© 2019. Published by The Company of Biologists Ltd | Development (2019) 146, dev168161. doi:10.1242/dev.168161

RESEARCH ARTICLE

Transcriptionally dynamic progenitor populations organised around a stable niche drive axial patterning

Filip J. Wymeersch[1,2], Stavroula Skylaki[3], Yali Huang[1], Julia A. Watson[1], Constantinos Economou[1], Carylyn Marek-Johnston[1], Simon R. Tomlinson[1] and Valerie Wilson[1,*]

ABSTRACT

The elongating mouse anteroposterior axis is supplied by progenitors with distinct tissue fates. It is not known whether these progenitors confer anteroposterior pattern to the embryo. We have analysed the progenitor population transcriptomes in the mouse primitive streak and tail bud throughout axial elongation. Transcriptomic signatures distinguish three known progenitor types (neuromesodermal, lateral/paraxial mesoderm and notochord progenitors; NMPs, LPMPs and NotoPs). Both NMP and LPMP transcriptomes change extensively over time. In particular, NMPs upregulate Wnt, Fgf and Notch signalling components, and many Hox genes as progenitors transit from production of the trunk to the tail and expand in number. In contrast, the transcriptome of NotoPs is stable throughout axial elongation and they are required for normal axis elongation. These results suggest that NotoPs act as a progenitor niche whereas anteroposterior patterning originates within NMPs and LPMPs.

KEY WORDS: Mouse, Neuromesodermal, Lateral and paraxial mesoderm, Notochord progenitors, Hox

INTRODUCTION

The anteroposterior axis of the vertebrate embryo emerges in a head-to-tail sequence from a growth zone termed the primitive streak in early embryos, and the tail bud in later embryos. This region, located at the posterior tip of the embryo, continuously produces the spinal cord, notochord, paraxial and lateral/ventral mesoderm over the anteroposterior axis (reviewed by Henrique et al., 2015; Wilson et al., 2009). The progenitors of spinal cord and paraxial mesoderm (neuromesodermal progenitors or NMPs) arise just before the start of somitogenesis, and are maintained in the caudal lateral epiblast (CLE) and node-streak border (NSB) of the primitive streak region and the chordoneural hinge (CNH) of the tail bud until elongation ceases at around 65 somite pairs (s) in the mouse. Population and clonal analyses indicate that these cells behave as a stem cell population (Cambray and Wilson, 2007; Tzouanacou et al., 2009). Specifically, they can produce progenitors that remain in the primitive streak and tail bud, as well as differentiated paraxial mesoderm or neurectoderm. Furthermore, NMPs transplanted from late to early embryos can reset their identity to produce more anterior segments of the axis (Cambray and Wilson, 2002; McGrew et al., 2008). Despite this functional stem cell-like behaviour, there are intriguing temporal changes in gene expression throughout the progenitor region, likely including NMPs (Cambray and Wilson, 2007; Gomez et al., 2008; Olivera-Martinez et al., 2012), and a recent study comparing individual NMPs with their descendant mesoderm at two developmental stages shows that some temporal differences occur in NMPs themselves (Gouti et al., 2017). This suggests that, similar to other 'stem cells' in the embryo that produce distinct differentiated phenotypes over time, such as neural or neural crest stem cells (Temple, 2001; White et al., 2001), NMPs do not strictly self-renew. It is thus unknown whether NMPs, which act as stem cells, in fact contain temporal anteroposterior patterning information.

A number of studies have highlighted a crucial role for Hox genes in anteroposterior axial patterning (reviewed by Deschamps and van Nes, 2005; Mallo et al., 2009). The sequential activation of Hox genes from paralogous group (PG) 1 to PG13 in any of the four vertebrate clusters (HoxA-D), depending on their position within the cluster, is a canonical property of Hox genes, termed temporal collinearity. Temporal collinearity has been demonstrated *in vitro* (Lippmann et al., 2015), and for a minority of mouse Hox genes *in vivo* (Izpisua-Belmonte et al., 1991; Soshnikova and Duboule, 2009; Tschopp et al., 2009). For those few that have been studied in detail, activation begins at the posterior primitive streak and spreads anteriorly (Forlani et al., 2003; Iimura and Pourquié, 2006). However, the spatiotemporal expression of most Hox genes in the mouse progenitor region, and specifically their expression in known progenitor types, remains unclear.

Two further populations of axial progenitors have been described, but characterised in less detail. Notochordal progenitors, which we term here 'NotoPs', are also retained for relatively long periods during axial elongation, in the ventral layer of the node at the anterior end of the primitive streak (Beddington, 1994; Wilson and Beddington, 1996). Although a comprehensive temporal fate map of the lateral plate mesoderm (LPM) has not been reported, fate maps at individual stages show that anterior and trunk LPM progenitors (LPMPs) are present in the primitive streak prior to somitogenesis (Castillo et al., 2016; Kinder et al., 1999; Smith et al., 1994; Taguchi et al., 2014). The posteriormost LPM, forming the peri-cloacal mesenchyme, is derived from LPMPs in the early somite-stage primitive streak (Cambray and Wilson, 2007; Wymeersch et al., 2016). The LPM (and potentially also its progenitors) provides important signals that regulate the transition from production of the trunk to the tail. Interestingly, signalling from NMPs to LPMPs may be important for sustaining and patterning the LPMP population (Aires et al., 2016; Jurberg et al.,

[1]MRC Centre for Regenerative Medicine, Institute for Stem Cell Research, School of Biological Sciences, University of Edinburgh, 5 Little France Drive, Edinburgh EH16 4UU, UK. [2]RIKEN Center for Biosystems Dynamics Research, 2-2-3 Minatojima-minamimachi, Chuo-ku, Kobe, Hyogo 650-0047, Japan. [3]Department of Biosystems Science and Engineering, ETH Zürich, 4058 Basel, Switzerland.

*Author for correspondence (v.wilson@ed.ac.uk)

F.J.W., 0000-0001-8999-4555; V.W., 0000-0003-4182-5159

Received 23 May 2018; Accepted 6 December 2018

Development (2019) 146, dev168161. doi:10.1242/dev.168161

2013). Aside from this interaction, however, the interplay between different progenitors as axis elongation proceeds is not clear.

Thus, despite an accumulating body of information, it still remains obscure: (1) when, and in which cells, do known patterning events, such as Hox acquisition, take place; and (2) what interactions occur between different axial progenitor populations?

To answer these questions, we determined the spatiotemporal transcriptome for axial progenitors throughout axis elongation. We find that NMPs, LPMPs and NotoPs show distinct expression profiles, whereas NMPs are similar to their immediate mesoderm-committed descendants. Furthermore, we show transcriptional changes occur in LPMPs and NMPs over time; in the latter, the major change occurs between early somitogenesis and completion of trunk morphogenesis. We also present evidence that NotoPs are a stable integrator of the behaviour of these progenitors.

RESULTS

Temporal differences in transcriptome predominate over progenitor identity

We collected RNA of microdissected embryonic regions according to Fig. 1A, and performed Illumina microarray hybridisation detecting ∼45,000 transcripts. Samples corresponded to regions of known or expected differential fate in and around the primitive streak and tail bud (Table 1; Fig. 1; Materials and Methods). Importantly, although each sample contained a mixture of cell types (Fig. S1A), comparison of samples containing target and non-target cell types allowed us to extract gene expression signatures in cell types of interest. To validate the accuracy and reproducibility of dissection, we performed qRT-PCR on a series of markers with known regional expression on independently dissected regions of the primitive streak (Fig. S1B). We also performed *in situ* hybridisation on primitive streak/tail buds from E9.5-13.5 for known markers (Fig. S2). These analyses showed, in all cases examined, similarity between the intensity values in the microarray analysis and the corresponding independently validated measurements.

We performed unsupervised hierarchical clustering on all samples, expecting separation into fate-based clusters: NMP-containing (NSB, CLE and CNH), exclusively mesoderm-fated NMP descendants (St1, P-CNH) and LPMP-containing regions (E7.5 PP and E8.5 St5). Unexpectedly, NMP regions were more related to their contemporary non-NMP neighbours than to NMP-containing samples at other stages. This suggests that NMPs are not highly transcriptionally divergent from their immediate mesoderm-committed descendants in the primitive streak and tail bud, and that, instead, embryonic age constitutes a distinct transcriptomic signature in NMPs and their descendants. Moreover, samples formed two major clusters: an early group corresponding to stages up to E8.5 and a late group composed of samples from E9.5 onwards. Within the 'early' grouping, the transcriptome of St5 was most divergent from other samples (Fig. 1B).

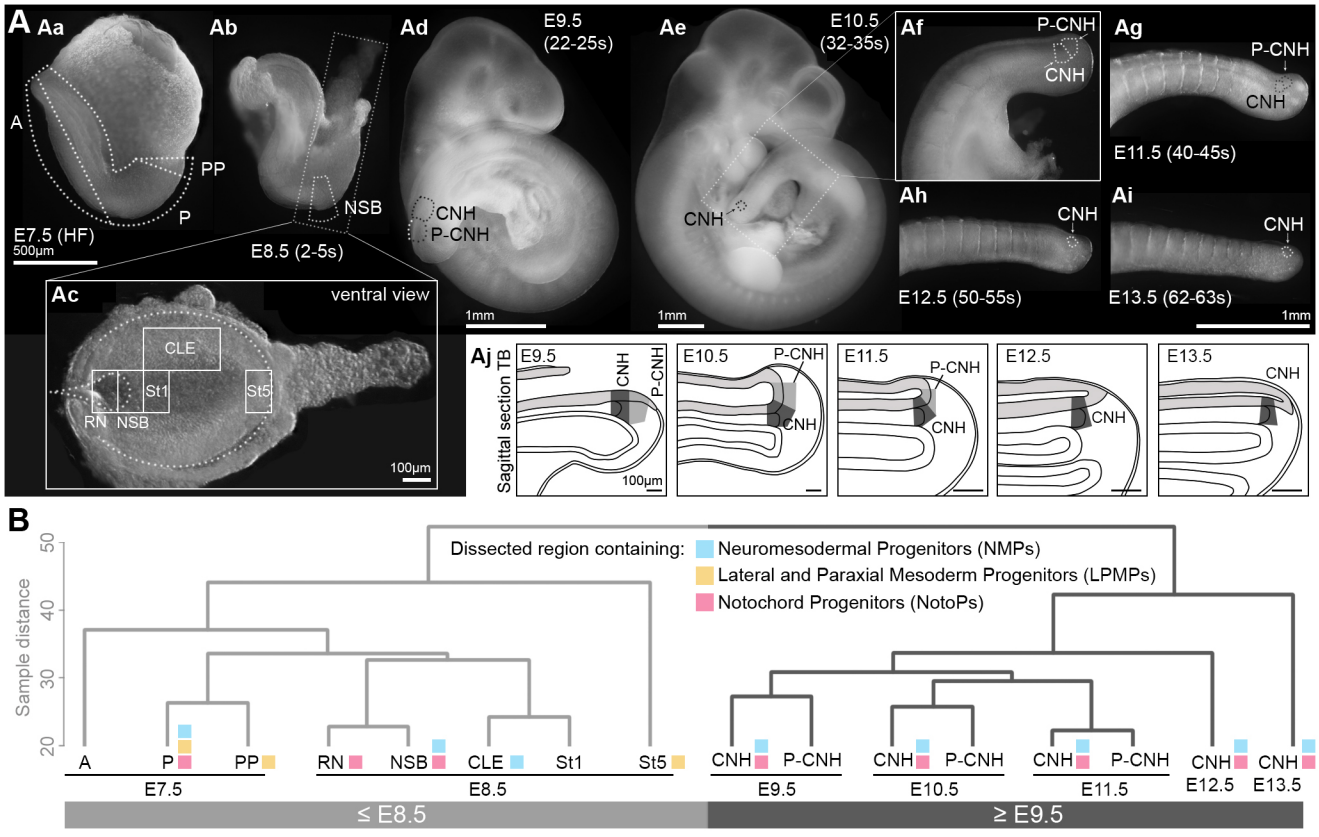


Fig. 1. Microarray analysis of primitive streak and tail bud regions. (A) Schema of dissected embryonic regions. (Aa) In headfold (HF) stage embryos, three regions were isolated: A, anterior; P, posterior; PP, proximal posterior. (Ab) At E8.5 (2-5 s), five regions were isolated as shown in Ac: RN, rostral node; NSB, node-streak-border; St1, rostral 1/5 of the streak; St5, caudalmost 1/5 of the streak; CLE, caudal lateral epiblast beside the rostral 3/5 length of the streak. (Ad-i) The chordoneural hinge (CNH) was dissected in E9.5-E13.5 tail buds. The region just posterior to the CNH (P-CNH) was dissected in E9.5-E11.5 embryos. (Aj) Schema of tail bud (TB) sections with the dissected domains. (B) Unsupervised hierarchical clustering of all samples. Coloured squares highlight the progenitors present in the samples: NMPs (blue), LPMPs (yellow) and NotoPs (pink).

