© 2014. Published by The Company of Biologists Ltd | Development (2014) 141, 4285-4297 doi:10.1242/dev.110908

RESEARCH ARTICLE

Mesogenin 1 is a master regulator of paraxial presomitic mesoderm differentiation

Ravindra B. Chalamalasetty[1], Robert J. Garriock[1], William C. Dunty, Jr[1], Mark W. Kennedy[1], Parthav Jailwala[2], Han Si[2] and Terry P. Yamaguchi[1,*]

ABSTRACT

Neuromesodermal (NM) stem cells generate neural and paraxial presomitic mesoderm (PSM) cells, which are the respective progenitors of the spinal cord and musculoskeleton of the trunk and tail. The Wnt-regulated basic helix-loop-helix (bHLH) transcription factor mesogenin 1 (Msgn1) has been implicated as a cooperative regulator working in concert with T-box genes to control PSM formation in zebrafish, although the mechanism is unknown. We show here that, in mice, Msgn1 alone controls PSM differentiation by directly activating the transcriptional programs that define PSM identity, epithelial-mesenchymal transition, motility and segmentation. Forced expression of Msgn1 in NM stem cells *in vivo* reduced the contribution of their progeny to the neural tube, and dramatically expanded the unsegmented mesenchymal PSM while blocking somitogenesis and notochord differentiation. Expression of *Msgn1* was sufficient to partially rescue PSM differentiation in *Wnt3a*$^{-/-}$ embryos, demonstrating that Msgn1 functions downstream of Wnt3a as the master regulator of PSM differentiation. Our data provide new insights into how cell fate decisions are imposed by the expression of a single transcriptional regulator.

KEY WORDS: bHLH transcription factor, Paraxial mesoderm, PSM, Differentiation, EMT, Motility, Wnt, Embryonic stem cell, Somite, Mouse

INTRODUCTION

The process of vertebrate body formation provides an excellent opportunity to understand the mechanisms driving stem cell differentiation because the body axis extends progressively during development and is sustained by a stem cell population found at the posterior end of the embryo. Neuromesodermal (NM) stem cells initially reside near the node and in the caudal lateral epiblast adjacent to the primitive streak (PS), and contribute to the somites and neural tube of the entire trunk and tail (Nicolas et al., 1996; Cambray and Wilson, 2007; Tzouanacou et al., 2009; Wilson et al., 2009). On the path to formation of the trunk musculoskeleton, NM stem cells must undergo several necessary processes and developmental transitions. Proper entry into the presomitic mesoderm (PSM) cell fate requires that epithelial NM stem cells undergo ingression in the PS, undertake epithelial-mesenchymal transition (EMT) to become mesoderm, adopt PSM cell identity, undergo lateral cell motility and finally repress neural cell fate. The segmentation clock is a characteristic trait of PSM cells as it is an essential component of a clock-and-wavefront mechanism that periodically divides the anterior PSM into the discrete segmental units known as somites (Dequéant and Pourquié, 2008). In mice, the PSM is continuously maintained at the posterior end by the addition of new paraxial mesoderm progenitors furnished by the PS (or the tailbud at later stages) until ~65 pairs of somites have been generated. These somites serve as the building blocks of the trunk and tail, forming the muscle, bone, cartilage and dermis. As so many separate processes are associated with acquisition of a PSM cell fate, it is reasonable to expect complex and diverse transcriptional control over these processes; however, how these processes are coordinated remains unknown.

Wnt signaling plays a pivotal role in the formation of the PSM. Loss-of-function mutations in *Wnt3a* or *Ctnnb1* (β-catenin) lead to a loss of paraxial mesoderm progenitors and their derivatives, the PSM and somites, whereas conditional gain-of-function mutations that stabilize β-catenin in the mesoderm result in a large expansion of the PSM (Takada et al., 1994; Cano et al., 2000; Dunty et al., 2008). Several transcriptional regulators lie downstream of Wnt signaling in the PS and, therefore, are excellent candidates to regulate PSM formation (Dunty et al., 2014). These include the T-box transcription factors brachyury (hereafter referred to by its symbol *T*) and *Tbx6*, and the basic helix-loop-helix (bHLH) transcription factor mesogenin 1 (*Msgn1*) (Yamaguchi et al., 1999; Wittler et al., 2007; Dunty et al., 2008; Chalamalasetty et al., 2011). Combinatorial expression of these factors defines nascent paraxial mesoderm (*T*$^{+}$*Tbx6*$^{+}$*Msgn1*$^{+}$) or differentiated PSM (*Tbx6*$^{+}$*Msgn1*$^{+}$). Null mutations in these genes lead to posterior phenotypes that bear similarities to the *Wnt3a*$^{-/-}$ phenotype, as they all display a loss of posterior somites and PSM (Fujimoto and Yanagisawa, 1979; Chapman and Papaioannou, 1998; Yoon et al., 2000; Yoon and Wold, 2000). However, gain-of-function studies on Tbx6 in mice (Wehn and Chapman, 2010) and T in embryonic stem cells (ESCs) (Lindsley et al., 2008) do not reveal strong mesoderm-inducing activity, raising the question as to what is the major positive regulator(s) of PSM fate.

In zebrafish, the core pathway known to initiate PSM identity involves co-regulation by Msgn1 and the Tbx6-related T-box transcription factor spadetail (Fior et al., 2012; Yabe and Takada, 2012), indicating a role for multiple transcription factors in PSM cell fate. However, we present evidence that, in mice, Msgn1 is necessary to maintain the PSM and might be sufficient for PSM specification. Using molecular, genetic and genomic approaches, we show that Msgn1 directly activates a gene regulatory network that orchestrates PSM differentiation, EMT and motility. Forced expression of Msgn1 in ESCs and *in vivo*, resulted in cells adopting PSM fates, demonstrating that Msgn1 is a master regulator of PSM differentiation.

RESULTS

Msgn1 is necessary and sufficient for PSM differentiation

Previous studies of *Msgn1*$^{-/-}$ embryos that analyzed the expression of the PSM marker *Tbx6* have been contradictory (Yoon and Wold, 2000; Nowotschin et al., 2012). Given that *Tbx6* is expressed in two

[1]Cancer and Developmental Biology Laboratory, Center for Cancer Research, NCI-Frederick, NIH, Frederick, MD 21702, USA. [2]CCRIFX Bioinformatics Core, Leidos Biomedical Research, FNLCR, Frederick, MD 21702, USA.

*Author for correspondence (yamagute@mail.nih.gov)

Received 1 April 2014; Accepted 3 September 2014

Development (2014) 141, 4285-4297 doi:10.1242/dev.110908

related populations of posterior mesoderm – the nascent paraxial mesoderm emerging from the PS (Fig. 1A,B, arrowhead) and the more anteriorly located differentiated PSM (dashed ovals, Fig. 1B) – we re-examined *Tbx6* expression in *Msgn1*$^{-/-}$ embryos at a stage [E8.5, 4-8 somite stage (ss)] that allowed us to distinguish these two populations (Fig. 1C,D). We found that *Tbx6* expression was dramatically reduced in *Msgn1*$^{-/-}$ embryos (Fig. 1C), particularly in the PSM domain (cf. dashed ovals, Fig. 1B,D). By contrast, *T* expression in immature paraxial mesoderm progenitors was expanded in the *Msgn1*$^{-/-}$ mesoderm compared to wild-type embryos (Fig. 1E-H). These results demonstrate that *Tbx6* expression is Msgn1-dependent and suggest that both Msgn1 and Tbx6 are necessary for PSM differentiation.

To determine whether *Msgn1* or *Tbx6* is sufficient to drive PSM differentiation, we generated inducible Tet-ON ESC lines to conditionally express *Msgn1* or *Tbx6*. Recombinant ESCs expressing a doxycycline (Dox)-inducible FLAG epitope-tagged Msgn1 (iF-Msgn1) were cultured in suspension for 2 days to initiate embryoid body (EB) formation and epiblast differentiation, before addition of Dox. Dox-mediated induction of *Msgn1* at this timepoint represents premature, ectopic expression because endogenous *Msgn1* is not expressed until day 4 of differentiation (supplementary material Fig. S1A-C). EBs were treated with Dox for 6-72 h before cells were harvested for RT-PCR analysis. Expression of *Msgn1* in differentiating EBs resulted in the rapid activation of the PSM gene *Tbx6*, and the anterior PSM and somite marker *Pax3*, but not the intermediate mesoderm marker *Osr1* (Mugford et al., 2008) (Fig. 1I). Pdgfrα is a cell surface marker of mesenchymal paraxial mesoderm progenitors (Sakurai et al., 2006; Takebe et al., 2006), and, together with Flk1 (Kdr – Mouse Genome Informatics), marks progenitors with broad mesodermal potential (Ema et al., 2006; Lindsley et al., 2008). Fluorescence-activated cell sorting (FACS) analysis of iF-Msgn1 EBs treated with Dox demonstrated that the induction of Msgn1 resulted in the rapid and preferential cell surface expression of Pdgfrα. Whereas only 1.36% of control cells were Pdgfrα^{+}, 62.8% were Pdgfrα^{+} after only 12 h of Dox treatment (arrowheads, supplementary material Fig. S1D). A similar percentage of Msgn1-expressing cells were Pdgfrα^{+} after 24 h of Dox administration, compared with 4.92% in untreated cells (arrowheads, Fig. 1J). A fivefold increase in Pdgfrα^{+}/Flk1^{+} double-positive cells was also detected in Msgn1-expressing cells (3.53%, compared with 0.69% in controls). Importantly, overexpression of FLAG-tagged Tbx6 (F-Tbx6) in EBs (supplementary material Fig. S1E-G) did not induce

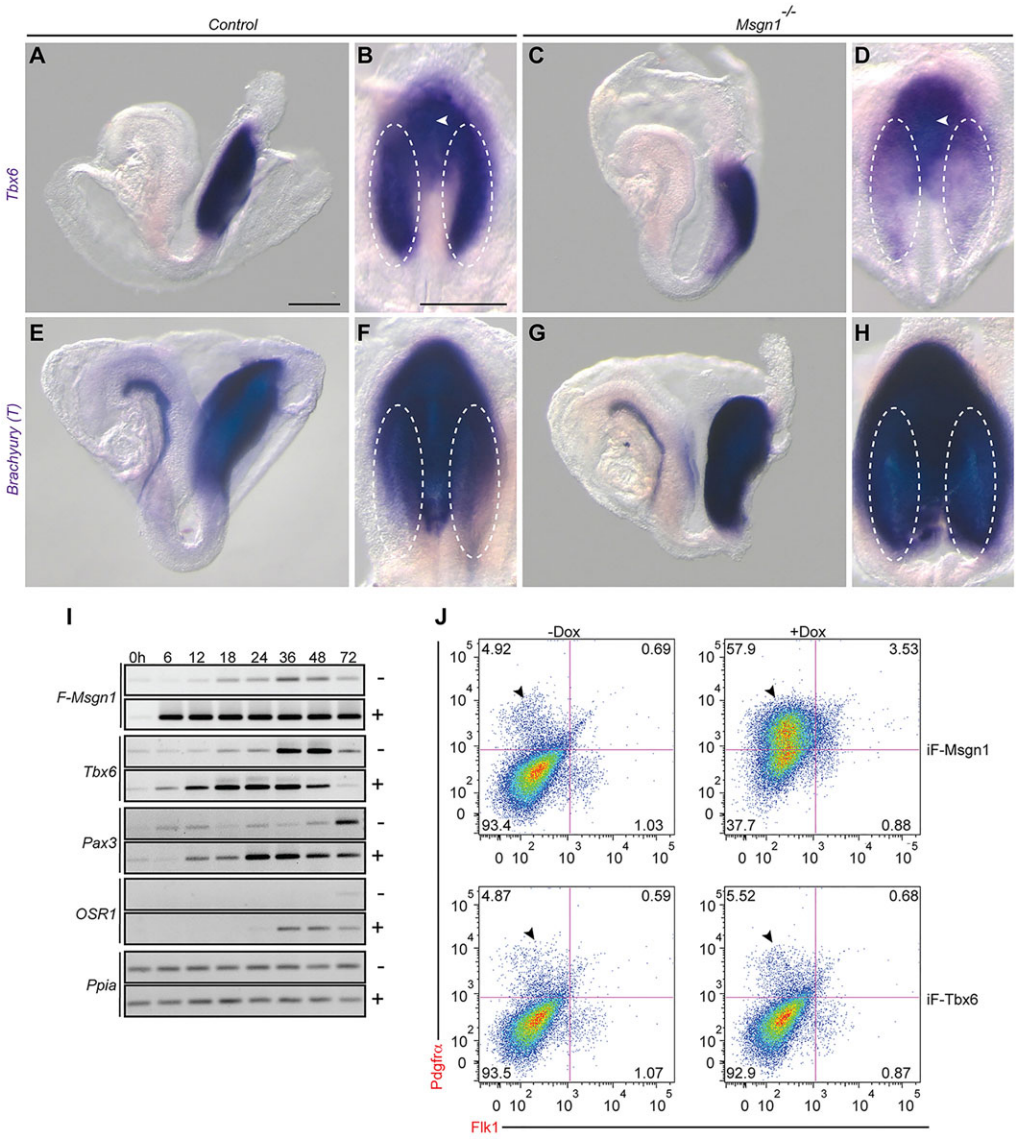

Fig. 1. Msgn1 promotes paraxial mesoderm formation. (A-H) WISH analysis of *Tbx6* (A-D) and *T* (E-H) mRNA expression in E8.5 control (A,E, lateral view; B,F ventral view) and *Msgn1*$^{-/-}$ (C,G, lateral view; D,H, ventral view) embryos. Scale bars: 200 µm. (I) iF-Msgn1 EBs treated with (+) or without (–) Dox from 0-72 h are analyzed for the expression of *Tbx6*, *Pax3* and *Osr1* by semi-quantitative RT-PCR. Cyclophilin A (*PpiA*) is used as a loading control. (J) Density plots of flow cytometry analysis for the expression of Pdgfrα and Flk1 in iF-Msgn1 and iF-Tbx6 EBs treated with (+) or without (–) Dox for 24 h. Numbers indicate the percentage of cells within each quadrant. Arrowheads indicate the section containing Pdgfrα-expressing paraxial mesoderm cells.

