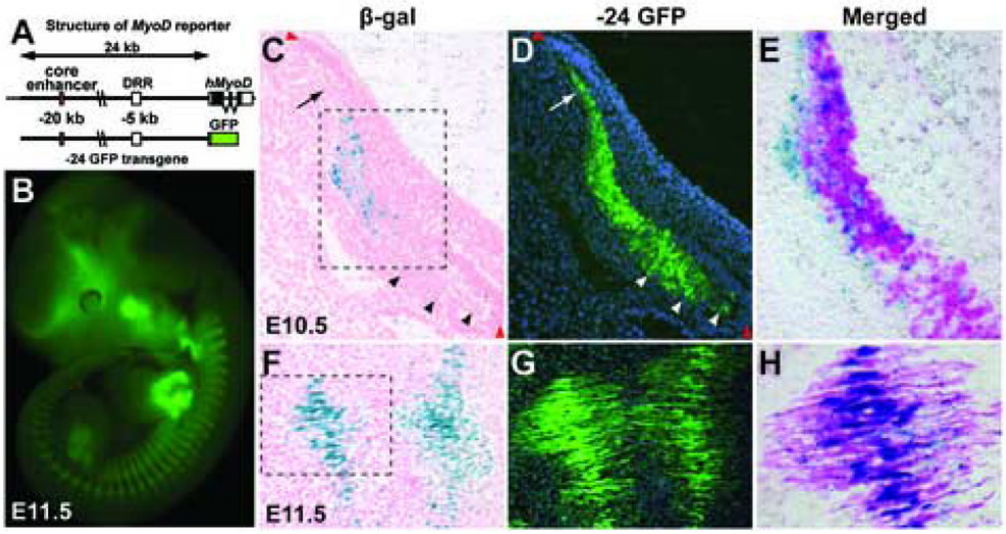

Fig. 6. *Gig1* is expressed in a subset of *MyoD*-positive myotomal cells

(A) Schematic representation of MyoDGFP (−24GFP) transgene. **(B)** The −24GFP transgene is expressed in all muscle-forming regions at E11.5, closely matching endogenous *MyoD* expression at this stage. Ectopic expression in the brain is likely the consequence of the site of transgene integration. **(C–E)** Representative transverse cryosection through an interlimb somite of an E10.5 $Gig1^{nlacZ/+}$ embryo. A subset of GFP+ epaxial myotomal cells are positive for β-gal (C–E). At E10.5, a sub-population of β-gal+ GFP − cells lies subjacent to the GFP+ myotome (E). The fluorescent signal in (D) was colorized red in (E) to better show X-gal and GFP co-localization. Staining is absent from the dorso-medal-most aspect of the myotome (arrow), the dermomyotomal lips (red arrowheads), and most of the hypaxial myotome (unstained hypaxial myotome is delineated by black or white arrowheads). **(F–G)** Sagittal section through the epaxial myotomes of two interlimb somites at E11.5. The X-gal and GFP signals appear qualitatively distinct because the $Gig1^{nlacZ}$ allele produces nuclear localized β-gal. **(H)** Image corresponding to the boxed area of (F) showing co-localization of β-gal and GFP signals. At E11.5, essentially all β-gal+ cells also express GFP. Sections were counterstained with DAPI.