Development (2019) 146, dev168161. doi:10.1242/dev.168161

Table 1. Predominant fate in dissected samples

Age	Stage	Region (abbreviation)		Predominant fate	
E7.5	HF	Anterior	A	Rostral CNS, including fore-, mid- and hindbrain	Cajal et al. (2012)
		Posterior	P	Rest of body	Lawson et al. (1991)*
		Proximal posterior	PP	Allantois and trunk lateral mesoderm up to hindlimb bud	Lawson et al. (1991)*; Forlani et al., (2003)
E8.5	2-5s	Rostral node	RN	Notochord	Cambray and Wilson (2007); Wilson and Beddington (1996); Kinder et al. (1999)
		Node-streak border	NSB	Ventral neurectoderm and somite mesoderm	Cambray and Wilson (2007)
		Caudal lateral epiblast	CLE	Neurectoderm and somite mesoderm	Cambray and Wilson (2007); Wymeersch et al. (2016)
		Anterior primitive streak	St1	Somite mesoderm	Cambray and Wilson (2007); Wilson and Beddington (1996)
		Posterior primitive streak	St5	Ventral mesoderm (cloaca)	Cambray and Wilson (2007); Wilson and Beddington (1996); Kinder et al. (1999); Wymeersch et al. (2016)
E9.5	22-25s	Chordoneural hinge	mCNH	n/d	
		Posterior to chordoneural hinge	pCNH	n/d	
E10.5	32-35s	Chordoneural hinge	mCNH	Neurectoderm and somite mesoderm	Cambray and Wilson (2002)
		Posterior to chordoneural hinge	pCNH	Somite mesoderm	McGrew et al. (2008)
E11.5	40-45s	Chordoneural hinge	mCNH	n/d	
		Posterior to chordoneural hinge	pCNH	n/d	
E12.5	50-55s	Chordoneural hinge	mCNH	Neurectoderm and somite mesoderm	Cambray and Wilson (2002)
E13.5	60-63s	Chordoneural hinge	mCNH	n/d	

Samples are shown in relation to the predominant fate in the region.
*Fate only assessed in streak-stage embryos; n/d, not determined.

Analysis of differential expression at E8.5 reveals spatial domains corresponding to progenitor subpopulations

To investigate transcriptomic differences in known fated regions at E8.5, we first analysed differentially expressed genes (DEGs; fold change (FC)≥1.5, $P \leq 0.05$) between different E8.5 samples (Fig. 2A). Consistent with the clustering analysis, St5 was most divergent from the other E8.5 samples, with over 300 unique DEGs. Upregulated genes included expected markers of the posterior streak (e.g. *Bmp4*, *Cdh5*, *Flt1*, *Hhex*, *Tbx3* and *Tbx4*) (Fong et al., 1996; Fujiwara et al., 2002; Naiche et al., 2011; Scialdone et al., 2016; Thomas et al., 1998). Genes regarded as markers of the primitive streak excluding node and notochord (e.g. *Fgf8*, *Fgf17*, *Wnt3a* and *Wnt5b*) were depleted in St5, underlining its distinct character from the rest of the streak (Cambray and Wilson, 2007; Maruoka et al., 1998; Takada et al., 1994). Genes specifically upregulated in the RN included markers of emergent notochord, such as *Cerl* (Belo et al., 1997). Markers of neural (e.g. *Olig3*) and somite (e.g. *Meox1)* differentiation were also enriched, consistent with the inclusion of emerging neurectoderm and incipient somites as minor cell populations expected in this sample (Fig. S1A). Two endoderm markers, *Sox17* and *Ctsh* (Chen et al., 2013; Kanai-Azuma et al., 2002), were specifically downregulated in the CLE, consistent with the dissection of endoderm away from the CLE. *Tbx6*, a known marker of the primitive streak midline and paraxial mesoderm determinant, and *Dll1*, a known target of *Tbx6* (White and Chapman, 2005), were the only genes showing enrichment in St1, reflecting the paraxial mesoderm fate of this region (Cambray and Wilson, 2007).

The above data provide broad validation of the microarray datasets and suggest that they are appropriate for a search for uniquely enriched transcripts in E8.5 NMPs. Nevertheless, no uniquely upregulated genes were detected in the CLE. Only four genes were upregulated specifically in the NSB. *In situ* hybridisation for those for which unique probes could be designed, *Bhmt2* and *Ccno*, confirmed their specific localisation in the NSB (Fig. 2B). However, their expression was confined to the ventral layer corresponding to the crown of the node, rather than to the NMP-containing dorsal layer. The levels of known NMP markers *T* and *Sox2* correlated well with protein levels measured by immunofluorescence, underlining the accuracy of dissection (Wymeersch et al., 2016) (Fig. 2D). Specifically, levels of *Sox2* correlated with neural fate in the sample, whereas levels of *T* reflected its high expression in the notochord and posterior streak, as well as anterior streak midline (Wymeersch et al., 2016). However, we found no significantly upregulated transcripts in both NSB and CLE, the two NMP-containing areas. Last, we examined the expression of another candidate NMP marker, *Mnx1* (Harrison et al., 1999) (Fig. 2C), which was expressed in the ventral node region at E8.5, but expressed in the E10.5 CNH, making it a potential late NMP marker. Thus, despite the identification of genes with both known and novel differential expression in LPMPs and NotoPs, our analysis did not identify single genes specifically enriched in all NMP-containing areas, either due to a lack of unique markers or because these were below threshold detection levels.

The regional expression of signalling molecules in the analysis above led us to systematically examine the spatial localisation of signalling pathway activity at E8.5. We analysed the Kyoto Encyclopedia of Genes and Genomes (KEGG; Kanehisa et al., 2016) components of the Wnt, Notch, retinoic acid (RA), Nodal, Hedgehog and BMP signalling pathways, which are active in the primitive streak and also involved in axial patterning (reviewed by Wilson et al., 2009). Hierarchical clustering of DEGs belonging to these signalling pathways (≥1.5FC across E8.5 samples) showed three broad domains where signalling molecules were expressed (Fig. 3A): (1) the RN and NSB; (2) the NSB, CLE and St1; and (3) St5. The RN-NSB domain was uniquely characterised by Shh pathway member upregulation, whereas the NSB-CLE-St1 domain showed upregulation of Notch pathway members. *Bmp4* and *Bmp7*

Development (2019) 146, dev168161. doi:10.1242/dev.168161

A

Unique DEGs in the E8.5 caudal region

	Up	Down
RN	1700011H14Rik, Acpl2, Arhgap24, Cer1, Ckb, Cldn11, Col2a1, Ctgf, Fgfr2, Foxc2, Fzd7, Irx3, Kynu, LOC100040592, Meis1, Meox1, Mt1, Myo6, Nbl1, Nin, Olig3, Pcsk9, Rfx4, Sfrp1, Six1, Spsb4, Tgm2, Tmem132e	Chst7, Cldn26, Dusp4, E330027M22Rik, Etv5, Evx1, Fbxl14, Fgf17, Fgf8, Fzd10, Gatsl3, Hoxa7, Klf5, Lmo2, Pid1, Rab8b, Rab8b, Rspo3, Wnt3a, Wnt5b
NSB	Aurkc*, Bhmt2, Ccno, Defcr3*	-
CLE	-	Sox17, Ctsh
St1	Dll1,Tbx6	-
St5	>200 genes	>100 genes

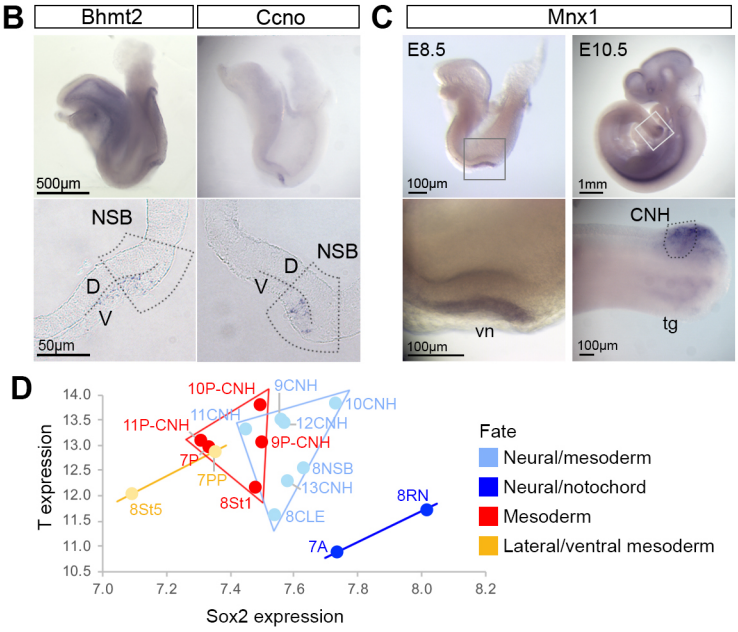


Fig. 2. Spatial analysis of the primitive streak region. (A) Unique DEGs to each E8.5 region (≥1.5-fold change compared with other samples; for unique DEGs in St5, see Table S1). Asterisk indicates that no specific *in situ* probe could be constructed. (B) *Bhmt2* and *Ccno* expression in the ventral (V) but not dorsal (D) NSB layer. (C) Whole-mount *in situ* hybridisation for *Mnx1*. tg, tail gut; vn, ventral node. (D) *Sox2* versus *T* intensity values for all samples. Abbreviations are preceded by embryo age; colours indicates fate. Regions of similar fate are enclosed by a coloured line.

were uniquely upregulated in St5. In addition to these unique pathway components, several Wnt, Fgf, RA and Nodal components were expressed in more than one region. However, the enrichment of individual pathway members respected these domains. Moreover, the Nodal response genes *Lefty1* and *Lefty2*, and the Hedgehog target *Ptch1* were upregulated in the RN-NSB domain, whereas Fgf and Wnt target genes (*Dusp6* and *Axin2* respectively) were upregulated in NSB-CLE-St1. Bmp targets *Id1*, *Id2* and *Id3* were upregulated in St5, indicating that these signalling pathways have localised activation patterns (Fig. 3B).

To determine whether these patterns were specific to signalling pathways known to operate in primitive streak morphogenesis, or reflected more global patterns of gene expression, we analysed all DEGs at E8.5 via hierarchical clustering and explored their function in the STRING online database (Franceschini et al., 2013) (Fig. S3A). This expanded analysis also grouped genes into the three major categories identified above, indicating that these domains contain cells with broadly similar transcriptomes. Interestingly, in the NSB-CLE-St1 cluster, we identified five genes in addition to *Dll1* with modest (<1.5x) upregulation in the midline streak relative to the CLE: *Aph1a*, *Ncstn*, *Ctbp2*, *Dvl1* and *Kat2a*, which are also associated with Notch signalling (KEGGID:04330). In the posterior primitive streak, we observed additional upregulated ligands and receptors involved in vascular development (GO.0072358), including *Pdgfra*, *Adora2b*, *Tgfb1*, *Bmp7*, *Fgf10*, *Efna1*, *Cxcl12*, *Wnt4/5b/6* and *Vegfa*, consistent with the lateral/ventral mesoderm origin of blood vessels (Fig. S3B; Table S2).

Thus, three distinct transcriptomic signatures characterise the E8.5 primitive streak region domains corresponding to the three known progenitor populations: RN-NSB (NotoPs), NSB-CLE-St1 (NMPs and their descendants) and St5 (LPMPs). These data further suggest that the three progenitor types respond to differing signalling pathways that are already known to be functional in axis development. Furthermore, the inclusion of the CLE and midline primitive streak (St1) in a single domain reflects the progression of NMPs towards mesoderm commitment. Intriguingly, it suggests that transcriptional change during this process is minor, although Notch signalling component transcription may increase along with mesoderm commitment.

The LPMP population undergoes temporal change

To examine temporal changes between E7.5-E8.5 LPMPs, we first examined the E7.5 transcriptome. Short Time-series Expression Miner (STEM) analysis (Ernst and Bar-Joseph, 2006) indicated several significantly enriched profiles. Expected enrichment of primitive streak-specific markers (e.g. *Evx1*, *Fgf8*, *T*) (Fig. 4A; Table S3) in the posterior region and neural/emergent notochord markers anteriorly (e.g. *Otx2*, *Pou3f1*, *Chrd* and *Foxd4*) (Iwafuchi-Doi et al., 2012; Tamplin et al., 2008) provided confidence in the validity of these E7.5 samples. In the proximal-posterior region

Development (2019) 146, dev168161. doi:10.1242/dev.168161

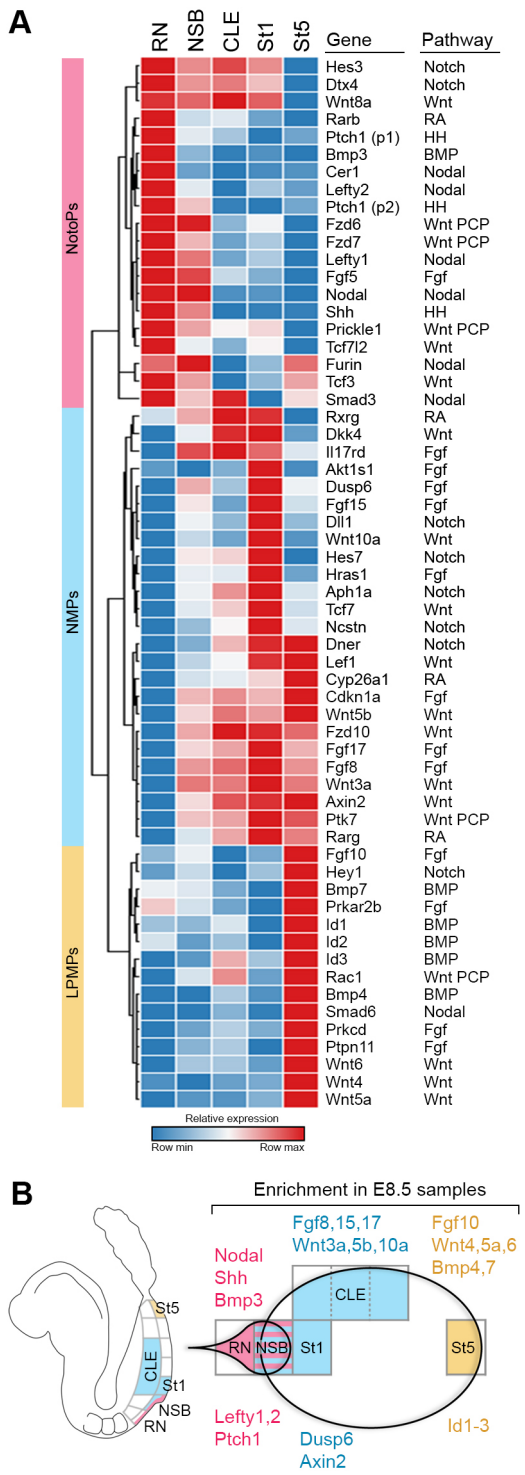


Fig. 3. Spatial domains correspond to progenitor subpopulations. (A) Hierarchical clustering of selected KEGG pathway components reveals three broad domains of signalling (≥1.5-fold change across E8.5 samples), in accordance with the progenitors they contain (NotoPs, NMPs or LPMPs). (B) Schema showing enrichment of signalling ligands and their transcriptional targets in the primitive streak region.

containing LPMPs, known posterior markers, including *Bmp4*, *Mesp1* and *Mixl1* (Arnold and Robertson, 2009), were enriched. Thus, transcriptome differences at E7.5 agree well with previously reported expression patterns, and suggest that LPMPs represent a transcriptionally distinct region at E7.5.