Development (2014) 141, 4285-4297 doi:10.1242/dev.110908

these mesoderm markers (Fig. 1J). These results suggest that Msgn1 is sufficient to elicit PSM differentiation from ESCs.

Genome-wide characterization of Msgn1 target genes

Microarrays were used to assess the effects of Msgn1 overexpression on global gene expression in EBs. Transcriptional profiles of iF-Msgn1 EBs at 12, 24 and 48 h after Dox treatment revealed that Msgn1 rapidly activated gene expression (NCBI GEO dataset GSE29848). After 24 h of Dox treatment, 85% (280/332, $P \leq 0.05$) of the unique, differentially expressed genes were upregulated >1.5 fold (supplementary material Fig. S2A and Table S1). Pathway analysis revealed that a highly significant number of the genes activated by Msgn1 in EBs were expressed in mesoderm (73/332, $P \leq 4.94 \times 10^{-39}$) (Fig. 2A, Table 1; supplementary material Table S1), somites, the PS and stem cells, and were associated with developmental signaling pathways including the Wnt/β-catenin pathway (Fig. 2B, Table 1). Remarkably, Msgn1 activated genes crucial for the development of multiple aspects of PSM development. Msgn1 activated key mesodermal progenitor markers, including *Tbx6*, *Pdgfra* and *Gata4* (Fig. 1I, Table 1), as well as genes that control EMT (*Snai1*, *Twist1*, *Zeb1*, *Zeb2* and *Foxc1*), segmentation and Notch signaling (*Epha4*, *Dll1*, *Dll3*, *Notch1*, *Lfng* and *Nrarp*), and planar cell polarity (*Wnt5a*, *Prickle1* and *Dact1*) (Table 1).

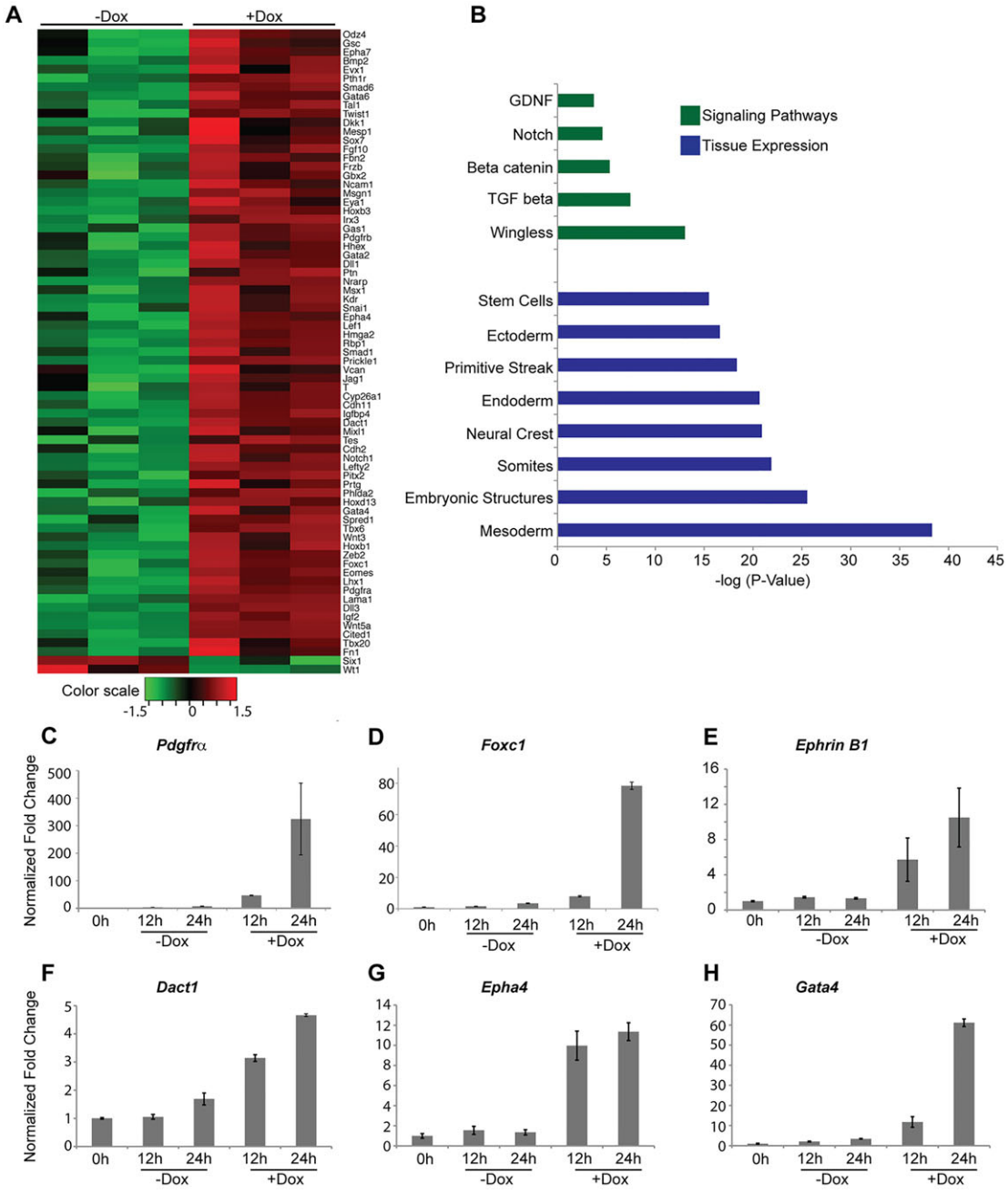

Fig. 2. Identification of the Msgn1 transcriptome in differentiating iF-Msgn1 EBs. (A) Hierarchical clustering of differentially expressed mesoderm-associated genes induced by F-Msgn1 (+Dox) for 24 h ($P \leq 0.05$; fold change ≥1.5). The color bar represents $\log_2$ intensity values. (B) Pathway and tissue expression analysis of all differentially expressed genes ($P \leq 0.05$, fold change≥1.5) induced by F-Msgn1 for 24 h. The *x*-axis shows the *P*-value ($-\log_{10}$). (C-H) qPCR analysis of select differentially expressed mesoderm genes in iF-Msgn1 EBs treated with (+) or without (–) Dox for 12 h and 24 h. The *y*-axis shows the normalized fold change. Results are mean±s.e.m.

Development (2014) 141, 4285-4297 doi:10.1242/dev.110908

Table 1. Pathway analysis of differentially expressed genes upon induction of Msgn1 for 24 h

Category	Differentially expressed genes
Epithelial to mesenchymal transition	*Twist1* (1.9), *Lef1* (1.8), *Zeb2* (2.5), *Zeb1* (1.6), *Snai1* (2.8), *Notch1* (2.0), *Wnt5a* (7.4), *Mmp2* (1.6), *Pdgfra* (6.2), *Foxc1* (1.8), *Fn1* (1.5), *Cdh2* (1.5), *Pdgfrb* (1.5)
Mesoderm expression	*Igfbp4* (3.2), *Gata4* (2.1), *Lefty2* (1.5), *Prickle1* (2.7), *Eomes* (2.9), *Smad1* (1.8), *Dact1* (2.5), *Dkk1* (1.6), *Eya1* (1.5), *Sox7* (2.2), *Wnt3* (1.8), *Kdr* (2.7), *Cdh11* (3.9), *Cyp26a1* (3.5), *Phlda2* (1.52), *Snai1* (2.8), *Hhex* (1.5), *Jag1* (1.6), *Fbn2* (2.4), *Tbx6* (1.9), *Nrarp* (2.3), *Wnt5a* (7.4), *Gbx2* (1.6), *Gas1* (1.7), *Foxc1* (1.8), *Smad6* (2.0), *Prtg* (1.8), *Evx1* (2.0), *Zeb2* (2.5), *Hoxb3* (2.1), *Lama1* (4.0), *Pth1r* (2.0), *Frzb* (2.0), *Vcan* (1.8), *Hoxd13* (1.7), *Gata2* (1.5), *Odz4* (1.7), *Dll1* (2.0), *Lef1* (1.8), *Bmp2* (1.9), *Hmga2* (2.0), *T* (1.6), *Cdh2* (1.5), *Pitx2* (1.5), *Ncam1*, (1.8), *Wt1* (−1.8), *Epha4* (2.9), *Tbx20* (1.7), *Fn1* (1.5), *Rbp1* (1.8), *Igf2* (3.3), *Six1* (−1.6), *Lhx1* (4.2), *Mixl1* (2.8), *Notch1* (2.0), *Mesp1* (1.5), *Msx1* (1.6), *Msgn1* (1.8), *Tes* (1.5), *Tal1* (2.0), *Pdgfra* (6.2), *Ptn* (1.6), *Epha7* (1.7), *Irx3* (1.7), *Hoxb1* (2.4), *Gsc* (1.9), *Twist1* (1.9), *Gata6* (2.1), *Spred1* (1.7), *Pdgfrb* (1.5), *Dll3* (6.3), *Fgf10* (2.4), *Cited1* (5.6)
Skeletal system development	*Csrnp1* (−1.5), *Smad1* (1.8), *Mmp2* (1.6), *Eya1* (1.5), *Pkdcc* (1.7), *Hoxb2* (1.7), *Fbn2* (2.4), *Lgals3* (−1.7), *Wnt5a* (7.4), *Gas1* (1.7), *Foxc1* (1.8), *Mgp* (−1.5), *Hoxb3* (2.1), *Pth1r* (2.0), *Frzb* (2.0), *Hoxd13* (1.7), *Smad3* (1.6), *Bmp2* (1.9), *Col1a2* (1.6), *Six1* (−1.5), *Msx1* (1.6), *Pdgfra* (6.2), *Hoxb1* (2.4), *Gsc* (1.9), *Twist1* (1.9), *Pdgfrb* (1.5), *Zeb1* (1.6), *Dll3* (6.3)
Wnt signaling pathway	*Igfbp4* (3.2), *Sp5* (2.5), *Prickle1* (2.7), *Lgr5* (1.7), *Dact1* (2.5), *Dkk1* (1.6), *Sox7* (2.2), *Wnt3* (1.8), *Cyp26a1* (3.5), *Snai1* (2.8), *Hhex* (1.5), *Tbx6* (1.9), *Nrarp* (2.3), *Lgals3* (−1.7), *Wnt5a* (7.4), *Gbx2* (1.6), *Prl2c2* (−2.0), *Foxc1* (1.8), *Peg3* (1.7), *Frzb* (2.0), *Olfm1* (1.9), *Lef1* (1.8), *Nkd1* (1.7), *Bmp2* (1.9), *T* (1.6), *Cdh2* (1.5), *Pitx2* (1.5), *Hprt* (5.3), *Wt1* (−1.8), *Nr0b1* (−1.9), *Six1* (−1.6), *Lhx1* (4.2), *Mixl1* (2.8), *Mesp1* (1.5), *Msx1* (1.6), *Irx3* (1.7), *Dixdc1* (1.5), *Nkd2* (1.6), *Pdzrn3* (1.7), *Rspo3* (2.6), *Gata6* (2.1), *Sp8* (1.9), *AW551984* (1.7), *Fgf10* (2.4)
TGFβ signaling pathway	*Igfbp4* (3.2), *Gata4* (2.1), *Serpinh1* (1.6), *Lefty2* (1.5), *Lmo1* (1.9), *Eomes* (2.9), *Smad1* (1.8), *Bmper* (2.0), *Snai1* (2.8), *Hhex* (1.5), *H2-Ab1* (1.6), *Fbn2* (2.4), *Foxc1* (1.8), *Smad6* (2.0), *Id1* (1.8), *Mgp* (−1.5), *Fli1* (4.8), *Zeb2* (2.5), *Fstl1* (2.1), *Vcan* (1.8), *Smad3* (1.6), *Ets1* (1.8), *Bmp2* (1.9), *Hmga2* (2.0), *T* (1.6), *Tbx20* (1.7), *Fn1* (1.5), *Col1a2* (1.6), *Six1* (−1.6), *Mixl1* (2.8), *Msx1* (1.6), *Twist1* (1.9), *Gata6* (2.1), *Zeb1* (1.6), *Zfp521* (1.7), *Fgf10* (2.4), *Cited1* (5.6)
Notch signaling pathway	*Dll1* (2.0), *Dll3* (6.3), *Notch1* (2.0), *Nrarp* (2.3), *Epha4* (2.9), *Lfng*, *Maml3* (Chalamalasetty et al., 2011)
Planar cell polarity pathway	*Wnt5a* (7.4), *Prickle* (2.7), *Dact1* (2.5)

Differentially expressed genes with $P \leq 0.05$ and a fold change≥1.5 are listed. Values in parentheses indicate fold changes. Where there is no value, there is no change.