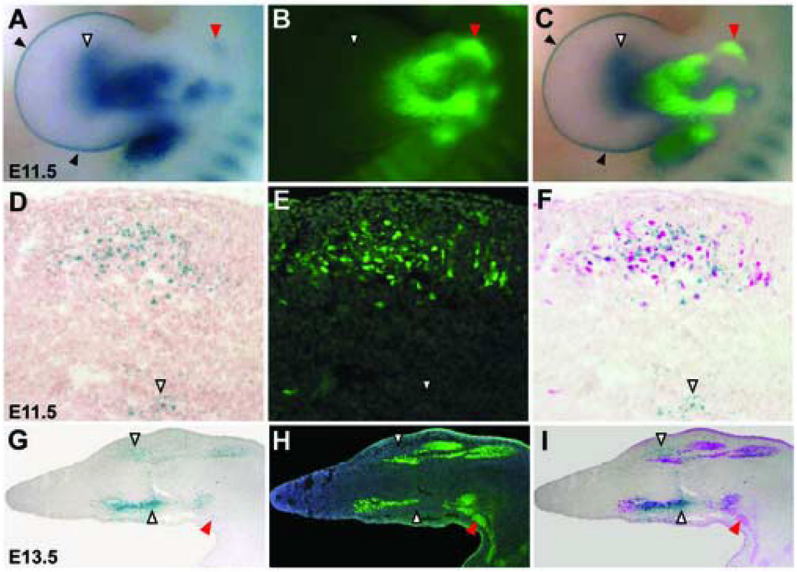

Fig. 7. *Gig1* and *MyoD* expression overlap in limb buds

(A–C) X-gal staining of a $Gig1^{nlacZ/+}$ whole mount forelimb bud at E11.5 showing overlap of X-gal staining and −24GFP transgene expression. The *Gig1* expression domain extends more distally than the domain of myogenesis at this stage (white arrowheads). Not all *MyoD*-positive muscle groups in the proximal limb bud are X-gal-positive (red arrowhead). Staining of the AER is visible in A and C (black arrowheads). **(D–F)** Transverse cryostat section through the mid-hindlimb bud showing a pre-muscle mass co-expressing *lacZ* and the −24GFP transgene. X-gal staining in pre-muscle masses was observed in both GFP-positive and GFP-negative cells (not shown). Small clusters of X-gal-positive cells in the central limb bud not associated with GFP-positive muscle precursors were also observed (white arrowhead). **(G–I)** A longitudinal section through an E13.5 hindlimb bud showing the relationship between β-gal and GFP expression domains. The X-gal signal associated with the muscle beds includes non-muscle mesenchyme (white arrowheads). Not all developing muscles in the proximal limb bud and body wall are X-gal-positive (e.g. red arrowhead). (A, D, G) bright-field images; (B, E, H) fluorescence images; (C, F, I) merge of X-gal and GFP signals. In F and I, the GFP signal was colorized red to better show GFP and X-gal co-localization. Sections were counterstained with DAPI.