We compared 97 genes specifically upregulated in the proximal posterior region at E7.5 with the 233 genes upregulated in St5 at E8.5. (Fig. 4B,C; Fig. S4; Table S4). A relatively small overlap between these sets (n=24) included known markers of the posterior primitive streak, including *Bmp4*, *Mixl1*, *Tbx4* and *Efna1* (Chapman et al., 1996; Duffy et al., 2006; Lawson et al., 1999; Pearce and Evans, 1999). Enriched GO terms for this overlapping subset of genes included blood vessel morphogenesis and regulation of epithelial cell migration (*Ctsh*, *Tbx4*, *Etv2*, *Ets1*, *Bmp4* and *Efna1*) (GO:0001568 and GO:0010632). Thus, the E7.5 PP and E8.5 St5 regions, which contain prospective lateral mesoderm, show prominent differential gene expression with a subset of shared gene expression that may be useful as markers of LPMPs.

Temporal changes in NMPs

Comparison of NMP-containing regions with their non-NMP neighbours between all samples (up in NSB±CLE and all CNH) did not identify any enriched transcripts. Instead, the NMP transcriptome changed over developmental time (Fig. 1B). We therefore further investigated the transcriptional changes occurring specifically in topologically equivalent NMP-containing regions (the NSB and CNH) over time. Using the E7.5 posterior region as a baseline, we allowed the DEGs at each stage to form self-organising maps (Spielman and Folch, 2015), which highlight gene expression changes as a matrix of patterns and allow visualisation of periods of flux and stability in gene expression profiles (Fig. 5A). This showed that a sharp shift in gene expression occurred between E8.5 and E9.5, and a less prominent change occurred at E12.5-E13.5. We further examined DEGs uniquely up- or downregulated in any one of these samples. The majority of these were found in the E8.5 NSB and E10.5 CNH, with a smaller set at E13.5 (Table S5). Interestingly, a large proportion of the genes that were upregulated at E8.5 were downregulated at E10.5, and vice versa (Fig. 5Ba). The profiles of these genes showed a reciprocal pattern with the majority of the change occurring between E8.5 and E9.5, whereas expression returned towards E7.5 levels after E10.5 (Fig. 5Bb,Bc). Furthermore, most of the genes that did not fall in the intersections between these categories (e.g. up at E8.5 but not down at E10.5) respected the above trend, although the level of up- or downregulation was less than the 1.5-fold cut-off (Fig. S5). In contrast, the genes changing at E13.5 showed no other consistent change earlier in development (Fig. 5C). Thus, the period between E8.5 and E10.5 marks a major transition in gene expression in NMP-containing regions, with most of the shift occurring in the first 24 h of this period.

The genes undergoing transition between E8.5 and E10.5 included known markers of the primitive streak and not the notochord. Furthermore, DEGs enriched in NMPs versus nascent mesoderm in a parallel single cell analysis showed prominent temporal differences (Gouti et al., 2017). These correlated well with our list of DEGs between E8.5 and E10.5 (Fig. S5C,D), supporting the idea that the temporal changes at the NSB and CNH are specific to NMPs and not NotoPs. Genes that were downregulated between E8.5 and E10.5 included pluripotency-associated markers, e.g. *Pou5f1*, *Klf5*, *Lin28*, *Dnmt3b* and *Zscan10* (Ng and Surani, 2011; Takahashi and Yamanaka, 2006; Wang et al., 2007), and markers of the early primitive streak, such as *Cdh1*, *Cdx1*, *Cited2* and *Fst* (Albano et al., 1994; Dunwoodie et al., 1998; Malaguti et al., 2013; Meyer and Gruss, 1993). Genes that were upregulated at E10.5 included members of signalling pathways known to be expressed widely in the primitive streak and tail bud, e.g. *Wnt5a* and *Fgf8* (Crossley and Martin, 1995; Yamaguchi et al., 1999).

Development (2019) 146, dev168161. doi:10.1242/dev.168161

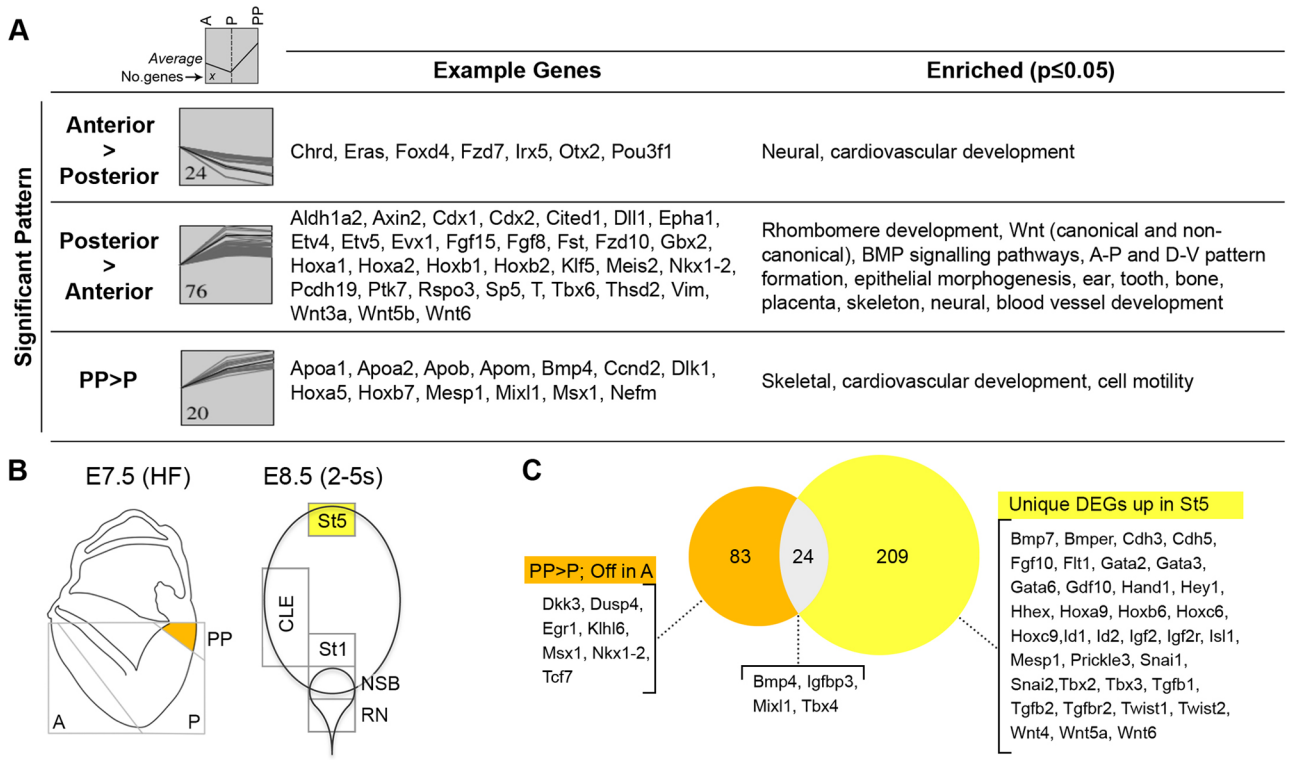


Fig. 4. Temporal transcriptomic changes in LPMPs. (A) Significant patterns at E7.5 and enriched terms ($P \leq 0.05$, defined by permutation test in STEM). Grey boxes show the number of genes in each pattern; the black line indicates their average. (B) Schematic diagram of headfold stage embryo (E7.5 HF) and E8.5 primitive streak area with LPMP-containing regions coloured in yellow. (C) Genes uniquely upregulated in E7.5 PP versus P (orange) and DEGs unique to E8.5 St5 (yellow) outnumber common genes (grey).

STEM analysis, to determine any additional temporal patterns, identified seven significantly enriched patterns (Fig. S6A; Table S3), of which two overlapped with the previously identified set of genes that were downregulated between E8.5 and E10.5 (n_{genes}=160 and 42). A group of 139 genes peaked at E10.5 and overlapped with those upregulated between E8.5 and E10.5. Further groups of 46, 43 and 98 genes showed variations of this pattern with slightly broader peaks, whereas 37 genes peaked at E13.5. The lack of any other patterns indicates that changes between E8.5 and E10.5 constitute the major transcriptional diversity in NMPs over time. A further set of 46 members of transcription factor families or signalling pathways (Hox, Fox, Tbx, Pou, Wnt, Fgf, Notch and RA) was identified by correlation with the pattern of a typical profile for the 139 genes peaking at E10.5, i.e. that of *Wnt5a* (Fig. S6B; Table S6). Combining these datasets generated a list of 313 genes upregulated between E8.5 and E10.5, the expression of which declined thereafter (Fig. S6C).

Analysis of genes upregulated between E8.5 and E10.5

A previous analysis in chick (Olivera-Martinez et al., 2014) focused on the differentiation of NMPs in the 'stem zone' (equivalent to the CLE) towards neural fates. The intersection of our compiled list of 313 genes upregulated at E10.5 with genes upregulated in the chick CLE versus the emerging neural tube contains 16 genes, most of which are known primitive streak markers (Fig. S7A), including Wnt pathway members and their targets (*Wnt5a*, *Wnt5b*, *Rspo3*, *T* and *Evx1*), Fgf pathway members and their targets (*Fgf8*, *Fgf18*, *Il17rd* and *Cyp26a1*), steroid signalling (*Greb1*), and epithelial-to-mesenchymal transition (*Zeb1*). Therefore, the genes upregulated at E10.5 represent a group of evolutionarily conserved primitive streak markers expressed in both NMPs and (as no marker uniquely defines NMPs) their committed NMP descendants.

GO term analysis (Fig. S6D) and manual annotation (Table S7) of this list identified genes associated with the Wnt, Fgf (Ras/Mapk/PI3K/Akt) and Notch signalling pathways, as well as the negative regulator of RA signalling *Cyp26a1*. Wnt, Fgf and Notch signalling pathways are known to be active in the primitive streak/tail bud, and are important for axial elongation, while downregulation of RA synthesis characterises the middle period of axial elongation. However, a coordinated quantitative peak in expression of these genes during trunk morphogenesis has never been reported. In addition, GO and KEGG terms associated with butanoate and steroid metabolism were also enriched. Members of several metabolic pathways were upregulated, including transcripts of the glycolytic enzymes *Eno3* and *Pgm2* recently shown to be enriched in the tail bud relative to anterior PSM, as part of a general upregulation of glycolysis (Oginuma et al., 2017). Furthermore, cell cycle regulators, extracellular matrix molecules (in particular those associated with microfibril formation) and chromatin modifiers were also upregulated in this cohort of genes (Fig. S6D; Table S7). This suggests that E9.5-E10.5 NMPs reach a maximal level of signalling, metabolic and transcriptional regulatory activity as they lay down the posterior trunk and anterior tail bud. Interestingly, the expression profile of the temporally upregulated genes correlates with an expansion in NMP numbers between E8.5 and E9.5, and a subsequent decline between E10.5 and E13.5 (Wymeersch et al., 2016). Several of the genes upregulated at E10.5 (70/313) are bound by Sox2 (13), T (17) or both transcription factors (40) in NMPs (Koch et al., 2017). Moreover, 20/313 genes were targets of β-catenin in human ES cells, and a larger subset was activated by