Msgn1-dependent gene activation was validated by qPCR analysis for select genes (Fig. 2C-H; Chalamalasetty et al., 2011). These data show that Msgn1 regulates the expression of key regulators of PSM fate.

To determine which of these genes represent direct targets of Msgn1, genomic Msgn1-binding sites in differentiating EBs were determined by chromatin immunoprecipitation sequencing (ChIP-Seq). A total of 4087 identified Msgn1-binding sites or 'peaks' were assigned to 1860 genes because they were found within or adjacent [−10 kb to Transcription Start Site (TSS); +5 kb to Transcription Termination Site (TTS)] to a gene. A significant number of peaks were within 2 kb upstream of the TSS, suggesting that Msgn1 predominantly bound to promoters (Fig. 3A,B; supplementary material Table S2). Integration of the ChIP-Seq and microarray data revealed that one-third (121/332 genes) of the genes that were differentially expressed upon Msgn1 induction were directly bound by Msgn1 (Fig. 3C; supplementary material Table S3). Gene ontology (GO) term analysis using the GREAT algorithm (McLean et al., 2010) showed enrichment for genes that are expressed in the PS and mesoderm, and those that regulate mesoderm differentiation and morphogenesis, germ layer formation and EMT, among others (Fig. 3D; supplementary material Fig. S2B and Table S2). Analysis of all 4087 Msgn1-binding site sequences for common sequence motifs identified an E-box motif (CANNTG), specifically CCATTTGT (Fig. 3E; supplementary material Fig. S2C), which is a well-characterized binding site for bHLH transcription factors (Jones, 2004). This Msgn1-binding site was located centrally within identified peaks (supplementary material Fig. S2D) and bound to Msgn1 in electrophoretic mobility shift assays (EMSAs) (supplementary material Fig. S3; Chalamalasetty et al., 2011). Several other interesting motifs were associated with Msgn1-binding sites, including binding sites for the Ets, T-box, Gata and Sp transcription factors (supplementary material Fig. S2E).

Msgn1 directly activates key regulators of PSM fate

The rapid expression of *Tbx6* and *Pdgfra* mRNA and protein in EBs upon Msgn1 induction (Fig. 1I; supplementary material Fig. S1D) is consistent with Msgn1 directly activating these genes. Indeed, ChIP-Seq revealed that Msgn1 bound to regulatory elements in both *Tbx6* (Fig. 4A) and *Pdgfra* (Fig. 4D). The *Tbx6+3913* 3′ enhancer contained four E-boxes (Fig. 4A), including two embedded consensus Msgn1-binding sites (E1 and E3) that bound (supplementary material Fig. S3A), and were activated by, Msgn1 *in vitro* (Fig. 4B). The *Tbx6* ChIP-seq peak was validated by ChIP-qPCR in iF-Msgn1 EBs (Fig. 4C). Similarly, six E-boxes (E1-6) were discerned in the *Pdgfra −2140* peak (Fig. 4D), including two consensus Msgn1-binding sites (E2 and E4) that bound Msgn1 (supplementary material Fig. S3B). Mutation of E2 showed that it was necessary for Msgn1-dependent activation of a *Pdgfra −2140* enhancer-luciferase reporter (Fig. 4E). Importantly, both *Tbx6* and *Pdgfra* were downregulated in the *Msgn1*$^{-/-}$ PSM (Fig. 1A-D, Fig. 4F-I), clearly demonstrating their dependence upon *Msgn1 in vivo*. These data show that Msgn1 positively regulates key mediators of the PSM cell fate.

Msgn1 is a direct regulator of EMT

Given that *Snai1* and other EMT genes are induced by Msgn1 (Table 1), and that *Snai1* is considered a master regulator of the EMT program (Batlle et al., 2000; Barrallo-Gimeno and Nieto, 2005), we examined the mechanism of *Snai1* activation by Msgn1. RT-PCR analysis validated the microarray data, showing that *Snai1* was rapidly and robustly activated in EBs within 6 h of Msgn1 induction (Fig. 5A). ChIP-seq analysis revealed that Msgn1 bound a previously described *Snai1* 3′ enhancer (Palmer et al., 2007) that contains two E-boxes (Fig. 5B). Co-transfection of Msgn1 expression vectors with *Snai1*+5148 luciferase reporter constructs showed that Msgn1 activated the reporter sevenfold (Fig. 5C). Moreover, binding of Msgn1 to the *Snai1* enhancer was confirmed by ChIP assays in both iF-Msgn1 EBs and PSM extracts using anti-Msgn1 antibody (Fig. 5D).

To determine whether *Snai1* expression *in vivo* depends upon Msgn1 and the Wnt3a/β-catenin signaling pathway, we examined *Snai1* expression in *Wnt3a*$^{-/-}$, conditional *Ctnnb1* loss-of-function

Development (2014) 141, 4285-4297 doi:10.1242/dev.110908

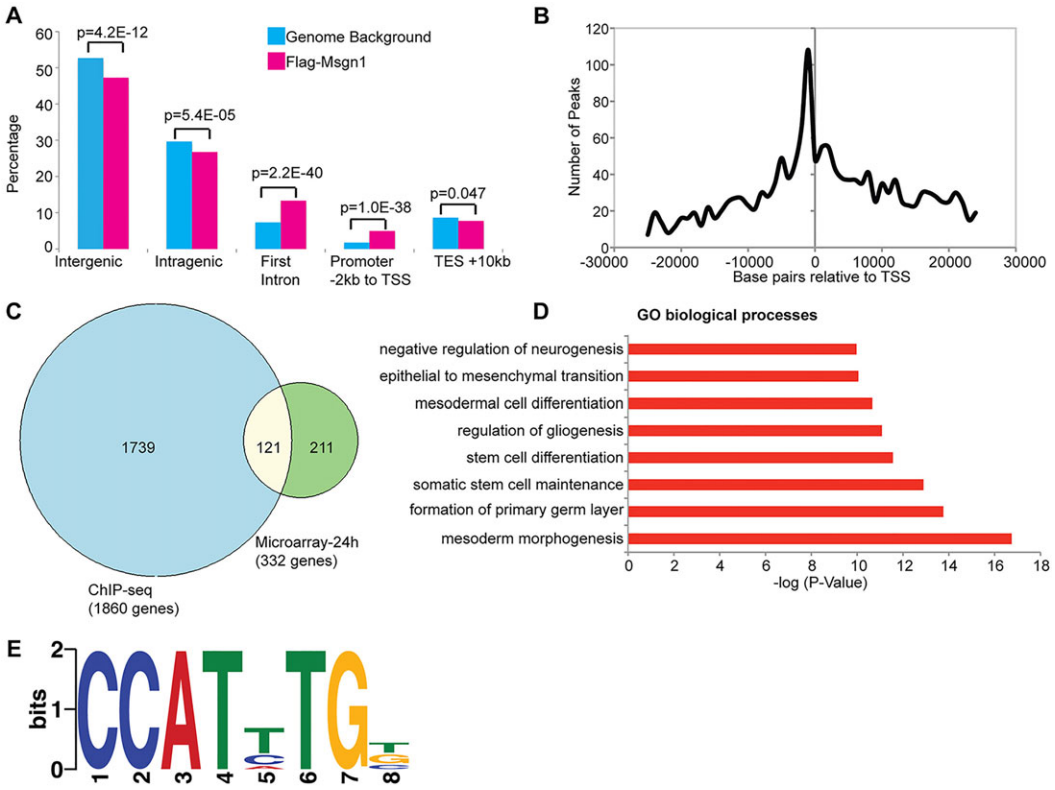

Fig. 3. ChIP-seq analysis of iF-Msgn1 EBs. (A) Genomic distribution of 4087 Msgn1 peaks (pink) compared with those from genome background (blue). For genome background, percentages of peaks from random sampling of similar length reads across the genome background are used. A binomial exact test was used to calculate the *P*-values. TSS, transcription start site; TES, transcription end site. (B) Msgn1 peaks within 25 kb upstream or downstream of the TSS are plotted. (C) Venn diagram depicting the intersection of 1860 genes bound by Msgn1 (blue circle) with 332 differentially expressed genes at the 24-h time point (green circle). The overlap is statistically significant ($P \leq 4.6\times10^{-80}$, hypergeometric probability test). (D) GO biological processes associated with genes bound by F-Msgn1 and analyzed by GREAT. The *x*-axis shows the *P*-value ($-\log_{10}$). (E) Discriminative DNA motif discovery (DREME) analysis of Msgn1 peaks identifies an E-box (underlined), CCATHTGB, as the top motif (E-value=7.6×10^{-279}). The motif logo displays nucleotide conservation at that position (measured in bits) and height of the symbol reflect the relative frequency at that position.

and gain-of-function mutants, and *Msgn1*$^{-/-}$ embryos. Whole-mount *in situ* hybridization (WISH) confirmed that *Snai1*, like *Msgn1* (Chalamalasetty et al., 2011), depended upon Wnt3a and β-catenin signaling for expression in the PSM (supplementary material Fig. S3C-F). Consistent with *Snai1* being a direct target of Msgn1, examination of *Msgn1* mutants showed that *Snai1* expression was absent from the *Msgn1*$^{-/-}$ PSM, but, as expected, was maintained in the allantois and head mesenchyme where Msgn1 is not normally expressed (Fig. 5E-H; supplementary material Fig. S3G).

Additionally, the ChIP-seq data showed that six additional genes implicated in the regulation of EMT [*Zeb2*, *Foxc1*, *Lef1*, *Notch1*, fibronectin (*Fn1*) and *Pdgfra*, see above] were directly bound by Msgn1. Quantitative real-time PCR (qPCR) analysis confirmed that Msgn1 expression in EBs resulted in the upregulation of several mesenchymal markers including N-cadherin (*Cdh2*), *Fn1*, *Foxc1*, *Mmp2*, *Twist1* and vimentin (*Vim*), whereas epithelial and junctional markers, such as E-cadherin (*Cdh1*), claudin 3 and claudin 7 (*Cldn3*, *Cldn7*), and occludin (*Ocln*) were downregulated (Fig. 6A; supplementary material Fig. S3H). As many of the targets of Msgn1 are also Snai1 targets, the results show that Msgn1 activates a redundant mechanism to initiate the EMT program (Cano et al., 2000; Guaita et al., 2002; Barrallo-Gimeno and Nieto, 2005).