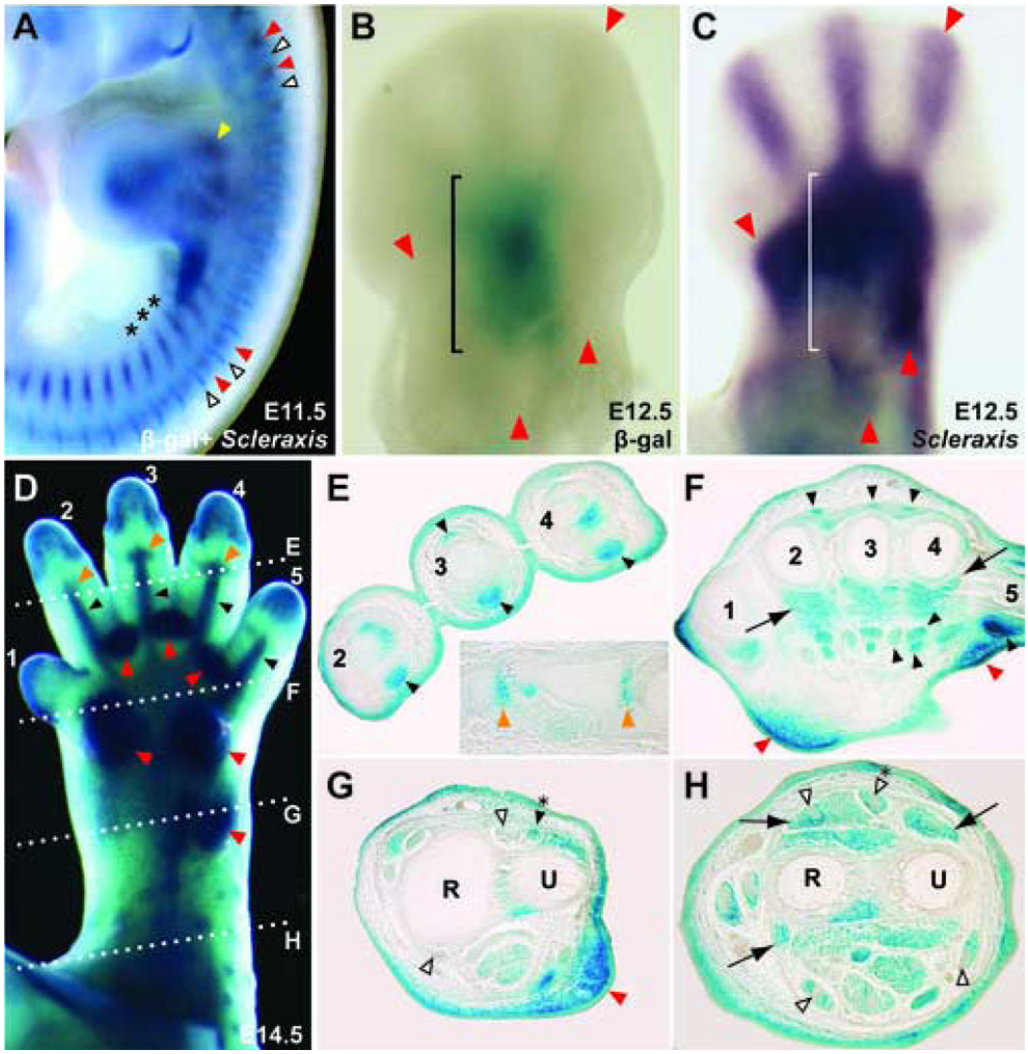

Fig. 8. *Gig1* is expressed in a subset of tendon precursors and definitive tendon rudiments in the limb

(A–C) Co-detection of X-gal and *Scleraxis* mRNA in *Gig1*$^{nlacZ/+}$ embryos. **(A)** In the trunk at E11.5, somitic X-gal (blue; white arrowheads) and *Scleraxis* (purple-brown; red arrowheads) signals are non-overlapping. *Scleraxis*, but not *Gig1*, is expressed in rib pre-cartilage (asterisks). In the limb, *Scleraxis* (yellow arrowhead) is expressed more proximally than *lacZ* (see panels B and C). **(B, C)** In the hindlimb bud at E12.5, X-gal staining (B) is restricted to the centro-medial portion of the limb bud (bracket), overlapping with *Scleraxis* expression (C), but is absent from more lateral, distal and proximal areas of *Scleraxis* expression (red arrowheads). The bracket in (C) corresponds to the limits of *Gig1* expression shown in (B). **(D)** Whole mount X-gal stain of the forelimb at E14.5. Ventral view is shown. X-gal staining is robust in the distal tendon rudiments (black arrowheads), developing footpads (red arrowheads), digital joints (orange arrowheads) and in the skin capping each of the digits (1–5). Dotted lines E–H correspond to the approximate position of the transverse paraffin sections shown in panels E–H. **(E)** Section through digits 2–4, showing intense X-gal staining in developing ventral tendons (arrowheads). A dorsal tendon (arrowhead) is also stained. The inset is a horizontal section of digit 3, showing X-gal staining of the joints (orange arrowheads). **(F)** Section through the proximal digits showing staining of ventral and dorsal tendons (e.g. arrowheads), developing musculature (e.g. arrows), and mesenchyme and overlying ectoderm

of the developing footpads (red arrowheads). **(G)** Section through the distal-radius/ulna. Some (e.g. black arrowhead) but not all (white arrowheads) tendon rudiments are stained at this level. An example of a tendon rudiment that is stained at this level, but not at more proximal levels (H), is marked with an asterisk. **(H)** In the proximal forelimb, tendon rudiments are not stained (white arrowheads). Developing muscles (e.g. black arrows) are stained. R, radius; U, ulna. In (E–H), dorsal is oriented toward the top of each panel.

Table 1

Major Sites of *Gig1* Expression at E11.5

Ectodermal Derivatives
Neural tube
Roof plate
Floor plate
Ventral neural tube[1]
Surface ectoderm[2,3]
Neural retina
Otic vesicle epithelium
Endolymphatic duct
Cervical sinus
Apical ectodermal ridge
Mesodermal Derivatives
Skeletal muscle (myotomes, muscle precursors of limb and head)
Tendon precursors
Notochord (posterior region of embryo)
Nephrogenic mesoderm
Flank mesenchyme just posterior to forelimb buds
Vascular wall[3]
Heart
Endocardial cushion mesenchyme[3]
Truncus arteriosus cushion mesenchyme[3]
Endodermal Derivatives
Thyroid rudiment
Gut epithelium (regional)
Neural Crest
Sympathetic chain ganglia

[1] Position of motor neurons

[2] More prominent at later stages

[3] May be of neural crest origin; marker studies required for unambiguous identification.