Development (2019) 146, dev168161. doi:10.1242/dev.168161

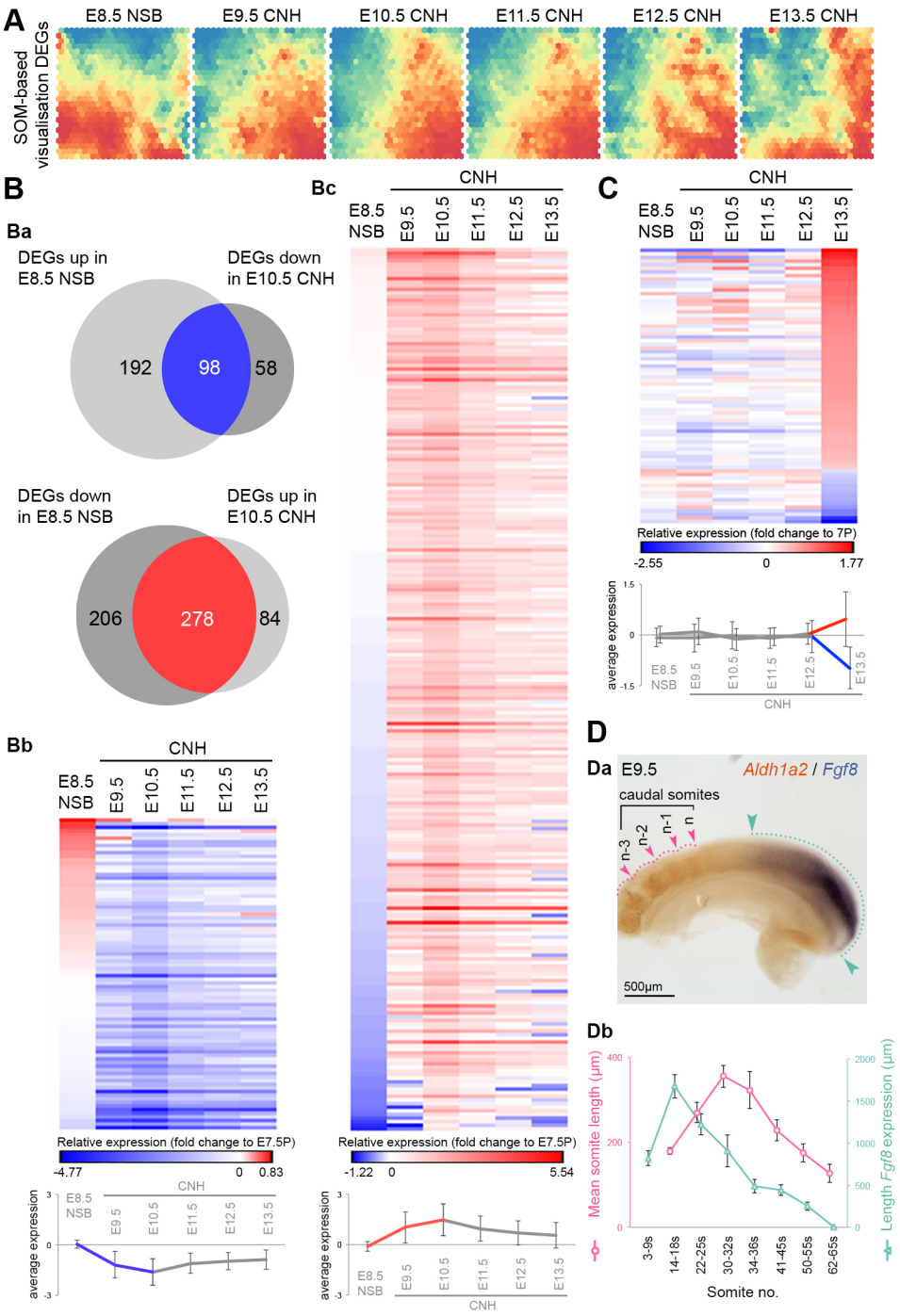


Fig. 5. Transcriptomic changes in NMPs during primitive streak-to-tail bud transition. (A) Self-organising map (SOM)-based visualisation of DEGs. (Ba) Venn diagrams with upregulated DEGs in the E8.5 NSB versus those downregulated in the E10.5 CNH and vice versa (≥1.5-fold change across NMP-containing samples). Expression heatmaps of downregulated (Bb) and upregulated DEGs (Bc) in NMP-containing regions. (C) DEGs in the E13.5 CNH. All heatmap values are relative to 7P, with the mean±s.d. shown below. (Da) *In situ* hybridisation in the E9.5 tail bud for *Fgf8* and *Aldh1a2*. (Db) The anteroposterior length of *Fgf8* expression domain and the length of the four caudalmost somites are shown in relation to somite number (n) and peak at E9.5 (14-18 s) and E10.5 (30-32 s), respectively. Data are mean±s.d. (see also Fig. S8).

CHIR99021-mediated Wnt/β-catenin stimulation of EpiSCs (Fig. S7) (Funa et al., 2015; Tsakiridis et al., 2014). This suggests that activation of Wnt/β-catenin signalling and its direct target T may account for some of the increase in expression of these genes. Interestingly, many β-catenin targets are also members or targets of the Fgf signalling pathway (*Fgf 8*, *Fgf17 Fgf18*, *Dusp6* and *Il17rd*) (Aulehla et al., 2003). Thus, activation of Wnt/β-catenin and/or Fgf signalling could account for this novel mid-trunk expression peak.

We confirmed via whole-mount *in situ* hybridisation that *Fgf8* reaches its maximum intensity of expression, as well as anteroposterior length of its expression domain, at E9.5-E10.5, immediately preceding a peak in somite size (Fig. 5D; Fig. S8). At this stage, expression of the RA synthetic enzyme *Aldh1a2* in the somites was maximally separated from the *Fgf8* expression domain at E9.5, suggesting that the known antagonistic relationship between Fgf8 and RA (Diez del Corral et al., 2003; Sirbu and Duester, 2006) scales with the peak and decline in NMP numbers.

Hox gene expression is upregulated between E8.5-10.5

Several of the genes that were most highly upregulated in NMP- and LPMP-containing regions between E8.5 and E10.5 were members of the Hox gene family, which regulate anteroposterior axial pattern (reviewed by Mallo et al., 2010). Wnt signalling has recently been shown to activate the 3′ (anteriorly expressed) part of the HoxA cluster and facilitate the activation of more 5′, Cdx-dependent, trunk Hox genes (Amin et al., 2016; Neijts et al., 2017). However, no

Development (2019) 146, dev168161. doi:10.1242/dev.168161

general upregulation has been reported *in vivo* between E8.5 and E10.5. Furthermore, chick Hox clusters are activated in a temporally collinear sequence *in vivo* (Denans et al., 2015) but information on the exact time of activation of these genes in mouse embryos is incomplete. We therefore examined our whole microarray dataset for patterns and timing of Hox activation and further upregulation. Hox genes can be broadly classified into anterior (paralogous groups, PG1-3), central (PG4-8) and posterior (PG9-13) subgroups (reviewed by Young and Deschamps, 2009), with PG13 (here referred to as 'terminal') proposed to precipitate the decision to stop axis elongation (Young et al., 2009). After filtering out probes that did not show activation above background (Fig. S9I), genes were grouped according to the above criteria (Fig. 6A).

PG1-3 genes were highly expressed in the E7.5 headfold stage primitive streak. Their expression declined gradually after E9.5 (Fig. 6B; Fig. S9A). The central Hox group (PG4-8) was active at headfold stage, and increased prominently on E8.5. Expression peaked on E9.5 and subsequently declined. PG9-12 expression was low at E7.5, slightly elevated in the E8.5 posterior streak and strongly upregulated at E9.5-E10.5, declining from E11.5 onwards (Fig. 6B; Fig. S9B,C). PG4 and PG9 genes showed profiles intermediate between PG1-3/5-8 and PG4-8/9-12, respectively. PG13 gene expression rose between E8.5 and E10.5, and declined thereafter (Fig. 6B; Fig. S9D). Measuring the extent of change relative to the previous day specifically in the NSB-CNH subpopulation showed that Hox gene expression conformed to the pattern described above (Fig. 6C). Taken together with published data (Forlani et al., 2003; Izpisua-Belmonte et al., 1991; Juan and Ruddle, 2003; Scotti and Kmita, 2012; Soshnikova and Duboule, 2009; Tschopp et al., 2009), this indicates that PG1-4 Hox genes are active before the first time point of E7.5, PG5-8 genes are activated around this time, PG9-12 genes are activated between E8.5 and E9.5, and the terminal Hox genes are activated around E9.5. Despite these generalities, subtle differences between members of a given paralogous group, and cluster-specific profiles were evident (Fig. S9E-H), e.g. the profiles of *Hoxc6*, *Hoxc9* and *Hoxc10* were strikingly similar, possibly reflecting cluster-specific regulation (Neijts et al., 2017, 2016).

As the NSB and CNH samples also contained NotoPs, we confirmed that temporal change in the Hox genes occurs specifically in NMPs by measuring the averaged expression of single E8.5-9.5

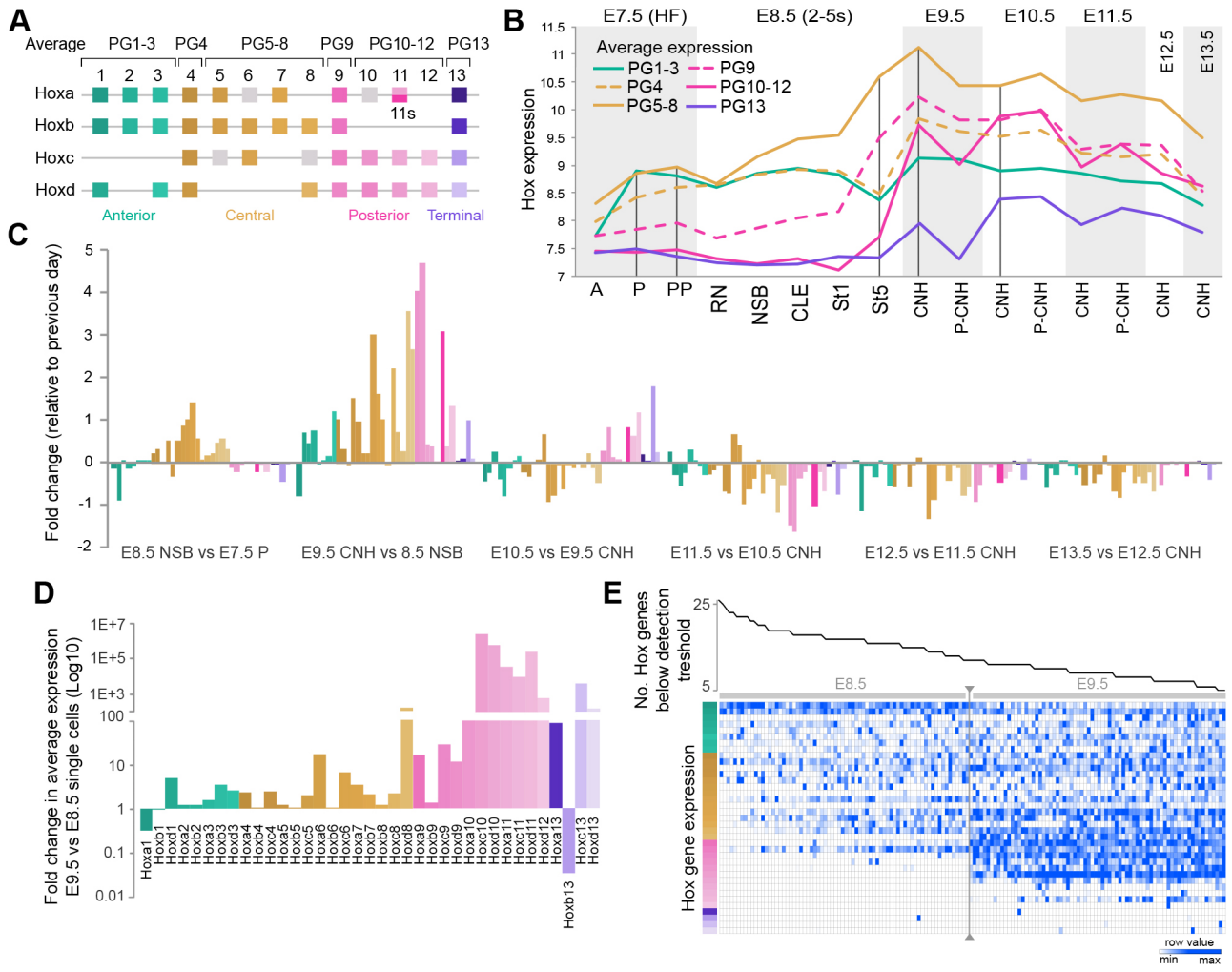


Fig. 6. Hox expression peaks during primitive streak-to-tail bud transition. (A) Hox gene classification into anterior (green), central (gold), posterior (pink) and terminal groups (purple). (B) Timing of Hox gene expression in different paralogous groups (PGs). Expression of individual genes is shown in Fig. S9A-D. (C) Relative levels of Hox transcripts in NMP-containing regions show highest Hox levels are reached between E8.5 and E9.5 in all but PG13 genes. Colours match those in A. (D) Single CLE cell (Gouti et al., 2017) Hox expression values averaged and represented as E9.5 versus E8.5 fold change. (E) Heatmap of single E8.5-E9.5 CLE cells, organised by the number of Hox genes expressed (columns) versus Hox genes (rows). Coloured bar represents the gene order shown in D.

Development (2019) 146, dev168161. doi:10.1242/dev.168161

CLE cells (Gouti et al., 2017) (Fig. 6D). Furthermore, ordering single cells by the number of Hox genes expressed above the detection threshold showed a remarkably consistent correlation with the position of a given Hox gene in the cluster (Fig. 6E), indicating temporal collinearity in individual CLE cells. Thus, Hox gene expression in NMPs shows a general upregulation between E8.5 and E9.5, with a gradual decline between E10.5 and E13.5 (Fig. 6B,C). Furthermore, temporal collinearity implies that anteroposterior patterning originates in NMPs and LPMPs.