The lack of *Snai1* expression in the *Msgn1*$^{-/-}$ PSM prompted us to examine embryo sections for defects in EMT, which could manifest as abnormal cell morphology, ingression or movement in the PS. Surprisingly, cells with mesenchymal morphology had clearly ingressed in the mutant PS (Fig. 6B,C). The morphology of these *Msgn1*$^{-/-}$ cells in the midline PS was indistinguishable from similarly located wild-type mesenchymal cells that had ingressed but had not yet moved from the PS, suggesting that the mutant cells display a cell motility defect. To directly test this, we assessed cell motility in PS explant cultures (Fig. 6D,E). In 18 of 21 control explants, motile cells migrated out of the explant, traveling distances of up to 1100 µm within 48 h of culture (Fig. 6F). These motile cells expressed the mesenchymal marker vimentin, and not the epithelial marker E-cadherin (Fig. 6G-J). By contrast, motile, vimentin-positive cells were much less abundant in the periphery of *Msgn1*$^{-/-}$ explants (n=14; Fig. 6K-N) and displayed significantly reduced migration, not traveling for more than ~300 µm (Fig. 6F). These results suggest that Msgn1 is required for the initiation of a motile mesenchymal state in PSM cells.

Msgn1 specifies PSM fates *in vivo*

To test whether Msgn1 can specify PSM fate *in vivo*, we generated a 'Tet-ON' *Msgn1*-expressing transgenic mouse line to be used in conjunction with Cre recombinase-expressing transgenic lines to precisely control the spatiotemporal expression of Msgn1 (supplementary material Fig. S4A). *T-Cre* mice were used to express Msgn1 in the trunk region, presumably in the NM stem cell that gives rise to both the neural tube and mesodermal lineages (Perantoni et al., 2005). Administration of Dox to *T-Cre*$^{tg/+}$*; R26-flox-*

Development (2014) 141, 4285-4297 doi:10.1242/dev.110908

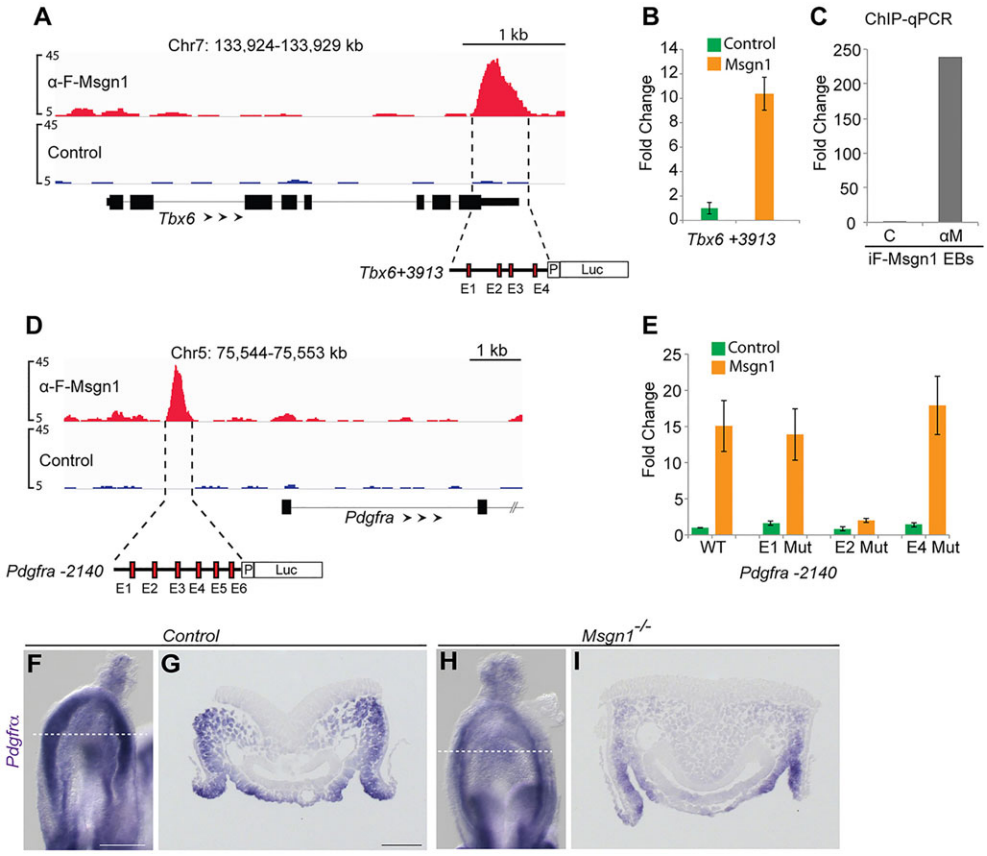

Fig. 4. The paraxial mesoderm regulators *Tbx6* and *Pdgfra* are direct target genes of Msgn1. (A) Msgn1-binding peaks (red) and input control (blue) at the *Tbx6* locus for the indicated genomic intervals (mm9 coordinates). The schematic (below) illustrates the four E-boxes (E1-4, red rectangles) identified in the *Tbx6*+3913 enhancer peak. Arrowheads reflect gene orientation and the *y*-axis indicates the peak height corresponding to the normalized fold enrichment. (B) Luciferase reporter assay of *Tbx6*+3913 enhancer co-transfected with pCS2 Control (green) or pCS2-Msgn1 (orange) expression constructs. The *y*-axis shows the normalized fold change; error bars show the s.d.; *n*=3. (C) ChIP-qPCR analysis of F-Msgn1 binding to *Tbx6+3913* enhancer using control 'C' and anti-Msgn1 'αM' ascites in iF-Msgn1 EBs. (D) Msgn1 binds the *Pdgfra* locus at −2140 bp, relative to TSS. (E) Luciferase reporter assay of the *Pdgfra* −2140 wild-type (WT) and E-box mutant enhancers co-transfected with control or Msgn1 expression constructs. The *y*-axis shows the normalized fold change; error bars show the s.d.; *n*=3. (F-I) WISH analysis of *Pdgfra* mRNA expression in control and *Msgn1*$^{-/-}$ embryos. G and I are cross sections through the PS [dashed white lines in F and H (ventral view)]. Scale bars: 200 µm (F,H), 50 µm (G,I).

rtTA-Ires-EGFP/$^{+}$*; lacZ-TRE-F-Msgn1*$^{tg/+}$ (hereafter referred to as *Tcre-F-Msgn1*^{GOF}, where GOF stands for gain-of-function) transgenic embryos led to the highly elevated expression of Msgn1 and the Msgn1 reporter β-galactosidase (β-gal) (Fig. 7A-F; supplementary material Fig. S4B-D) throughout the trunk neural tube and paraxial mesoderm.

Using EGFP (rtTA-Ires-EGFP) expression as a permanent cell lineage tracer, the ectopic expression of Msgn1 in NM stem cells and their progeny resulted in a reduction in the number of EGFP-expressing cells contributing to the neural tube. Although 62% (*n*=5) of neural cells were EGFP^{+} in controls, only 22% (*n*=5) were EGFP^{+} when Msgn1 was induced (Fig. 7A-G), suggesting that Msgn1 expression diverted NM progeny away from a neural fate. Nevertheless, F-Msgn1 was still detected in the cells that remained in the neural tube. These cells were Sox2^{+} and Tbx6^{-}, suggesting that they retained neural character (supplementary material Fig. S4E-L). To investigate whether the reduced number of EGFP^{+} neural cells was caused by Msgn1-induced cell death, we examined *Tcre-F-Msgn1*^{GOF} embryos for expression of the apoptosis initiator protein caspase-3. Although caspase-3 was detected in 2/5 and 3/3 mutants at E9.5 and E10.5, respectively, it was mostly observed in mesoderm cells and did not colocalize with F-Msgn1 (supplementary material Fig. S5), suggesting that the cell death is a secondary consequence of Msgn1 expression.

The reduced number of NM stem cell progeny contributing to the neural tube was accompanied by a correspondingly dramatic increase in the number of *Tcre-F-Msgn1*^{GOF} PSM cells (Fig. 7H-N), and by the enlarged mRNA and protein expression domains of the PSM markers *Dll3* and *Tbx6* (Fig. 7O-T; supplementary material Fig. S4I-L). The expansion of the PSM came at the expense of somites because segment boundaries were absent and the expression of the somite polarity marker *Uncx4.1* (*Uncx* – Mouse Genome Informatics) was reduced or absent (Fig. 7O-T). *T* expression in the germ layer progenitors of the PS was also expanded (Fig. 7U,X). Thus, forced expression of Msgn1 in NM stem cells resulted in fewer progeny becoming neural, and more cells becoming PSM, which is consistent with Msgn1 specifying the PSM fate.

We took advantage of *T-Cre* expression in the posterior notochord (Perantoni et al., 2005) to test whether ectopic expression of Msgn1 could induce notochord progenitors to take up a PSM fate. Msgn1 expression in notochord progenitors severely disrupted the expression of the notochord marker *T* (Fig. 7U-Z) in the posterior midline of *Tcre-F-Msgn1*^{GOF} embryos treated with Dox (Fig. 7X-Z). Sections revealed that the *T*$^{+}$ midline notochord, which normally divides the adjacent Tbx6^{+} PSM into two paraxial domains (Fig. 7H-J,V,W), was absent in *Tcre-F-Msgn1*^{GOF} embryos (Fig. 7Y,Z) and was replaced by a large swath of Msgn1^{+} Tbx6^{+} PSM (Fig. 7K-M). Sections further revealed that the *Tcre-F-Msgn1*^{GOF} neural tube was smaller and displayed abnormal ventral morphology (cf. Fig. 7V,Y), consistent with the absence of a floor plate due to a lack of notochord-derived ventralizing signals (Dodd et al., 1998). These results suggest that the ectopic expression of Msgn1 suppressed notochord fate and promoted a PSM fate.

Msgn1 functions downstream of Wnt in mesoderm formation

Wnt3a initiates PSM formation in the trunk region of the embryo. As Msgn1 lies downstream of Wnt3a, we tested whether Msgn1 was sufficient to induce PSM differentiation independent of Wnt activity. iF-Msgn1 and control iCre EBs were treated with the secreted Wnt antagonist Dkk1 to inhibit endogenous Wnt signaling (Fig. 8A). Dkk1 suppressed the number of Pdgfrα^{+} cells in uninduced cells (12.5% in samples without Dox compared with 5% in samples without Dox and with Dkk1) but had no effect on

Development (2014) 141, 4285-4297 doi:10.1242/dev.110908

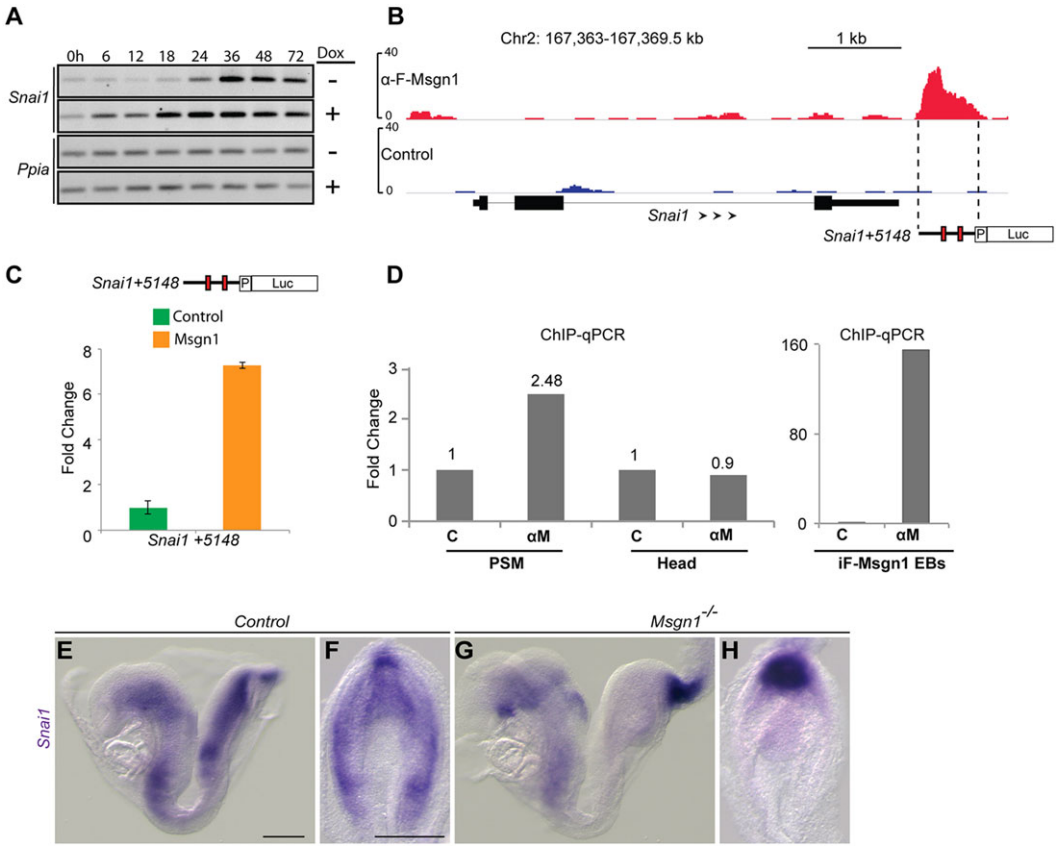

Fig. 5. The EMT master regulator *Snai1* is a direct target of Msgn1. (A) Semi-quantitative PCR analysis of *Snai1* expression in iF-Msgn1 EBs treated with (+) or without (–) Dox from 0-72 h. Cyclophilin A (*PpiA*) is used as a loading control. (B) Msgn1 binds to the +5148 bp enhancer of *Snai1*. The schematic (below) shows the two E-boxes (red rectangles) identified in the *Snai1* +5148 enhancer peak. Arrowheads indicate orientation and *y*-axis denotes peak height, corresponding to the normalized fold enrichment. (C) Luciferase reporter assay of *Snai1* +5148 enhancer co-transfected with control or Msgn1 expression constructs. The *y*-axis shows the normalized fold change; error bars show the s.d.; *n*=3. (D) ChIP-qPCR analysis of Msgn1 enrichment on *Snai1* +5148 locus using control 'C' and anti-Msgn1 'αM' ascites in E9.5 PSM and head extracts and in iF-Msgn1 EBs. (E-H) WISH analysis of *Snai1* mRNA in E8.5 control and *Msgn1*$^{-/-}$ embryos (E,G, lateral view; F,H, ventral view). Scale bars: 200 µm.