Ventral NSB cells are a transcriptomically stable, quiescent and static population essential for axis elongation

As both NMP and LPMP transcriptomes change with time, we investigated whether the same was true of the NotoP population. At E8.5, genes expressed in the RN-NSB domain included known markers of NotoPs. We therefore searched for genes enriched in RN, NSB and all CNH samples that would constitute NotoP markers. This gene set overlapped extensively with genes enriched in cells expressing the NotoP marker Foxa2 (Tamplin et al., 2011) (Fig. S10A; Table S8). Indeed, about half of the genes fulfilling these criteria (23/41) were previously identified markers of the node and emergent notochord (Fig. 7A). Within the remaining half, we validated one of these novel potential NotoP markers, Timp3, a metalloproteinase inhibitor, via antibody staining (Fig. 7B). Interestingly, the levels of expression of these 41 genes in each NotoP-containing sample were relatively stable over time (Fig. 7C). This raised the possibility that the NotoP population, which at early stages coincides with the organiser of the neuraxis and at all stages

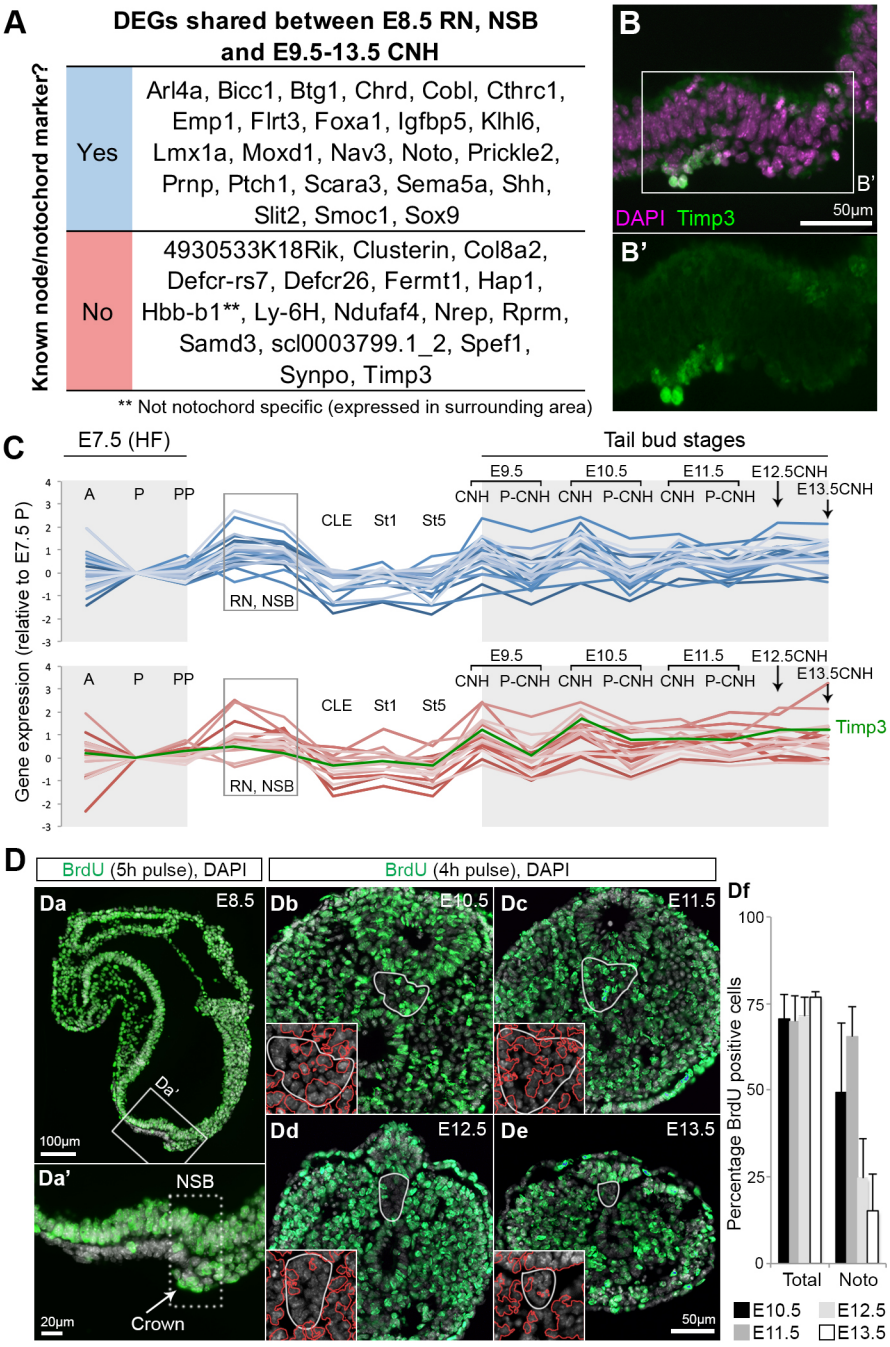


Fig. 7. Quiescence and transcriptional stability of NotoPs. (A) DEGs shared between E8.5 RN and NSB and E9.5-13.5 CNH samples highlight known, and reveal potential novel, node/notochord markers. Two genes, *Sp6* and *Etsrp71* (*Etv2*), were specifically downregulated. (B) Immunostaining for Timp3 shows elevated expression in the ventral node layer. (C) Genes in A plotted over time. Known markers are in blue; potential novel markers are in red; Timp3, green. (D) Immunohistochemistry detecting incorporated BrdU (green) in E8.5 (2-3 s) embryo and tail sections. White outline indicates the posterior notochord domain. Insets contain DAPI-stained nuclei showing the position of BrdU labelled cells (red). (Df) Quantification of $BrdU^+$ cells in tail sections. Data are mean±s.d.; noto, notochord (see also Fig. S10B).

Development (2019) 146, dev168161. doi:10.1242/dev.168161

contacts NMPs, may stabilise behaviours of adjacent populations throughout axial elongation.

Consistent with this possibility, previous reports have shown that the ventral node and notochordal plate contain slow-dividing or quiescent cells at E7.5-E8.5 (Bellomo et al., 1996), although they are more proliferative at E9.5 (Ukita et al., 2009). To determine the proliferative characteristics of the crown and later notochordal plate, we analysed cells in S-phase after bromodeoxyuridine (BrdU) labelling. Both the crown at early somite stages and the ≥E10.5 notochordal plate contained a mixture of labelled and unlabelled cells. (Fig. 7Da). Nevertheless the proportion of NotoPs in S-phase relative to other tail bud regions during tail development was low (15-60% versus 70% in the surrounding tissues; Fig. 7Db-Df; Fig. S10B). Although the mouse node has been fate mapped as a whole, the fate of the posterior crown region, which lies just ventral to the NMPs, is unknown. Control DiI label of the ventral node resulted in the expected descendants along the length of the notochord as far posteriorly as the notochord end, as well as the dorsal hindgut (Fig. 8Aa-Ad). Descendants of the crown also populated the dorsal hindgut and notochord, although in a more posterior region than the whole node (Fig. 8Ae-Ai; Fig. S11A). This suggests that crown cells contain NotoPs whose exit from the progenitor region is delayed relative to NotoPs in the rostral node. This is consistent with homotopic grafts of the whole NSB, where descendant cells were found in the notochordal plate and posterior part of the notochord (Cambray and Wilson, 2007). The gut was unlabelled in these grafts, suggesting that DiI additionally labels a neighbouring population of dorsal/posterior endoderm progenitors, whereas the cells in contact with the NSB are exclusively notochord progenitors. Consistently, descendants of the crown remained in contact with NMPs (Fig. 8B, Fig. S11B). Thus, the stable gene expression, low proliferation and retention of NotoP descendants in the posterior end of the notochord suggests that NotoPs are ideally placed to provide stable environmental signals to NMPs.

Highly localised electroporation of mouse epiblast cells is accompanied by a small region of cell death mainly in the adjacent outer layer of cells, i.e. the endoderm or the exposed notochordal plate (Fig. S11C; Huang et al., 2015). We exploited this localised cell ablation to investigate the role of the ventral NSB in axial elongation. Electroporation of control CLE or primitive streak resulted in normal development and widespread distribution of electroporated cell descendants, according to their expected fates: neurectoderm and mesoderm from CLE; and mesoderm from primitive streak (Fig. 8Ca-f). Hindgut labelling was also observed, probably due to plasmid uptake by the endoderm (Fig. S12). In contrast, electroporating the NSB resulted in sparsely labelled embryos (Fig. 8Cg-n; Fig. S12), exhibiting two distinct phenotypes (Table S9): severe (Class I; n=6), where axis elongation halted immediately despite apparently viable electroporated putative NMPs dorsal to the crown (Fig. 8Cg-j); and milder (Class II; n=13), where embryos failed to turn, the anteroposterior axis was moderately foreshortened and kinked, and the notochordal plate was wider, ending further anteriorly than in control embryos (Fig. 8Ck-n,D). Electroporated cell descendants populated neurectoderm and occasionally mesoderm (Fig. 8Cj,n; Fig. S12), and were rarely found in the progenitor region. Class II phenotypes were recapitulated in embryos where a small area including the ventral NSB layer had been manually removed, but not in controls where an equivalent area of endoderm under the mid-primitive streak was removed (Fig. S13). Thus, the E8.5 crown region of the ventral node is essential for normal axial elongation. Taken together, these data suggest that the E8.5 crown, and its descendant the posterior notochordal plate, provide a stable environment important for axis elongation, and thus may constitute the equivalent of a 'niche' for NMPs.

DISCUSSION

Comprehensive spatiotemporal analysis of progenitor populations provides novel insights into the progressive production of tissues along the anteroposterior and mediolateral axes of the mouse embryo. We identified characteristic transcriptomes of three known progenitor populations, the NMPs, LPMPs and NotoPs, and discovered major transcriptional shifts in the NMP and LPMP populations during axis elongation. In contrast, the adjacent NotoP population has a largely unchanging trancriptome over this time period, and we propose that they act as a 'niche' for NMPs (Fig. 9).

NMPs, LPMPs and NotoPs are defined by distinct transcriptomes

The three spatial domains identified by transcriptome analysis at E8.5 correspond well with the three axial progenitor cell types previously identified by fate-mapping studies (Cambray and Wilson, 2007; Wilson and Beddington, 1996; Wymeersch et al., 2016). Cells of the CLE have an almost identical transcriptomic profile to their immediate mesoderm-committed descendants in the streak midline. Interestingly, comparison of the transcriptome of NMPs in the chick CLE (stem zone) with their immediate neural-committed descendants in the pre-neural tube also reveals that the majority of changes occur after neural commitment, during differentiation of the pre-neural tube to the neural tube (Olivera-Martinez et al., 2014). As gene expression and function are extensively conserved between mouse and chick, NMP commitment to either neurectoderm or mesoderm may involve only minor transcriptional differences.

Consistent with this idea, the only ≥1.5-fold DEGs between NMPs and mesoderm-committed primitive streak cells are *Tbx6* and *Dll1*, which are upregulated in the anterior primitive streak. Expression of these genes is instrumental in paraxial mesoderm differentiation. Tbx6 enforces paraxial mesoderm differentiation of presumptive NMPs: null mutations in this gene lead to the neural differentiation of already ingressed prospective paraxial mesoderm (Chapman and Papaioannou, 1998; Takemoto et al., 2011). A pivotal position of Tbx6 in mesoderm commitment of NMPs is also suggested by the reciprocal expression of Tbx6 and the NMP marker Sox2 *in vitro* (Gouti et al., 2017). *Dll1*, a Notch ligand involved in somite differentiation, is a known target of Tbx6 (Hrabe de Angelis et al., 1997; White and Chapman, 2005). As Tbx6 also functions downstream of Notch signalling (White et al., 2005), these data suggest Notch signalling promotes paraxial mesoderm commitment of NMPs, while Dll1 upregulation may further reinforce it. Furthermore, several genes that are <1.5-fold upregulated in the primitive streak are also associated with Notch signalling (Fig. S3B), supporting a role for this pathway in mesoderm commitment of NMPs.

Temporal change in LPMPs and NMPs

Our analysis of E7.5 and E8.5 LPMPs highlighted profound temporal changes. Although the posterior primitive streak analysed in our study at headfold stage has not been fate mapped, the posterior streak at the slightly earlier bud stage gives rise to interlimb LPM and, at low frequency, to cells in the E8.5 posterior primitive streak (Kinder et al., 1999; Smith et al., 1994). Therefore there is probably an overlap between LPMPs at E7.5 and E8.5. However, it is unclear whether the divergent transcriptome between

Development (2019) 146, dev168161. doi:10.1242/dev.168161

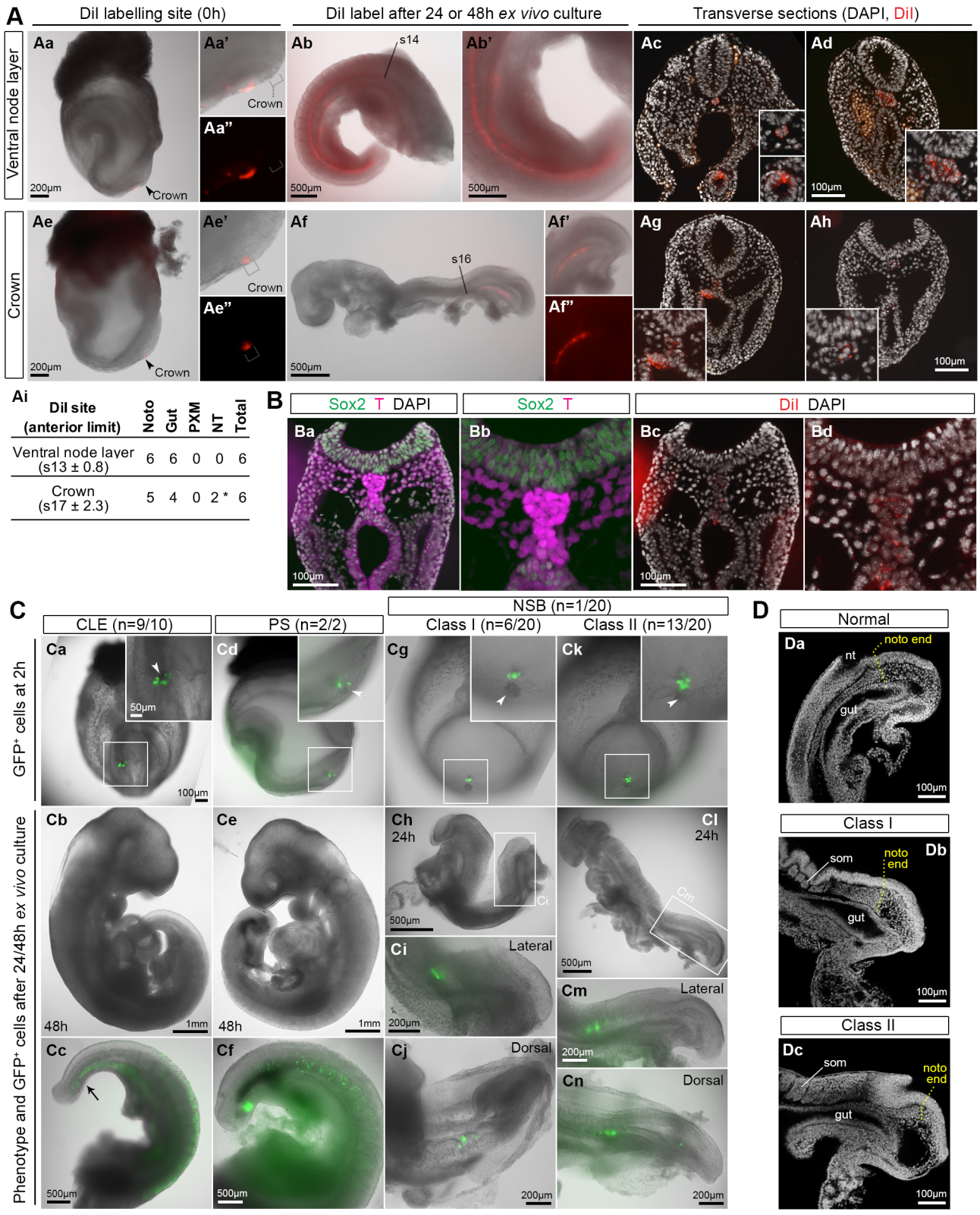


Fig. 8. See next page for legend.

the two regions reflects a single population undergoing maturation (i.e. progressive temporal change in progenitors) or two largely separate cell populations.

In contrast, clonal and population fate mapping shows that late-stage NMPs are largely derived from earlier NMPs (Cambray and Wilson, 2007; Tzouanacou et al., 2009). Therefore, the temporal change in NMP-containing regions implies the maturation of individual NMPs throughout axial elongation. This indicates that the changing intrinsic properties of NMPs (indicated by collinear Hox expression in NMPs) may result in regionalisation of their

Development (2019) 146, dev168161. doi:10.1242/dev.168161

Fig. 8. NotoPs are essential for correct axis elongation. (A) DiI labelling of the ventral node or crown at E8.5 (2-5 s; Aa,Ae). (Aa′,Ae′) Magnified view of the ventral node region. (Aa″,Ae″) Red channel showing DiI labelling. The same embryos are shown after 48 h (Ab) or 24 h (Af) *ex vivo* culture. Ab′ and Af′ show magnified views of Ab and Af respectively. (Af″) Red channel showing DiI labelling. DiI was found in the notochord and dorsal gut (Ac,Ad,Ag,Ah). Insets in Ac, Ad, Ag and Ah show magnified views of DiI labelling in notochord and gut. (Ai) DiI labelling per embryo after culture with average anterior limit (±s.d.). The anterior limit in 4/6 crown-labelled embryos was in the presomitic mesoderm. The anterior limit in each of these embryos is denoted as (total somites/embryo) +1. As the presomitic mesoderm contains ∼7 presomites, the average anterior limit may be posterior to s17. Noto, notochord; NT, neural tube; PXM, paraxial mesoderm. Asterisk indicates 2/6 embryos had a minor contribution in the posterior neural tube. (B) Sox2/T immunostained section of embryo labelled with DiI in the crown after 24 h in culture ($n_{embryos}$=3). (C) Electroporation of GFP-containing plasmid in the CLE (Ca-Cc), primitive streak (Cd-Cf) and NSB (Cg-Cn) of E8.5 (2-5 s) embryos with n, the number of embryos developing normally/total cultured. (Ca-c,Cd-f) Representative CLE- and primitive streak-electroporated embryos, respectively, after 2 or 48 h. (Cg-j,Ck-n) Representative Class I and Class II embryos, respectively, 24 h after NSB electroporation. Arrowheads indicate cell death after electroporation. Black arrow in Cc indicates hindgut label (see Fig. S12). (D) Sagittal confocal sections through CLE-electroporated (Da), and NSB-electroporated class I (Db) and II (Dc) embryos after 24 h, with the notochord (noto) end shown in yellow. nt, neural tube; som, somite.

differentiated derivatives along the axis. Interestingly, Hox expression promotes limb bud outgrowth (Davis et al., 1995; Kmita et al., 2005), suggesting that the general upregulation of Hox genes between E8.5 and E9.5 may be related to the peak in NMP numbers at E9.5 and subsequent expansion of somite size.

A stable niche for axial progenitors

Our observation that ablation of the posterior crown of the node impairs axis elongation recalls experiments carried out in chick (Charrier et al., 1999), which show that at early somitogenesis stages, ablation of both layers of an area approximately equivalent to the mouse NSB also leads to termination of axial elongation. Together, these experiments point to an essential, evolutionarily conserved role of the crown cells in axial elongation. Interestingly, transplantation of cells from the predominantly transient CLE population to the NSB leads to their retention in the CNH (Wymeersch et al., 2016), the only region that is serially transplantable between generations of cultured embryos and can thus be considered resident (Cambray and Wilson, 2002). Therefore, the crown cells may provide a 'niche' that anchors NMPs in the progenitor region. This does not rule out other roles such as in left/right patterning, suggested by the failure of crown-ablated embryos to turn. Although it remains to be determined whether the later notochordal plate organises axial elongation after E8.5 in the same way as the crown cells, the aberrant tail phenotype of embryos where NotoPs are missing (Abdelkhalek et al., 2004) suggests that they are required for later axial elongation.

This 'niche' for NMPs at the node/streak border is the direct descendant of the Spemann organiser equivalent in the earlier embryo: the node (Beddington, 1994; Kinder et al., 1999). The *Xenopus* Spemann organizer, like the mouse node, consists almost exclusively of notochord progenitors (Beddington, 1994; Smith and Slack, 1983), suggesting that it is the equivalent of the NotoP population in mouse. Our finding that NMPs show dynamic gene expression changes, including in Hox expression, while the underlying notochordal plate does not, suggests a model in which anteroposterior axial patterning, via the sequential expression of progressively more posterior genes, is intrinsic to NMPs rather than the organiser. These data are interesting in the context of experiments showing that the *Xenopus* Spemann organiser

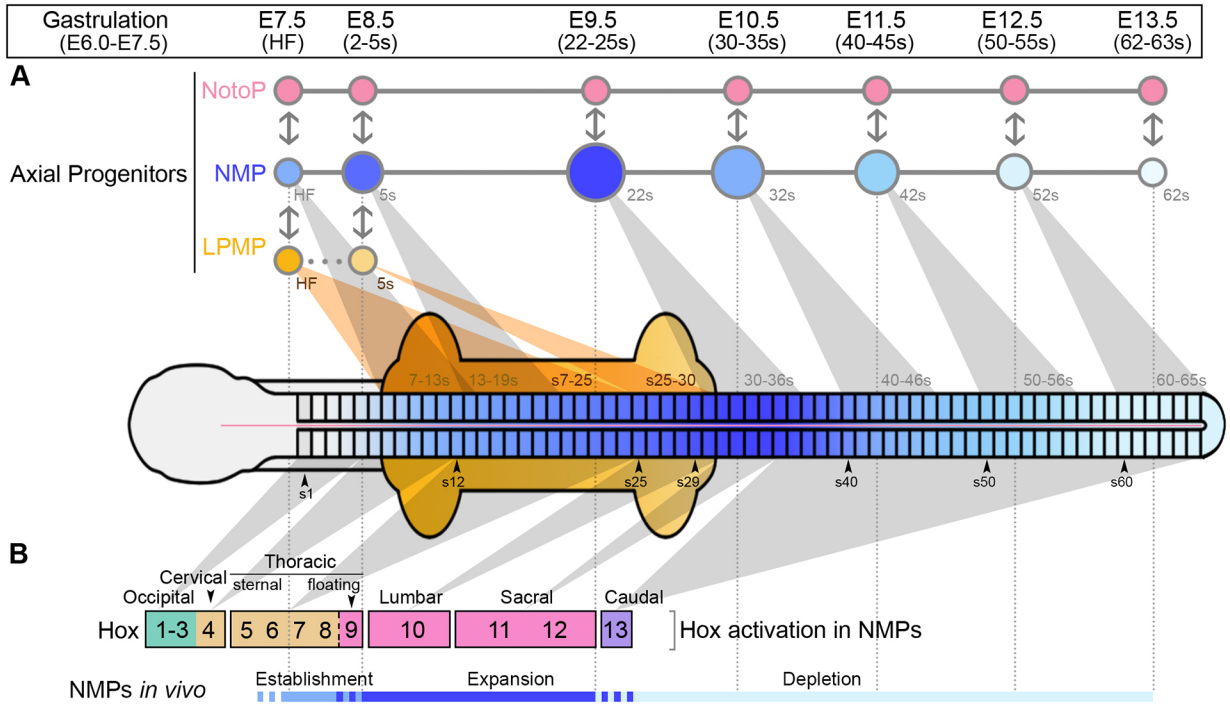


Fig. 9. Summary. Model showing the progression of NMP (blue), LPMP (yellow) and NotoP (pink) populations during axis elongation, the activation of Hox genes and regulatory changes in NMPs over time. (A) Changes in size of the respective populations are indicated by diameter of the circles at each stage; colour changes represent transcriptomic shifts. Contribution of NMPs and LPMPs to specific axial levels is indicated by grey and orange shading, respectively (e.g. an NMP at the 22 s stage will contribute to axial structures at ∼30-36 s) (data from Cambray and Wilson, 2007; Castillo et al., 2016; Tam, 1986; Wymeersch et al., 2016). Double-headed arrows indicate interactions between populations. (B) Activation of Hox genes in NMPs and the vertebrae they pattern (based on Burke et al., 1995; Kuratani, 2009). Colours indicate the paralogous group classifications shown in Fig. 6.

Development (2019) 146, dev168161. doi:10.1242/dev.168161

induces anterior neural character in overlying ectoderm but not detailed anteroposterior pattern (Jansen et al., 2007). Instead, pre-ingression non-organizer mesoderm (the topological equivalent of NMPs and LPMPs), shows intrinsic anteroposterior patterning (reviewed by Durston et al., 2010).

Transfer to a 'young' NMP environment can reset 'old' NMP expression (McGrew et al., 2008). How can this observation be reconciled with an intrinsic timing mechanism in NMPs? Community effects may operate, whereby the number of similar surrounding cells determines whether introduced cells self-differentiate or integrate with their surroundings (Gurdon, 1988; Huang et al., 2012, 2015; McGrew et al., 2008; Trainor and Krumlauf, 2000). This is supported by the observation that, in this population, lack of Cdx2 can be overcome by neighbouring cells (Bialecka et al., 2010), and in zebrafish, axial progenitors (which presumably include NMPs) create a Wnt-dependent environment whereby Brachyury mutation can be tolerated (Martin and Kimelman, 2010). In this scenario, the intrinsic timing of NMPs would respond to changes in local extracellular signalling.

In conclusion, we hypothesise that the vertebrate NotoP population, besides producing notochord, serves as a stable point for organisation of NMPs throughout axial elongation, whereas the NMPs, via Wnt-dependent community effects, undergo maturation. This leads to an expansion of progenitor numbers, an increase in their expression of Wnt, Fgf and Notch signalling pathway components, as well as a quantitative increase in Hox genes of all categories from E8.5-E9.5; this sets up the progenitor pool for sacral/caudal somite production and the activation of terminal Hox genes.

MATERIALS AND METHODS

Mouse strains, staging and husbandry

Wild-type, outbred MF1 mice were used for microarray samples. *sGFP* conditional reporter transgenic (Gilchrist et al., 2003) or MF1 mice were used for electroporation. All mice were maintained on a 12 h light/12 h dark cycle. For timed matings, noon on the day of finding a vaginal plug was designated as E0.5. Staging of early mouse embryos was carried out according to Downs and Davies (1993). All animal experiments were performed under the UK Home Office project license PPL60/4435, approved by the Animal Welfare and Ethical Review Panel of the University of Edinburgh and within the conditions of the Animals (Scientific Procedures) Act 1986.