Msgn1-induced samples (cf. 57% in samples treated only with Dox and 55.8% in samples treated with Dox and Dkk1). In controls, Dkk1 treatment suppressed Pdgfrα irrespective of Dox treatment. Thus, Msgn1 is capable of promoting PSM differentiation in ESCs even in the absence of Wnt activity.

To examine whether Msgn1 was sufficient to restore PSM differentiation *in vivo* and rescue the *Wnt3a*$^{-/-}$ mutant phenotype, we expressed F-Msgn1 in *Wnt3a*$^{-/-}$ NM stem cells using T-Cre (Fig. 8B-K). *Wnt3a*$^{-/-}$ mutants lack posterior paraxial mesoderm (Fig. 8B,C,E,F), do not express the somite and PSM marker *Meox1* in the posterior embryo (Fig. 8H,J) and instead display a duplicate neural tube (Fig. 8F). Forced expression of *Msgn1* restored paraxial mesoderm to *Wnt3a*$^{-/-}$ embryos (cf. Fig. 8F,G) and posteriorly extended the *Meox1* expression domain (cf. Fig. 8J,K), suggesting that Msgn1 is sufficient to drive PSM differentiation in the absence of Wnt3a. Msgn1 expression did not fully rescue posterior axis extension, presumably owing to a requirement for Wnt3a to maintain trunk progenitors. Interestingly, the duplicate neural tube was still observed despite the forced expression of Msgn1 (Fig. 8G), suggesting that the expression of Msgn1 alone is insufficient to suppress neural specification.

DISCUSSION

We have shown that Msgn1 is a master regulatory switch for paraxial PSM fate. Msgn1 possesses the extraordinary capacity to convert progenitor cell populations into paraxial mesoderm and is essential for paraxial mesoderm formation throughout the majority of the vertebrate trunk axis. Of the 65 total pairs of somites that form in the mouse embryo, only 8-10 develop in *Msgn1*$^{-/-}$ mutants, demonstrating that *Msgn1* is essential for proper trunk PSM and somite development. It is likely that the anterior somites form under the control of a separate genetic program, as suggested by the many mouse mutants that display defects in PSM formation and axis truncation but do not have defects in the anterior-most 8-10 somites (Takada et al., 1994; Chapman and Papaioannou, 1998; Galceran et al., 1999; Dunty et al., 2008, 2014; Nowotschin et al., 2012). By contrast, *Pdgfra*$^{-/-}$ mutants show defects in the most rostral somites (Soriano, 1997). Our demonstration that *Pdgfra* is a direct target of Msgn1 predicts that the rostral somites should also be affected in *Msgn1*$^{-/-}$ embryos. The absence of a *Pdgfra*$^{-/-}$-like phenotype in these mutants suggests that the rostral somites that form in *Msgn1* mutants could arise owing to redundancies between Msgn1 and the closely related Mesp1 and Mesp2 bHLH factors (Yoon et al., 2000).

Wnt3a regulates EMT and PSM cell motility through Msgn1

The PSM is a transitory musculoskeletal population formed from a trunk stem cell population. PSM cells are uniquely identified by their early mesodermal progenitor state and their motile mesenchymal phenotype. Our extensive gene-regulatory analysis of pathways downstream of Msgn1 provides a unique look at how a

Development (2014) 141, 4285-4297 doi:10.1242/dev.110908

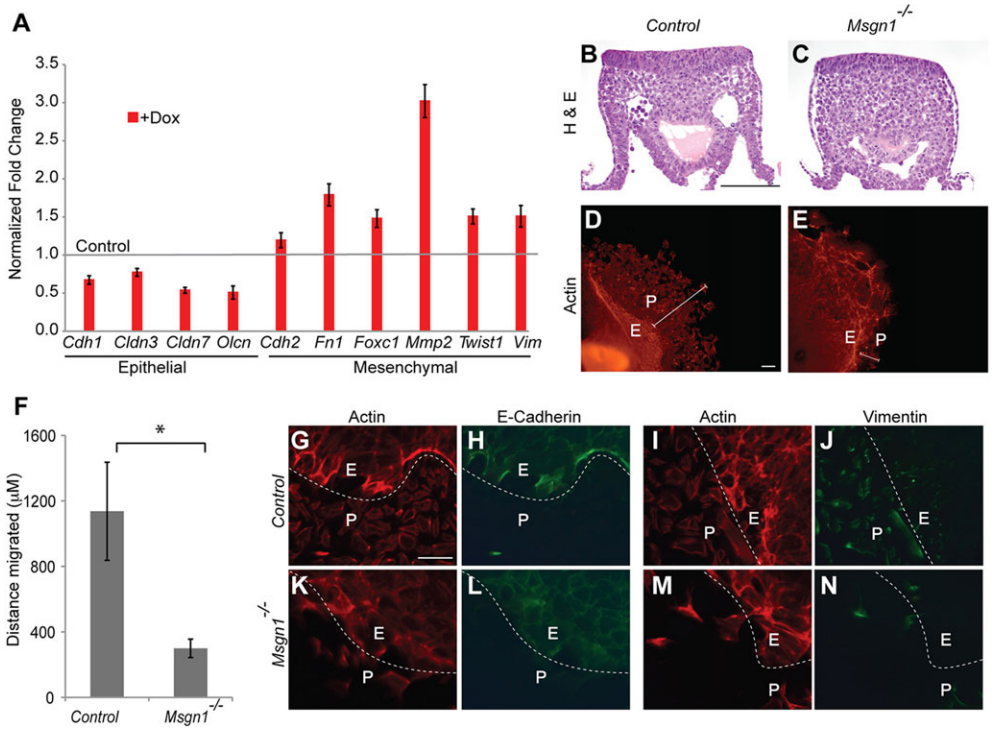

Fig. 6. Msgn1 is required for EMT and motility. (A) qPCR analysis of epithelial and mesenchymal genes expressed in Dox-treated (+Dox) iF-Msgn1 EBs after 48 h. The *y*-axis shows the normalized fold change; error bars are the s.d. (B,C) Histological analysis (H&E) of PS cross-sections of control (n=7) and $Msgn1^{-/-}$ (n=9) embryos at E8.5. (D,E) Rhodamine-phalloidin staining of control (D, n=21) and $Msgn1^{-/-}$ (E, n=14) tail bud explants dissected at the 18-22 somite stage and cultured for 48 h. The distance migrated from explant (E) to periphery (P) is indicated by white line. (F) The average distance migrated in control (n=13) and $Msgn1^{-/-}$ (n=7) explants is shown on *y*-axis. $^{*}P \leq 9.88\times10^{-5}$, two-sample unequal variance Student's *t*-test. (G-N) Immunostaining of control and $Msgn1^{-/-}$ explants cultured for 48 h and stained with Rhodamine-phalloidin, and anti-E-cadherin and anti-vimentin antibodies. The dashed white line demarcates the explant (E) from periphery (P) region. Scale bars: 100 μm.

master regulator functions to activate pathways essential for the many features that define cell fate. Our data show that Msgn1 can act as a multifaceted master regulator through activation of other master regulators. Previous analyses of $Wnt3a^{-/-}$ embryos have shown that migrating mesenchymal PSM cells are missing in the mutant PS (Yoshikawa et al., 1997). *Snai1* is downregulated in *Wnt3a* and *Ctnnb1* loss-of-function mutants (this work; Dale et al., 2006) and is induced by recombinant Wnt3a treatment of ESCs (R.B.C and T.P.Y., unpublished). Our data show that *Msgn1* – a Wnt3a target gene – is a major developmental regulator of *Snai1* and other EMT genes. Given that *Snai1* is a direct target of Msgn1, the loss of *Snai1* expression in *Wnt3a* and *Ctnnb1* mutants is presumably secondary to the loss of *Msgn1* expression in these mutants. Thus, Msgn1 is an excellent candidate to transduce the Wnt3a/β-catenin signal that regulates EMT and cell movement during gastrulation and axis extension. Interestingly, Msgn1 directly activates many of the EMT pathway genes that Snai1 activates, suggesting that Msgn1 ensures the activation of the EMT and motility programs in PSM cells through a feed-forward mechanism involving *Snai1*. Notably, $Msgn1^{-/-}$ mesodermal progenitors still ingress and display mesenchymal morphology in the PS, but are unable to move out of the PS. This argues that the Msgn1-Snai1 axis primarily functions to drive cell movement rather than EMT. The demonstration that Msgn1 regulates cell movements in zebrafish is consistent with this conclusion; however, the mechanisms appear to be substantially different because Msgn1 suppresses *snail1a* in fish (Fior et al., 2012). Alternatively, it is possible that additional Wnt3a target genes compensate for the loss of *Msgn1* and initiate, but fail to complete, the EMT process.

An enormous interest in EMT, and the reverse mesenchymal-epithelial transition (MET), has developed in recent years owing to the important roles that these evolutionarily conserved developmental processes play in embryogenesis, stem cell reprogramming, tumor progression and metastasis (Esteban et al., 2012; Lim and Thiery, 2012; De Craene and Berx, 2013). The only known tissue and developmental stage that *Msgn1* is expressed is in the mesenchymal cells of the PSM, making Msgn1 a remarkably stage- and tissue-specific transcription factor. Our studies demonstrate that Msgn1 regulates the EMT and PSM differentiation programs by binding to a unique E-box sequence that is distinct from the binding sites for Twist1 (Eckert et al., 2011), MyoD and NeuroD1 (Fong et al., 2012). Other non-bHLH EMT transcription factors, such as the Zn-finger homeodomain-containing factors Zeb1 and Zeb2, also function by binding to E-boxes and closely related Z-boxes (Burk et al., 2008). Interestingly, Msgn1-binding sites can be found clustered with binding sites for these EMT transcription factors in transcriptional enhancers of EMT genes suggesting that differentiation factors and EMT transcription factors target the same enhancers. For example, Msgn1 activates *Pdgfra* expression through the same highly conserved enhancer that is targeted by Twist1 during mammary tumor metastasis (Eckert et al., 2011). In addition to binding sites for Msgn1 and Twist1, this enhancer also contains binding sites for MyoD1 and Zeb1. These results are consistent with the EMT genetic program lying embedded within broader differentiation programs, and suggest a mechanism for how transcription factors that specify unique cellular phenotypes can similarly modulate EMT. In that sense, the EMT program is highly receptive to regulation by bHLH transcription factors to promote mesenchymal cell behavior.