Microdissection and sample preparation

Microdissection of embryonic regions was performed as described previously (Cambray and Wilson, 2007; Wymeersch et al., 2016). Embryonic regions of a single type were pooled to constitute one sample with at least two replicate samples per embryonic region (Fig. S1A). Specifically, we collected three regions in E7.5 embryos: an anterior neural-fated region (A); and a posterior region comprising the rest of the embryo, including the primitive streak (P) and the posteriormost primitive streak (PP). At E8.5, we analysed the rostral node (RN), node-streak border (NSB), the rostral 1/5 of the primitive streak (St1), the rostral 3/5 of the caudal lateral epiblast (CLE) and the posterior 1/5 of the primitive streak (St5). The CLE samples correspond to the L1-3 region in Wymeersch et al. (2016), where the underlying presomitic mesoderm and endoderm were dissected away from the ectoderm. Regions at subsequent stages up to E11.5 included the chordoneural hinge (CNH), and the region immediately posterior to the CNH (P-CNH). At E12.5-E13.5, owing to the small size of the P-CNH region, only the CNH region was collected. Whole regions were isolated rather than germ layer-dissected tissue to ensure as fast a workflow as possible. This meant that along with the target cell types, several expected minor populations were present. For example, endoderm was present in all <E8.5 samples, except CLE, whereas surface ectoderm was expected to be present in St5 and E13.5 CNH. However, in no case was a non-target tissue uniquely associated with a single cell type of interest, and therefore the data could be used to draw conclusions about expression profiles in the target cell types.

Microarray analysis

RNA was isolated using the RNeasy Micro Kit (Qiagen) and labelled and amplified using the Illumina TotalPrep RNA Amplification Kit (Life Technologies). The sample concentrations and quality were determined using a 2100 bioanalyzer (Agilent). Samples were loaded on six MouseWG-6 v2.0 Expression BeadChip arrays (Illumina). Data normalisation was performed using the lumi package in the R statistical environment (Du et al., 2008). Pre-processing steps consisted of a background adjustment, followed by Variance-Stabilizing Transformation (VST) and Robust Spline Normalization (RSN). A final quality control step was carried out to detect outliers, and probes that were not expressed in any samples were filtered (23,569 out of 45,281 probes; detection P-value<0.01). ComBat analysis (Chen et al., 2011) was used to correct for any batch effects. DEGs were identified using the limma package (parameters: BH with fold-change ≥1.5 and FDR ≤0.05; Smyth, 2004). Marker expression in individual microdissected pieces has been shown previously (Cambray and Wilson, 2007; Wymeersch et al., 2016) and is validated by qRT-PCR and *in situ* hybridisation here (Fig. S1; Fig. S2). Hierarchical clustering was performed using the Morpheus visualisation tool (software.broadinstitute.org/morpheus). To assign genes of interest to a specific signalling pathway or cellular process, we used the Kyoto Encyclopedia of Genes and Genomes (Kanehisa et al., 2016) and the STRING database (Franceschini et al., 2013).

In situ hybridisation

Whole-mount *in situ* hybridisation was performed as described previously (Wilkinson, 1998) except that proteinase K treatment was empirically adjusted according to embryo size and stage (time between 5-20 min). Riboprobes were designed against *Bhmt2* (NM_022884.1, nt1189-1917), *Ccno* (NM_001081062.1, nt79-1048), *Rspo3* (NM028351.3, nt628-1740) and *Sall4* (NM201396.2, nt515-984) mRNA sequences. Other riboprobes used included: *Aldh1a2* (Zhao et al., 1996), *Dusp6* (Dickinson et al., 2002), *Fgf8* (Mahmood et al., 1995), *Fgf17* (Maruoka et al., 1998), *Mnx-1* (Szumska et al., 2008), *Shh* (Echelard et al., 1993) and *Wnt3a* (Takada et al., 1994). Measurements of somite length and *Fgf8* and *Aldh1a2* expression domain length were performed on whole-mount images using Volocity software (Perkin Elmer).

Quantitative RT-PCR

For microarray validation, ∼10-15 independently dissected regions of the primitive streak were pooled to make up one sample. Total RNA was isolated using a RNeasy microkit (Qiagen) and cDNA synthesis performed using SuperScript III (Life Technologies). qRT-PCR was performed using Light Cycler 480 SYBR Green I Master Mix (Roche). Expression values were normalized to the expression of the TATA-box binding protein (TBP). Primer sequences can be found in Table S10.

Immunohistochemistry

Embryo cryosectioning, staining and immunofluorescence was performed as described previously (Huang et al., 2012). Primary antibodies (supplier, catalogue number and working concentration) were: anti-Timp3-loop1 (Abcam; ab39184; 5 µg/ml), anti-Sox2 (Abcam; ab92494; 1:200), anti-T (R&D; AF2085; 1 mg/ml) and anti-GFP (Abcam; ab13970; 10 mg/ml). For S-phase analysis, E8.5 (2-5 s) embryos were cultured *ex vivo* in rat serum-containing medium at 37°C for 5 h, containing 31 µg/ml BrdU (BD Biosciences) (Bellomo et al., 1996). Similarly, tail buds (including the PSM and the last two formed somite pairs) from E10.5-E13.5 embryos were cultured for 4 h, but in N2B27 culture medium (Invitrogen). Samples were fixed overnight in 4% paraformaldehyde in PBS at 4°C and cryosectioned. Antigen retrieval was performed with 10 mM sodium citrate (pH 6.0) for 10 min (Tang et al., 2007) and sections were stained with a BrdU Labelling and Detection Kit I (Roche). Cells were counted using Photoshop (Adobe) and ImageJ software (NIH).

Embryo manipulations

Fluorescent cell tracking was performed with CellTracker CM-DiI (Thermo Fisher Scientific) as described previously (Wilson and Beddington, 1996). Electroporation of pCAG-GFP or pCAG-Cre:GFP plasmids was performed

on E8.5 (2-5 s) wild-type or *sGFP* embryos, respectively, using an optimized electroporation method to target small numbers of cells (Huang et al., 2015). Cell death analysis was performed with DRAQ7 dye (Abcam) according to manufacturer's instructions. *Ex vivo* whole-mount embryo culture was performed as described previously (Copp and Cockroft, 1990). After 24/48 h, cultured embryos were dissected, imaged and scored on phenotype and GFP contribution. Scoring criteria for Class I embryos were: failure to elongate and turn; head truncation or malformation; kinked neural tube; and small somites. Class II embryo criteria were: failure to turn; moderate elongation with somite formation; kinked neural tube; open posterior neural plate; and/or small tail bud. Additional electroporation details can be found in Table S9. To remove the ventral cell layer at the NSB or St3, a sharp glass needle was inserted from posterior to the region of interest and pulled ventrally to separate ventral and dorsal cells. The free cell layer was trimmed to remove the crown or St3 endoderm, after which DiI was poured on the site, labelling all exposed cells (Fig. S13).

Image analysis

Whole-mount embryo images were taken on a Nikon AZ100 (Nikon) or Leica M165 FC microscope (Leica). A wide-field Olympus BX61 or Zeiss Observer microscope with fluorescence optics were used to capture images of immunostained cryosections. Confocal imaging was acquired on a Leica TCS SP8 platform (Leica). Image processing was carried out using Adobe Photoshop (Adobe Systems) and ImageJ software (Schneider et al., 2012).

Acknowledgements

We thank Sally Lowell, Anestis Tsakiridis, Guillaume Blin, James Briscoe, Alfonso Martinez-Arias, Donal O'Carroll and all V.W. lab members for critical reading of the manuscript; Carol Manson, John Agnew and Laura Borthwick for help with animal maintenance; Ron Wilkie for technical assistance; and Anne Wiblin and Abcam for advice and reagents.

Competing interests

The authors declare no competing or financial interests.

Author contributions

Conceptualization: F.J.W., C.E., V.W.; Methodology: F.J.W., S.S., Y.H., J.A.W., C.E., C.M.-J.; Validation: F.J.W., C.E.; Formal analysis: F.J.W., S.S., V.W.; Investigation: F.J.W., Y.H., J.A.W.; Resources: F.J.W., S.S.; Data curation: F.J.W., S.S., S.R.T.; Writing - original draft: F.J.W., V.W.; Writing - review & editing: F.J.W., V.W.; Visualization: F.J.W., S.S.; Supervision: V.W.; Project administration: V.W.; Funding acquisition: V.W.

Funding

This work was supported by the Medical Research Council [MR/K011200 and G080297] and a School of Biological Sciences/University of Edinburgh research studentship awarded to F.J.W. Deposited in PMC for immediate release.

Data availability

The raw and normalised data have been deposited in GEO under accession number GSE120870.

Supplementary information

Supplementary information available online at
http://dev.biologists.org/lookup/doi/10.1242/dev.168161.supplemental

References

Abdelkhalek, H. B., Beckers, A., Schuster-Gossler, K., Pavlova, M. N., Burkhardt, H., Lickert, H., Rossant, J., Reinhardt, R., Schalkwyk, L. C., Muller, I. et al. (2004). The mouse homeobox gene Not is required for caudal notochord development and affected by the truncate mutation. *Genes Dev.* **18**, 1725-1736.

Aires, R., Jurberg, A. D., Leal, F., Nóvoa, A., Cohn, M. J. and Mallo, M. (2016). Oct4 is a key regulator of vertebrate trunk length diversity. *Dev. Cell* **38**, 262-274.

Albano, R. M., Arkell, R., Beddington, R. S. and Smith, J. C. (1994). Expression of inhibin subunits and follistatin during postimplantation mouse development: decidual expression of activin and expression of follistatin in primitive streak, somites and hindbrain. *Development* **120**, 803-813.

Amin, S., Neijts, R., Simmini, S., van Rooijen, C., Tan, S. C., Kester, L., van Oudenaarden, A., Creyghton, M. P. and Deschamps, J. (2016). Cdx and T brachyury co-activate growth signaling in the embryonic axial progenitor niche. *Cell Rep* **17**, 3165-3177.

Arnold, S. J. and Robertson, E. J. (2009). Making a commitment: cell lineage allocation and axis patterning in the early mouse embryo. *Nat. Rev. Mol. Cell Biol.* **10**, 91-103.

Aulehla, A., Wehrle, C., Brand-Saberi, B., Kemler, R., Gossler, A., Kanzler, B. and Herrmann, B. G. (2003). Wnt3a plays a major role in the segmentation clock controlling somitogenesis. *Dev. Cell* **4**, 395-406.

Beddington, R. S. (1994). Induction of a second neural axis by the mouse node. *Development* **120**, 613-620.

Bellomo, D., Lander, A., Harragan, I. and Brown, N. A. (1996). Cell proliferation in mammalian gastrulation: the ventral node and notochord are relatively quiescent. *Dev. Dyn.* **205**, 471-485.

Belo, J. A., Bouwmeester, T., Leyns, L., Kertesz, N., Gallo, M., Follettie, M. and De Robertis, E. M. (1997). Cerberus-like is a secreted factor with neutralizing activity expressed in the anterior primitive endoderm of the mouse gastrula. *Mech. Dev.* **68**, 45-57.

Bialecka, M., Wilson, V. and Deschamps, J. (2010). Cdx mutant axial progenitor cells are rescued by grafting to a wild type environment. *Dev. Biol.* **347**, 228-234.

Burke, A. C., Nelson, C. E., Morgan, B. A. and Tabin, C. (1995). Hox genes and the evolution of vertebrate axial morphology. *Development* **121**, 333-346.

Cajal, M., Lawson, K. A., Hill, B., Moreau, A., Rao, J., Ross, A., Collignon, J. and Camus, A. (2012). Clonal and molecular analysis of the prospective anterior neural boundary in the mouse embryo. *Development* **139**, 423-436.

Cambray, N. and Wilson, V. (2002). Axial progenitors with extensive potency are localised to the mouse chordoneural hinge. *Development* **129**, 4855-4866.

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

Castillo, S. D., Tzouanacou, E., Zaw-Thin, M., Berenjeno, I. M., Parker, V. E., Chivite, I., Mila-Guasch, M., Pearce, W., Solomon, I., Angulo-Urarte, A. et al. (2016). Somatic activating mutations in Pik3ca cause sporadic venous malformations in mice and humans. *Sci. Transl. Med.* **8**, 332ra343.

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.

Chapman, D. L., Garvey, N., Hancock, S., Alexiou, M., Agulnik, S. I., Gibson-Brown, J. J., Cebra-Thomas, J., Bollag, R. J., Silver, L. M. and Papaioannou, V. E. (1996). Expression of the T-box family genes, Tbx1-Tbx5, during early mouse development. *Dev. Dyn.* **206**, 379-390.

Charrier, J. B., Teillet, M. A., Lapointe, F. and Le Douarin, N. M. (1999). Defining subregions of Hensen's node essential for caudalward movement, midline development and cell survival. *Development* **126**, 4771-4783.

Chen, C., Grennan, K., Badner, J., Zhang, D., Gershon, E., Jin, L. and Liu, C. (2011). Removing batch effects in analysis of expression microarray data: an evaluation of six batch adjustment methods. *PLoS ONE* **6**, e17238.

Chen, A. E., Borowiak, M., Sherwood, R. I., Kweudjeu, A. and Melton, D. A. (2013). Functional evaluation of ES cell-derived endodermal populations reveals differences between Nodal and Activin A-guided differentiation. *Development* **140**, 675-686.

Copp, A. J. and Cockroft, D. L. (1990). Dissection and culture of postimplantation embryos. In *Postimplantation Mammalian Embryos: A Practical Approach*. (ed. A. J. Copp and D. L. Cockroft), pp. 15-40. IRL Press.