Msgn1 is a potent activator of PSM differentiation

Careful assessment of $Msgn1^{-/-}$ embryos has revealed that primitive mesodermal progenitors form but are unable to differentiate into PSM or move away from the PS. This explains the grossly enlarged tailbud observed in $Msgn1^{-/-}$ mutants as being the result of the combined inability of immature mesodermal progenitors to undergo both differentiation and migration. *Tbx6* mutants also display a similarly enlarged tailbud (Chapman and Papaioannou, 1998). These findings in mice contrast with zebrafish studies where an enlarged tailbud is only seen in the absence of both *msgn1* and *spadetail* suggesting perhaps that the basic regulatory networks that define the PSM phenotype has diverged to some extent between mammals and fish (Fior et al., 2012; Yabe and Takada, 2012). In mice, $Wnt3a^{-/-}$

Development (2014) 141, 4285-4297 doi:10.1242/dev.110908

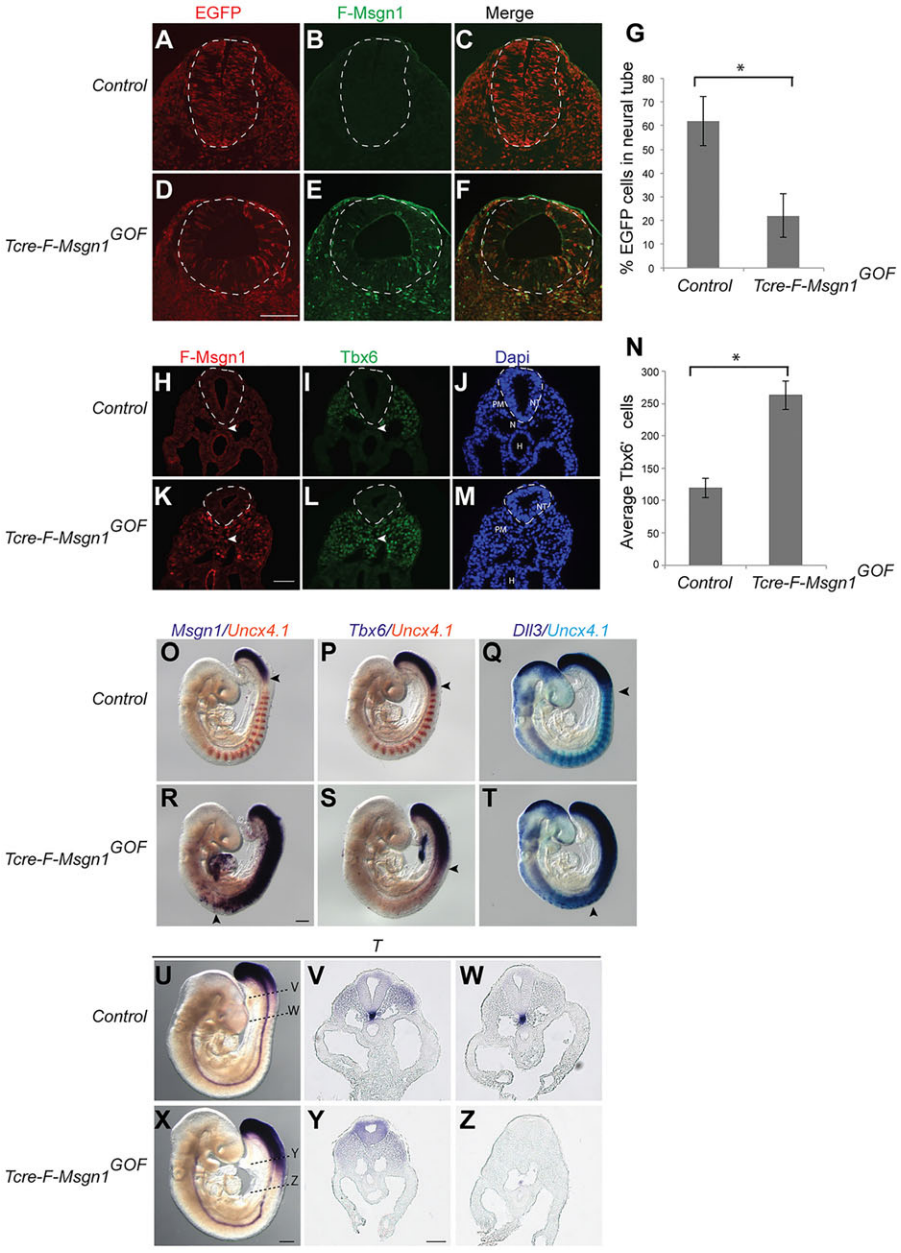

Fig. 7. Ectopic expression of F-Msgn1 in NM stem cells excludes cells from the neural tube and specifies PSM fates. (A-F) Immunodetection of EGFP (A,D) and F-Msgn1 (B,E) in cross-sections through the trunk neural tube (dashed lines) of E10.5 control (*T-Cre*$^{tg/+}$*; R26-flox-rtTA-Ires-EGFPl*$^{+}$) (A-C) and *T-cre-F-Msgn1*^{GOF} mutant embryos (*T-Cre*$^{tg/+}$*; R26-flox-rtTA-Ires-EGFPl*$^{+}$*; lacZ-TRE-F-Msgn1*$^{tg/+}$) (D-F). (G) Quantitative lineage tracing analysis of *T-Cre* labeled cells in neural tubes of control (A, *n*=5) and *T-cre-F-Msgn1*^{GOF} mutant (D, *n*=5) are shown. *$P \leq -2.18\times10^{-4}$, two-sample unequal variance Student's *t*-test. (H-M) Immunodetection of F-Msgn1 and Tbx6 protein in cross-sections through the PSM of E9.5 control (H-J) and *Tcre-F-Msgn1*^{GOF} mutant (K-M) embryos. NT, neural tube; PM, paraxial mesoderm; N, notochord; H, hindgut. Arrowheads denote the notochord. (N) Quantitative analysis of Tbx6^{+} cells in control (I, *n*=5) and *Tcre-F-Msgn1*^{GOF} mutants (L, *n*=5). *$P \leq -7.38\times10^{-6}$. (O-T) WISH analysis of *Msgn1* (O,R), *Tbx6* (P,S) and *Dll3* (Q,T), with the posterior somite marker *Uncx4.1* (orange in O,P,R and S; blue in Q and T) in E9.5 controls and *Tcre-F-Msgn1*^{GOF} mutants. (U-Z) WISH of *T* mRNA expression in control (U-W) and *Tcre-F-Msgn1*^{GOF} mutants (X-Z). V,W, cross-sections in U and Y; Z, cross-sections in X at the indicated embryonic locations. Scale bars: 100 µm (D,Y), 50 µm (K), 200 µm (R,X).

and $T^{-/-}$ embryos do not form a tailbud and are completely devoid of NM stem cells and posterior mesoderm by E9.5 (Wilkinson et al., 1990; Takada et al., 1994; Yamaguchi et al., 1999; Yoon and Wold, 2000). Thus, *Wnt3a* and *T* define a class of genes necessary for mesodermal progenitor formation and maintenance, whereas *Msgn1* and *Tbx6* belong to a separate class that is necessary for mesodermal differentiation (Fig. 9). Consistent with Msgn1 being the ultimate effector of PSM differentiation, *Msgn1* is not expressed in *Wnt3a*$^{-/-}$, $T^{-/-}$ or *Tbx6*$^{-/-}$ embryos (Wittler et al., 2007; Chalamalasetty et al., 2011; Nowotschin et al., 2012), whereas *Wnt3a*, *T* and, at least in mesodermal progenitors, *Tbx6*, continue to be expressed in *Msgn1*$^{-/-}$ mouse embryos (this work; Yoon and Wold, 2000; but see Nowotschin et al., 2012). The continued expression of *Wnt3a*, *T* and *Tbx6* in *Msgn1*$^{-/-}$ mutants demonstrates that their expression is insufficient for PSM differentiation in the absence of *Msgn1*.

Msgn1 promotes PSM differentiation through the direct activation of a PSM genetic program. Msgn1 activates *Tbx6*, *Snai1*, *Foxc1*, *Pdgfra* and *Dact1*, which are all essential for mesodermal development. Through the activation of multiple mediators of PSM cell fate, Msgn1 broadly instills PSM characteristics on mesodermal progenitors. A key feature of this activation of PSM fate is the parallel inhibition of competing cell fates. NM stem cells have the capacity to differentiate into both neural and mesodermal cells; however, in addition to activating the PSM program, Msgn1 also activates a program that inhibits entry into the neural cell fate. Specifically, Msgn1 directly activates the *Tbx6* transcription factor, which is a known inhibitor of the neural determinant *Sox2*, in PSM tissue. Indeed, misexpression of Sox2 in paraxial mesoderm is sufficient to generate ectopic neural tube formation (Takemoto et al., 2011). Tbx6 suppresses *Sox2* and neural fates through a neural-specific N1 enhancer in the *Sox2* locus (Takemoto et al., 2011). Thus, the reprogramming of neurons to a PSM fate likely requires the expression of at least two genes, *Msgn1* to specify PSM fates and *Tbx6* to suppress neural fates. Although we have shown that Msgn1 is

Development (2014) 141, 4285-4297 doi:10.1242/dev.110908

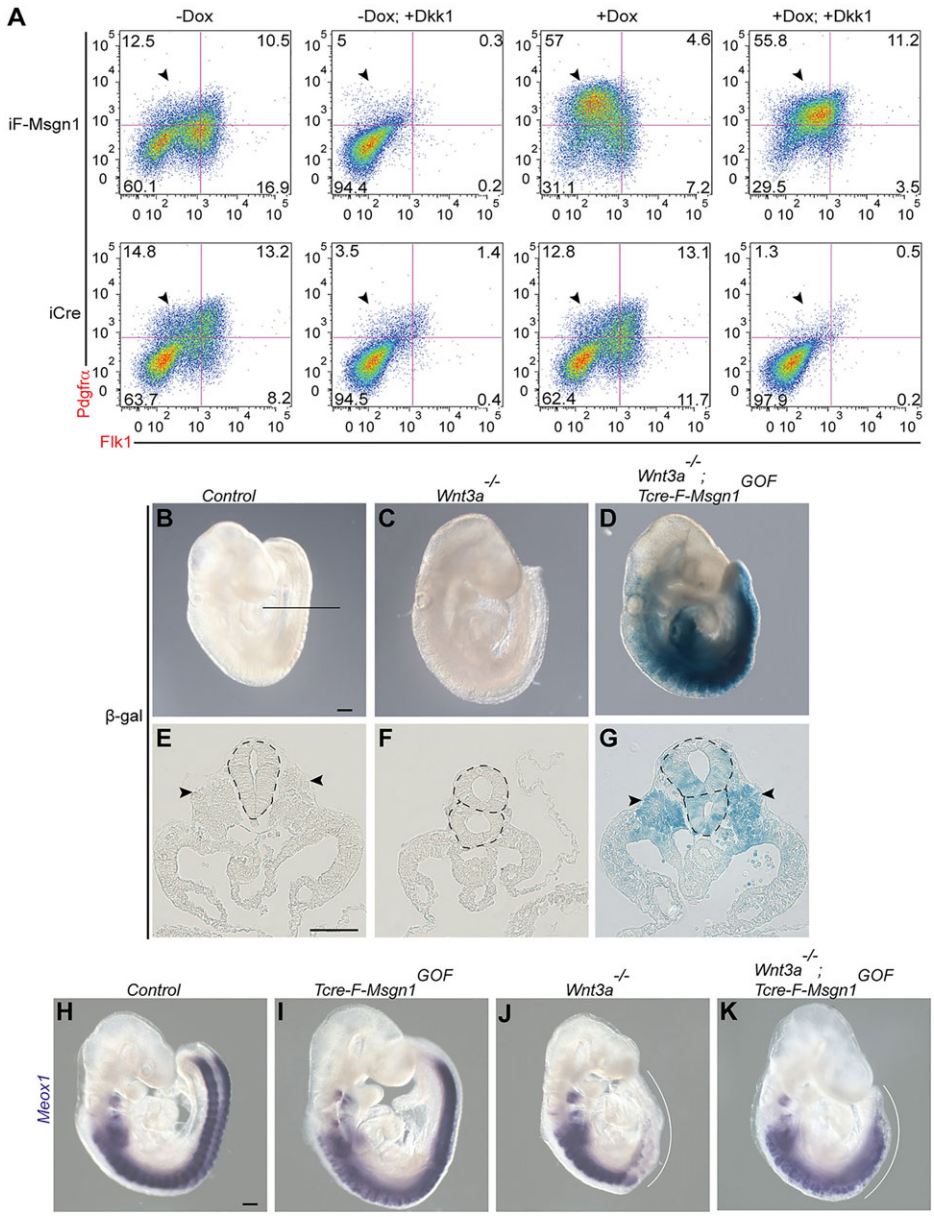

Fig. 8. Msgn1 rescues paraxial mesoderm in *Wnt3a*$^{-/-}$ mutants. (A) Flow cytometry analysis of iF-Msgn1 or control iCre EBs treated with (+) or without (–) Dkk1 and/or Dox for 48 h for the expression of Pdgfrα and Flk1. The arrowheads indicate Pdgfrα^{+} cells. (B-G) β-Gal staining of E9.5 control (B), *Wnt3a*$^{-/-}$ (C), and *Wnt3a*$^{-/-}$; *Tcre-F-Msgn1*^{GOF} (D) mutants. (E-G) These panels represent corresponding cross sections of embryos depicted in B-D, taken at the indicated location (black line, B). Arrowheads indicate paraxial mesoderm. (H-K) WISH analysis of *Meox1* in control (H), *Tcre-F-Msgn1*^{GOF} (I), *Wnt3a*$^{-/-}$ (J), and *Wnt3a*$^{-/-}$; *Tcre-F-Msgn1*^{GOF} (K) mutants. Curved white line in J, K highlights the rescued posterior paraxial mesoderm domain. Scale bars: 200 µm (B,H), 100 µm (E).

sufficient to activate *Tbx6* transcription in luciferase reporter assays in heterologous cell lines, it is apparently insufficient in neurons to reprogram neural fates or requires higher levels of expression than we were able to achieve experimentally. Nevertheless, given the powerful ability of Msgn1 to function as a transcriptional regulator, it will be interesting to test whether Msgn1 is capable of directly reprogramming human somatic cells into musculoskeletal stem cells (i.e. PSM), given that the PSM gives rise to several cell types, including skeletal muscle, dermis, bone and cartilage, that could be used in regenerative therapies for aging or musculoskeletal disease.

MATERIALS AND METHODS

Mice and embryos

The *Wnt3a*^{tm1Amc} (Takada et al., 1994), *Msgn1*^{tm1Wb} (Yoon and Wold, 2000), *Ctnnb1*^{tm2Kem} (Brault et al., 2001), *Ctnnb1*^{tm1Mmt} (Harada et al., 1999), Gt(Rosa) 26Sor^{tm1Sor} (Soriano, 1999), B6.Cg-*Gt (ROSA) 26Sortm1*$^{(rtTA,EGFP)Nagy}$/J (Belteki et al., 2005) and Tg(T-cre)1Lwd (Perantoni et al., 2005) mice were used to generate the allelic combinations featured in this study. For conditional expression of Flag-tagged Msgn1 in mice, *LacZ-TRE-F-Msgn1*$^{Tg/+}$ transgenic mice were generated by standard pronuclear injection methods. Detailed experimental details on the generation and characterization of *LacZ-TRE-F-Msgn1*$^{Tg/+}$ mice can be found in the methods in the supplementary materials. Embryos were dissected at E8.5 (4-8 somites) and E9.5 (18-22 somites). All animal experiments were performed in accordance with the guidelines established by the NCI-Frederick Animal Care and Use Committee.

Whole-mount *in situ* hybridization, β-gal staining and immunohistochemistry

Whole-mount *in situ* hybridization (WISH) experiments and probe synthesis (*Msgn1*, *Tbx6*, *Snai1*, *Meox1*, *T*, *Pdgfra*, *Uncx4.1* and *Dll3*) were carried out as described previously (Biris and Yamaguchi, 2014). All observed phenotypes were reproduced in at least three separate embryos unless otherwise specified. β-gal staining of cells and embryos was performed according to standard protocols (Whiting et al., 1991). For immunodetection on frozen sections, mouse embryos were fixed in 4% paraformaldehyde in PBS for 30 min at room temperature, incubated in 20% sucrose overnight at 4°C, embedded in OCT, and sectioned to a thickness of 5 µm. Immunodetection was performed as described previously (Dunty et al., 2014). Detailed experimental details can be found in the methods in the supplementary materials.

Development (2014) 141, 4285-4297 doi:10.1242/dev.110908

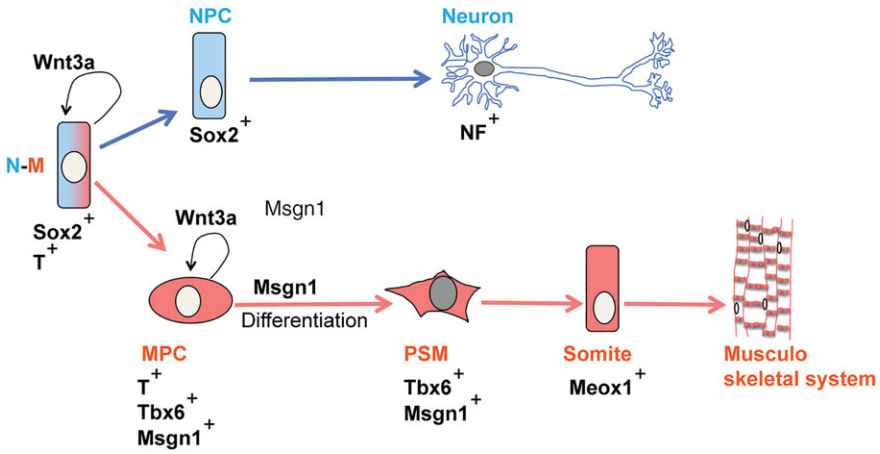

Fig. 9. Msgn1 is a master regulator of PSM differentiation. Bipotential NM stem cells, defined by Sox2 and T expression, generate the progenitors of the trunk spinal cord (NPC, neural progenitor cells) and musculoskeletal system (MPC, mesodermal progenitor cells). Wnt3a maintains both the NM stem cell and the MPC. In our model, the Wnt3a/β-catenin pathway activates the target genes *T* and *Msgn1* in NM stem cells in the PS to specify paraxial mesoderm fate (Yamaguchi et al., 1999; Dunty et al., 2008; Chalamalasetty et al., 2011). Msgn1 initiates the genetic programs that define PSM identity and cell behavior by activating downstream master regulators of EMT and motility (*Snai1*) and neural suppression (*Tbx6*) among others. *Msgn1* and *Tbx6* expression must be extinguished for PSM cells to undergo MET and segmentation to form somites and, ultimately, to differentiate into skeletal muscle, bone, cartilage, tendons, endothelial cells and dermis.

Tailbud explant migration assay

E9.5 18-22 ss embryos were dissected at room temperature in PBS. Excised tailbuds extended anteriorly from the tail tip up to, but not including, the hindgut. Isolated explants were transferred by pipette onto fibronectin-coated eight-well μ-slide dishes (Ibidi) containing 300 μl of Dulbecco's modified Eagle's medium (DMEM) and 10% fetal bovine serum (FBS) (Gibco), and were cultured for 48 h at 37°C in 5% CO_2. For immunocytochemistry analysis, explants were fixed in 4% PFA in PBS for 15 min at room temperature, washed with PBS and processed for antibody staining. Pilot experiments revealed that after the initial attachment and spreading of the explant on the dish, the tight cell-cell contacts visualized by intense phalloidin and E-cadherin staining remained intact in the ectoderm over the 48 h culture and readily defined the explant boundaries. Mesenchymal cells migrated radially away from the explant as individual cells and were defined by a low-level of phalloidin staining, and being negative for E-cadherin and positive for vimentin staining.

Cloning, ESC culture and differentiation

1×Flag-Msgn1, 3×Flag-Msgn1-Ires2-AcGFP and 3×Flag-Tbx6-Ires2-AcGFP were cloned into p2lox vector to generate iF-Msgn1 or iF-Tbx6 ESCs. Site-specific recombination in A2lox Cre ESCs was achieved as described previously (Iacovino et al., 2011). ESCs were cultured on mitomycin C-treated mouse embryo fibroblasts (MEFs) in DMEM, 15% FBS, 1× MEM non-essential amino acids, penicillin-streptomycin (all Life Technologies), 10^3 units/ml ESGRO (Chemicon) and 100 μM β-mercaptoethanol. For differentiation, ESCs were separated from MEFs and grown in suspension cultures for 2 days to form EBs. Dox was added on day 2 to induce Msgn1 and samples were harvested from 6-72 h post-induction, depending on the experiment. Detailed experimental details can be found in the methods in the supplementary materials.

EMSA assays

0.8 ng of DIG-labeled oligonucleotides (DIG Gel Shift Assay Kit, Roche, catalog number 03353591910) were incubated with Msgn1 and DNA-protein complexes were migrated on 6% DNA retardation gels (Invitrogen) and analyzed on Hybond-N+ membrane (Amersham, catalog number RPN303B) according to manufacturer's instructions. Detailed experimental details can be found in the methods in the supplementary materials.

Luciferase assays

200 ng of luciferase reporter and 5 ng of *Renilla* constructs were co-transfected with 50 ng of expression construct per well in 24-well plates. Luciferase assays were performed according to Dual Luciferase Assay Kit (Promega) instructions. Detailed experimental details can be found in the methods in the supplementary materials.

Reverse transcription and qPCR

RNA isolation, reverse transcription and qPCR were set up as described previously (Chalamalasetty et al., 2011). Total RNA was isolated using RNeasy mini kit (Qiagen), reverse transcription and qPCR were set up using iScript cDNA synthesis kit and SsoAdvanced Universal SYBR Green Supermix (Bio-Rad), respectively. Oligos used in this study are described in supplementary material Table S5. Detailed experimental details can be found in the methods in the supplementary materials.

Flow cytometry analysis

EBs were washed in PBS and briefly trypsinized with 0.25% Trypsin-EDTA for 1 min to dissociate into single cells. After neutralizing in DMEM, 10% FBS medium, cells were washed in ice-cold SORT buffer (PBS, penicillin-streptomycin, 0.2% Fraction-V BSA) and stained with antibodies for 20 min on ice. Cells were washed twice with SORT buffer and fixed in 1% PFA. For analysis, cells were passed through cell strainer and analyzed using a BD FACS machine. Raw data was imported to FlowJo software for extensive analysis. Antibodies used in flow cytometry are described in Table S4 in the supplementary materials.

Gene expression profiling, ChIP-qPCR and ChIP-Seq studies

For microarray analysis, day 2 EBs formed from iF-Msgn1 ESC were treated with 1 μg/ml of Dox and samples were subsequently collected at 12 h, 24 h, and 48 h time points. Microarray and ChIP-seq analyses were performed as described previously (Chalamalasetty et al., 2011). Briefly, ChIP-seq was performed on day 2 EBs treated with Dox for an additional 36 h, fixed in 1% formaldehyde solution for 15 min, quenched with 0.125 M glycine and flash-frozen. Nuclear extracts were incubated with anti-Flag M2 antibody (Sigma), DNA fragments were purified and amplified using Illumina ChIP-seq DNA sample prep kit (Illumina), and DNA libraries were sequenced on a Genome Analyzer II. 35-nt sequence reads were mapped to the mouse genome (build mm9) using the Eland algorithm. Validation of ChIP-seq peaks was done by ChIP-qPCR on iF-Msgn1 EBs and E9.5 PSM extracts using control (Clone: NS-1; Sigma) and anti-Msgn1 ascites (Chalamalasetty et al., 2011) as described previously. Detailed experimental details can be found in the methods in the supplementary materials.

Accession numbers

Microarray and ChIP-seq data reported in this paper are deposited in NCBI GEO database with accession numbers GSE29848 and GSE55263, respectively.

Acknowledgements

We are grateful to J. Yoon and M. Lewandoski for Msgn1 and T-Cre mice, respectively, and to M. Kyba for A2lox ESCs reagents. We thank L. Feigenbaum for

generating transgenic animals, Ruth Wolfe for assistance with animal husbandry, and the Flow Cytometry Core at NCI-Frederick for assistance with FACS experiments. We are grateful to A. Perantoni and M. Lewandoski for critiquing the manuscript.

Competing interests

The authors declare no competing financial interests.

Author contributions

R.B.C. and T.P.Y. conceived the project. R.B.C. and T.P.Y. designed and analyzed the experiments. R.J.G. performed the immunohistochemistry experiments in supplementary material Fig. S4 and *lacZ* sections in Fig. 8. W.C.D., P.J. and H.S. assisted R.B.C. with genomics and analysis. M.W.K. performed some of the luciferase assays. T.P.Y., R.B.C. and R.J.G. wrote the manuscript.

Funding

This research was supported by the Intramural Research Program of the National Institutes of Health, National Cancer Institute, Center for Cancer Research. Deposited in PMC for release after 12 months.

Supplementary material

Supplementary material available online at
http://dev.biologists.org/lookup/suppl/doi:10.1242/dev.110908/-/DC1

References

Barrallo-Gimeno, A. and Nieto, M. A. (2005). The Snail genes as inducers of cell movement and survival: implications in development and cancer. *Development* **132**, 3151-3161.

Batlle, E., Sancho, E., Francí, C., Domínguez, D., Monfar, M., Baulida, J. and García De Herreros, A. (2000). The transcription factor snail is a repressor of E-cadherin gene expression in epithelial tumour cells. *Nat. Cell Biol.* **2**, 84-89.

Belteki, G., Haigh, J., Kabacs, N., Haigh, K., Sison, K., Costantini, F., Whitsett, J., Quaggin, S. E. and Nagy, A. (2005). Conditional and inducible transgene expression in mice through the combinatorial use of Cre-mediated recombination and tetracycline induction. *Nucleic Acids Res.* **33**, e51.

Biris, K. K. and Yamaguchi, T. P. (2014). Two-color in situ hybridization of whole-mount mouse embryos. *Methods Mol. Biol.* **1092**, 17-30.

Brault, V., Moore, R., Kutsch, S., Ishibashi, M., Rowitch, D. H., McMahon, A. P., Sommer, L., Boussadia, O. and Kemler, R. (2001). Inactivation of the beta-catenin gene by Wnt1-Cre-mediated deletion results in dramatic brain malformation and failure of craniofacial development. *Development* **128**, 1253-1264.

Burk, U., Schubert, J., Wellner, U., Schmalhofer, O., Vincan, E., Spaderna, S. and Brabletz, T. (2008). A reciprocal repression between ZEB1 and members of the miR-200 family promotes EMT and invasion in cancer cells. *EMBO Rep.* **9**, 582-589.

Cambray, N. and Wilson, V. (2007). Two distinct sources for a population of maturing axial progenitors. *Development* **134**, 2829-2840.

Cano, A., Pérez-Moreno, M. A., Rodrigo, I., Locascio, A., Blanco, M. J., del Barrio, M. G., Portillo, F. and Nieto, M. A. (2000). The transcription factor snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression. *Nat. Cell Biol.* **2**, 76-83.

Chalamalasetty, R. B., Dunty, W. C., Jr, Biris, K. K., Ajima, R., Iacovino, M., Beisaw, A., Feigenbaum, L., Chapman, D. L., Yoon, J. K., Kyba, M. et al. (2011). The Wnt3a/beta-catenin target gene Mesogenin1 controls the segmentation clock by activating a Notch signalling program. *Nat. Commun.* **2**, 390.

Chapman, D. L. and Papaioannou, V. E. (1998). Three neural tubes in mouse embryos with mutations in the T-box gene Tbx6. *Nature* **391**, 695-697.

Dale, J. K., Malapert, P., Chal, J., Vilhais-Neto, G., Maroto, M., Johnson, T., Jayasinghe, S., Trainor, P., Herrmann, B. and Pourquié, O. (2006). Oscillations of the snail genes in the presomitic mesoderm coordinate segmental patterning and morphogenesis in vertebrate somitogenesis. *Dev. Cell* **10**, 355-366.

De Craene, B. and Berx, G. (2013). Regulatory networks defining EMT during cancer initiation and progression. *Nat. Rev. Cancer* **13**, 97-110.

Dequéant, M. L. and Pourquié, O. (2008). Segmental patterning of the vertebrate embryonic axis. *Nat. Rev. Genet.* **9**, 370-382.

Dodd, J., Jessell, T. M. and Placzek, M. (1998). The when and where of floor plate induction. *Science* **282**, 1654-1657.

Dunty, W. C., Jr, Biris, K. K., Chalamalasetty, R. B., Taketo, M. M., Lewandoski, M. and Yamaguchi, T. P. (2008). Wnt3a/beta-catenin signaling controls posterior body development by coordinating mesoderm formation and segmentation. *Development* **135**, 85-94.

Dunty, W. C., Jr, Kennedy, M. W. L., Chalamalasetty, R. B., Campbell, K. and Yamaguchi, T. P. (2014). Transcriptional profiling of Wnt3a mutants identifies Sp transcription factors as essential effectors of the Wnt/beta-catenin pathway in neuromesodermal stem cells. *PLoS ONE* **9**, e87018.

Eckert, M. A., Lwin, T. M., Chang, A. T., Kim, J., Danis, E., Ohno-Machado, L. and Yang, J. (2011). Twist1-induced invadopodia formation promotes tumor metastasis. *Cancer Cell* **19**, 372-386.

Ema, M., Takahashi, S. and Rossant, J. (2006). Deletion of the selection cassette, but not cis-acting elements, in targeted Flk1-lacZ allele reveals Flk1 expression in multipotent mesodermal progenitors. *Blood* **107**, 111-117.

Esteban, M. A., Bao, X., Zhuang, Q., Zhou, T., Qin, B. and Pei, D. (2012). The mesenchymal-to-epithelial transition in somatic cell reprogramming. *Curr. Opin. Genet. Dev.* **22**, 423-428.

Fior, R., Maxwell, A. A., Ma, T. P., Vezzaro, A., Moens, C. B., Amacher, S. L., Lewis, J. and Saude, L. (2012). The differentiation and movement of presomitic mesoderm progenitor cells are controlled by Mesogenin 1. *Development* **139**, 4656-4665.

Fong, A. P., Yao, Z., Zhong, J. W., Cao, Y., Ruzzo, W. L., Gentleman, R. C. and Tapscott, S. J. (2012). Genetic and epigenetic determinants of neurogenesis and myogenesis. *Dev. Cell* **22**, 721-735.

Fujimoto, H. and Yanagisawa, K. O. (1979). Effect of the T-mutation on histogenesis of the mouse embryo under the testis capsule. *J. Embryol. Exp. Morphol.* **50**, 21-30.

Galceran, J., Farinas, I., Depew, M. J., Clevers, H. and Grosschedl, R. (1999). Wnt3a-/–like phenotype and limb deficiency in Lef1(-/-)Tcf1(-/-) mice. *Genes Dev.* **13**, 709-717.

Guaita, S., Puig, I., Franci, C., Garrido, M., Dominguez, D., Batlle, E., Sancho, E., Dedhar, S., de Herreros, A. G. and Baulida, J. (2002). Snail induction of epithelial to mesenchymal transition in tumor cells is accompanied by MUC1 repression and ZEB1 expression. *J. Biol. Chem.* **277**, 39209-39216.

Harada, N., Tamai, Y., Ishikawa, T., Sauer, B., Takaku, K., Oshima, M. and Taketo, M. M. (1999). Intestinal polyposis in mice with a dominant stable mutation of the beta-catenin gene. *EMBO J.* **18**, 5931-5942.

Iacovino, M., Bosnakovski, D., Fey, H., Rux, D., Bajwa, G., Mahen, E., Mitanoska, A., Xu, Z. and Kyba, M. (2011). Inducible cassette exchange: a rapid and efficient system enabling conditional gene expression in embryonic stem and primary cells. *Stem Cells* **29**, 1580-1588.

Jones, S. (2004). An overview of the basic helix-loop-helix proteins. *Genome Biol.* **5**, 226.

Lim, J. and Thiery, J. P. (2012). Epithelial-mesenchymal transitions: insights from development. *Development* **139**, 3471-3486.

Lindsley, R. C., Gill, J. G., Murphy, T. L., Langer, E. M., Cai, M., Mashayekhi, M., Wang, W., Niwa, N., Nerbonne, J. M., Kyba, M. et al. (2008). Mesp1 coordinately regulates cardiovascular fate restriction and epithelial-mesenchymal transition in differentiating ESCs. *Cell Stem Cell* **3**, 55-68.

McLean, C. Y., Bristor, D., Hiller, M., Clarke, S. L., Schaar, B. T., Lowe, C. B., Wenger, A. M. and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. *Nat. Biotechnol.* **28**, 495-501.

Mugford, J. W., Sipilä, P., McMahon, J. A. and McMahon, A. P. (2008). Osr1 expression demarcates a multi-potent population of intermediate mesoderm that undergoes progressive restriction to an Osr1-dependent nephron progenitor compartment within the mammalian kidney. *Dev. Biol.* **324**, 88-98.

Nicolas, J. F., Mathis, L., Bonnerot, C. and Saurin, W. (1996). Evidence in the mouse for self-renewing stem cells in the formation of a segmented longitudinal structure, the myotome. *Development* **122**, 2933-2946.

Nowotschin, S., Ferrer-Vaquer, A., Concepcion, D., Papaioannou, V. E. and Hadjantonakis, A.-K. (2012). Interaction of Wnt3a, Msgn1 and Tbx6 in neural versus paraxial mesoderm lineage commitment and paraxial mesoderm differentiation in the mouse embryo. *Dev. Biol.* **367**, 1-14.

Palmer, M. B., Majumder, P., Green, M. R., Wade, P. A. and Boss, J. M. (2007). A 3′ enhancer controls snail expression in melanoma cells. *Cancer Res.* **67**, 6113-6120.

Perantoni, A. O., Timofeeva, O., Naillat, F., Richman, C., Pajni-Underwood, S., Wilson, C., Vainio, S., Dove, L. F. and Lewandoski, M. (2005). Inactivation of FGF8 in early mesoderm reveals an essential role in kidney development. *Development* **132**, 3859-3871.

Sakurai, H., Era, T., Jakt, L. M., Okada, M., Nakai, S., Nishikawa, S. and Nishikawa, S.-I. (2006). In vitro modeling of paraxial and lateral mesoderm differentiation reveals early reversibility. *Stem Cells* **24**, 575-586.

Soriano, P. (1997). The PDGF alpha receptor is required for neural crest cell development and for normal patterning of the somites. *Development* **124**, 2691-2700.

Soriano, P. (1999). Generalized lacZ expression with the ROSA26 Cre reporter strain. *Nat. Genet.* **21**, 70-71.

Takada, S., Stark, K. L., Shea, M. J., Vassileva, G., McMahon, J. A. and McMahon, A. P. (1994). Wnt-3a regulates somite and tailbud formation in the mouse embryo. *Genes Dev.* **8**, 174-189.

Takebe, A., Era, T., Okada, M., Jakt, L. M., Kuroda, Y. and Nishikawa, S.-I. (2006). Microarray analysis of PDGFR alpha+ populations in ES cell differentiation culture identifies genes involved in differentiation of mesoderm and mesenchyme including ARID3b that is essential for development of embryonic mesenchymal cells. *Dev. Biol.* **293**, 25-37.

Takemoto, T., Uchikawa, M., Yoshida, M., Bell, D. M., Lovell-Badge, R., Papaioannou, V. E. and Kondoh, H. (2011). Tbx6-dependent Sox2 regulation determines neural or mesodermal fate in axial stem cells. *Nature* **470**, 394-398.

Tzouanacou, E., Wegener, A., Wymeersch, F. J., Wilson, V. and Nicolas, J.-F. (2009). Redefining the progression of lineage segregations during mammalian embryogenesis by clonal analysis. *Dev. Cell* **17**, 365-376.

Development (2014) 141, 4285-4297 doi:10.1242/dev.110908

Wehn, A. K. and Chapman, D. L. (2010). Tbx18 and Tbx15 null-like phenotypes in mouse embryos expressing Tbx6 in somitic and lateral plate mesoderm. *Dev. Biol.* **347**, 404-413.

Whiting, J., Marshall, H., Cook, M., Krumlauf, R., Rigby, P. W., Stott, D. and Allemann, R. K. (1991). Multiple spatially specific enhancers are required to reconstruct the pattern of Hox-2.6 gene expression. *Genes Dev.* **5**, 2048-2059.

Wilkinson, D. G., Bhatt, S. and Herrmann, B. G. (1990). Expression pattern of the mouse T gene and its role in mesoderm formation. *Nature* **343**, 657-659.

Wilson, V., Olivera-Martinez, I. and Storey, K. G. (2009). Stem cells, signals and vertebrate body axis extension. *Development* **136**, 1591-1604.

Wittler, L., Shin, E.-h., Grote, P., Kispert, A., Beckers, A., Gossler, A., Werber, M. and Herrmann, B. G. (2007). Expression of Msgn1 in the presomitic mesoderm is controlled by synergism of WNT signalling and Tbx6. *EMBO Rep.* **8**, 784-789.

Yabe, T. and Takada, S. (2012). Mesogenin causes embryonic mesoderm progenitors to differentiate during development of zebrafish tail somites. *Dev. Biol.* **370**, 213-222.

Yamaguchi, T. P., Takada, S., Yoshikawa, Y., Wu, N. and McMahon, A. P. (1999). T (Brachyury) is a direct target of Wnt3a during paraxial mesoderm specification. *Genes Dev.* **13**, 3185-3190.

Yoon, J. K. and Wold, B. (2000). The bHLH regulator pMesogenin1 is required for maturation and segmentation of paraxial mesoderm. *Genes Dev.* **14**, 3204-3214.

Yoon, J. K., Moon, R. T. and Wold, B. (2000). The bHLH class protein pMesogenin1 can specify paraxial mesoderm phenotypes. *Dev. Biol.* **222**, 376-391.

Yoshikawa, Y., Fujimori, T., McMahon, A. P. and Takada, S. (1997). Evidence that absence of Wnt-3a signaling promotes neuralization instead of paraxial mesoderm development in the mouse. *Dev. Biol.* **183**, 234-242.