Crossley, P. H. and Martin, G. R. (1995). The mouse Fgf8 gene encodes a family of polypeptides and is expressed in regions that direct outgrowth and patterning in the developing embryo. *Development* **121**, 439-451.

Davis, A. P., Witte, D. P., Hsieh-Li, H. M., Potter, S. S. and Capecchi, M. R. (1995). Absence of radius and ulna in mice lacking hoxa-11 and hoxd-11. *Nature* **375**, 791-795.

Denans, N., Iimura, T. and Pourquie, O. (2015). Hox genes control vertebrate body elongation by collinear Wnt repression. *eLife* **4**, e04379.

Deschamps, J. and van Nes, J. (2005). Developmental regulation of the Hox genes during axial morphogenesis in the mouse. *Development* **132**, 2931-2942.

Dickinson, R. J., Eblaghie, M. C., Keyse, S. M. and Morriss-Kay, G. M. (2002). Expression of the ERK-specific MAP kinase phosphatase PYST1/MKP3 in mouse embryos during morphogenesis and early organogenesis. *Mech. Dev.* **113**, 193-196.

Diez del Corral, R., Olivera-Martinez, I., Goriely, A., Gale, E., Maden, M. and Storey, K. (2003). Opposing FGF and retinoid pathways control ventral neural pattern, neuronal differentiation, and segmentation during body axis extension. *Neuron* **40**, 65-79.

Downs, K. M. and Davies, T. (1993). Staging of gastrulating mouse embryos by morphological landmarks in the dissecting microscope. *Development* **118**, 1255-1266.

Du, P., Kibbe, W. A. and Lin, S. M. (2008). lumi: a pipeline for processing Illumina microarray. *Bioinformatics* **24**, 1547-1548.

Duffy, S. L., Steiner, K. A., Tam, P. P. L. and Boyd, A. W. (2006). Expression analysis of the Epha1 receptor tyrosine kinase and its high-affinity ligands Efna1 and Efna3 during early mouse development. *Gene Expr. Patterns* **6**, 719-723.

Dunwoodie, S. L., Rodriguez, T. A. and Beddington, R. S. P. (1998). Msg1 and Mrg1, founding members of a gene family, show distinct patterns of gene expression during mouse embryogenesis. *Mech. Dev.* **72**, 27-40.

Durston, A. J., Jansen, H. J. and Wacker, S. A. (2010). Review: time-space translation regulates trunk axial patterning in the early vertebrate embryo. *Genomics* **95**, 250-255.

Echelard, Y., Epstein, D. J., St-Jacques, B., Shen, L., Mohler, J., McMahon, J. A. and McMahon, A. P. (1993). Sonic hedgehog, a member of a family of putative signaling molecules, is implicated in the regulation of CNS polarity. *Cell* **75**, 1417-1430.

Ernst, J. and Bar-Joseph, Z. (2006). STEM: a tool for the analysis of short time series gene expression data. *BMC Bioinformatics* **7**, 191.

Fong, G.-H., Klingensmith, J., Wood, C. R., Rossant, J. and Breitman, M. L. (1996). Regulation of flt-1 expression during mouse embryogenesis suggests a role in the establishment of vascular endothelium. *Dev. Dyn.* **207**, 1-10.

Forlani, S., Lawson, K. A. and Deschamps, J. (2003). Acquisition of Hox codes during gastrulation and axial elongation in the mouse embryo. *Development* **130**, 3807-3819.

Franceschini, A., Szklarczyk, D., Frankild, S., Kuhn, M., Simonovic, M., Roth, A., Lin, J., Minguez, P., Bork, P., von Mering, C. et al. (2013). STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res.* **41**, D808-D815.

Fujiwara, T., Dehart, D. B., Sulik, K. K. and Hogan, B. L. (2002). Distinct requirements for extra-embryonic and embryonic bone morphogenetic protein 4 in the formation of the node and primitive streak and coordination of left-right asymmetry in the mouse. *Development* **129**, 4685-4696.

Funa, N. S., Schachter, K. A., Lerdrup, M., Ekberg, J., Hess, K., Dietrich, N., Honore, C., Hansen, K. and Semb, H. (2015). beta-catenin regulates primitive streak induction through collaborative interactions with SMAD2/SMAD3 and OCT4. *Cell Stem Cell* **16**, 639-652.

Gilchrist, D. S., Ure, J., Hook, L. and Medvinsky, A. (2003). Labeling of hematopoietic stem and progenitor cells in novel activatable EGFP reporter mice. *Genesis* **36**, 168-176.

Gomez, C., Özbudak, E. M., Wunderlich, J., Baumann, D., Lewis, J. and Pourquié, O. (2008). Control of segment number in vertebrate embryos. *Nature* **454**, 335-339.

Gouti, M., Delile, J., Stamataki, D., Wymeersch, F. J., Huang, Y., Kleinjung, J., Wilson, V. and Briscoe, J. (2017). A gene regulatory network balances neural and mesoderm specification during vertebrate trunk development. *Dev. Cell* **41**, 243-261.

Gurdon, J. B. (1988). A community effect in animal development. *Nature* **336**, 772-774.

Harrison, K. A., Thaler, J., Pfaff, S. L., Gu, H. and Kehrl, J. H. (1999). Pancreas dorsal lobe agenesis and abnormal islets of Langerhans in Hlxb9-deficient mice. *Nat. Genet.* **23**, 71-75.

Henrique, D., Abranches, E., Verrier, L. and Storey, K. G. (2015). Neuromesodermal progenitors and the making of the spinal cord. *Development* **142**, 2864-2875.

Hrabe de Angelis, M., McIntyre, J., II and Gossler, A. (1997). Maintenance of somite borders in mice requires the Delta homologue Dll1. *Nature* **386**, 717-721.

Huang, Y., Osorno, R., Tsakiridis, A. and Wilson, V. (2012). In Vivo differentiation potential of epiblast stem cells revealed by chimeric embryo formation. *Cell Rep.* **2**, 1571-1578.

Huang, Y., Wilkie, R. and Wilson, V. (2015). Methods for precisely localized transfer of cells or DNA into early postimplantation mouse embryos. *J. Vis. Exp.* **106**, e53295.

Iimura, T. and Pourquié, O. (2006). Collinear activation of Hoxb genes during gastrulation is linked to mesoderm cell ingression. *Nature* **442**, 568-571.

Iwafuchi-Doi, M., Matsuda, K., Murakami, K., Niwa, H., Tesar, P. J., Aruga, J., Matsuo, I. and Kondoh, H. (2012). Transcriptional regulatory networks in epiblast cells and during anterior neural plate development as modeled in epiblast stem cells. *Development* **139**, 3926-3937.

Izpisua-Belmonte, J. C., Falkenstein, H., Dolle, P., Renucci, A. and Duboule, D. (1991). Murine genes related to the Drosophila AbdB homeotic genes are sequentially expressed during development of the posterior part of the body. *EMBO J.* **10**, 2279-2289.

Jansen, H. J., Wacker, S. A., Bardine, N. and Durston, A. J. (2007). The role of the Spemann organizer in anterior-posterior patterning of the trunk. *Mech. Dev.* **124**, 668-681.

Juan, A. H. and Ruddle, F. H. (2003). Enhancer timing of Hox gene expression: deletion of the endogenous Hoxc8 early enhancer. *Development* **130**, 4823-4834.

Jurberg, A. D., Aires, R., Varela-Lasheras, I., Novoa, A. and Mallo, M. (2013). Switching axial progenitors from producing trunk to tail tissues in vertebrate embryos. *Dev. Cell* **25**, 451-462.

Kanai-Azuma, M., Kanai, Y., Gad, J. M., Tajima, Y., Taya, C., Kurohmaru, M., Sanai, Y., Yonekawa, H., Yazaki, K., Tam, P. P. et al. (2002). Depletion of definitive gut endoderm in Sox17-null mutant mice. *Development* **129**, 2367-2379.

Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M. and Tanabe, M. (2016). KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* **44**, D457-D462.

Kinder, S. J., Tsang, T. E., Quinlan, G. A., Hadjantonakis, A. K., Nagy, A. and Tam, P. P. (1999). The orderly allocation of mesodermal cells to the extraembryonic structures and the anteroposterior axis during gastrulation of the mouse embryo. *Development* **126**, 4691-4701.

Kmita, M., Tarchini, B., Zakany, J., Logan, M., Tabin, C. J. and Duboule, D. (2005). Early developmental arrest of mammalian limbs lacking HoxA/HoxD gene function. *Nature* **435**, 1113-1116.

Koch, F., Scholze, M., Wittler, L., Schifferl, D., Sudheer, S., Grote, P., Timmermann, B., Macura, K. and Herrmann, B. G. (2017). Antagonistic activities of Sox2 and brachyury control the fate choice of neuro-mesodermal progenitors. *Dev. Cell* **42**, 514-526 e517.

Kuratani, S. (2009). Modularity, comparative embryology and evo-devo: developmental dissection of evolving body plans. *Dev. Biol.* **332**, 61-69.

Lawson, K. A., Meneses, J. J. and Pedersen, R. A. (1991). Clonal analysis of epiblast fate during germ layer formation in the mouse embryo. *Development* **113**, 891-911.

Lawson, K. A., Dunn, N. R., Roelen, B. A., Zeinstra, L. M., Davis, A. M., Wright, C. V., Korving, J. P. and Hogan, B. L. (1999). Bmp4 is required for the generation of primordial germ cells in the mouse embryo. *Genes Dev.* **13**, 424-436.

Lippmann, E. S., Williams, C. E., Ruhl, D. A., Estevez-Silva, M. C., Chapman, E. R., Coon, J. J. and Ashton, R. S. (2015). Deterministic HOX patterning in human pluripotent stem cell-derived neuroectoderm. *Stem Cell Reports* **4**, 632-644.

Mahmood, R., Bresnick, J., Hornbruch, A., Mahony, C., Morton, N., Colquhoun, K., Martin, P., Lumsden, A., Dickson, C. and Mason, I. (1995). A role for FGF-8 in the initiation and maintenance of vertebrate limb bud outgrowth. *Curr. Biol.* **5**, 797-806.

Malaguti, M., Nistor, P. A., Blin, G., Pegg, A., Zhou, X. and Lowell, S. (2013). Bone morphogenic protein signalling suppresses differentiation of pluripotent cells by maintaining expression of E-Cadherin. *Elife* **2**, e01197.

Mallo, M., Vinagre, T. and Carapuco, M. (2009). The road to the vertebral formula. *Int. J. Dev. Biol.* **53**, 1469-1481.

Mallo, M., Wellik, D. M. and Deschamps, J. (2010). Hox genes and regional patterning of the vertebrate body plan. *Dev. Biol.* **344**, 7-15.

Martin, B. L. and Kimelman, D. (2010). Brachyury establishes the embryonic mesodermal progenitor niche. *Genes Dev.* **24**, 2778-2783.

Maruoka, Y., Ohbayashi, N., Hoshikawa, M., Itoh, N., Hogan, B. L. M. and Furuta, Y. (1998). Comparison of the expression of three highly related genes, Fgf8, Fgf17 and Fgf18, in the mouse embryo. *Mech. Dev.* **74**, 175-177.

McGrew, M. J., Sherman, A., Lillico, S. G., Ellard, F. M., Radcliffe, P. A., Gilhooley, H. J., Mitrophanous, K. A., Cambray, N., Wilson, V. and Sang, H. (2008). Localised axial progenitor cell populations in the avian tail bud are not committed to a posterior Hox identity. *Development* **135**, 2289-2299.

Meyer, B. I. and Gruss, P. (1993). Mouse Cdx-1 expression during gastrulation. *Development* **117**, 191-203.

Naiche, L. A., Arora, R., Kania, A., Lewandoski, M. and Papaioannou, V. E. (2011). Identity and fate of Tbx4-expressing cells reveal developmental cell fate decisions in the allantois, limb, and external genitalia. *Dev. Dyn.* **240**, 2290-2300.

Neijts, R., Amin, S., van Rooijen, C., Tan, S., Creyghton, M. P., de Laat, W. and Deschamps, J. (2016). Polarized regulatory landscape and Wnt responsiveness underlie Hox activation in embryos. *Genes Dev.* **30**, 1937-1942.

Neijts, R., Amin, S., van Rooijen, C. and Deschamps, J. (2017). Cdx is crucial for the timing mechanism driving colinear Hox activation and defines a trunk segment in the Hox cluster topology. *Dev. Biol.* **422**, 146-154.

Ng, H. H. and Surani, M. A. (2011). The transcriptional and signalling networks of pluripotency. *Nat. Cell Biol.* **13**, 490-496.

Oginuma, M., Moncuquet, P., Xiong, F., Karoly, E., Chal, J., Guevorkian, K. and Pourquie, O. (2017). A gradient of glycolytic activity coordinates FGF and Wnt signaling during elongation of the body axis in amniote embryos. *Dev. Cell* **40**, 342-353 e310.

Olivera-Martinez, I., Harada, H., Halley, P. A. and Storey, K. G. (2012). Loss of FGF-dependent mesoderm identity and rise of endogenous retinoid signalling determine cessation of body axis elongation. *PLoS Biol.* **10**, e1001415.

Olivera-Martinez, I., Schurch, N., Li, R. A., Song, J., Halley, P. A., Das, R. M., Burt, D. W., Barton, G. J. and Storey, K. G. (2014). Major transcriptome re-organisation and abrupt changes in signalling, cell cycle and chromatin regulation at neural differentiation in vivo. *Development* **141**, 3266-3276.

Pearce, J. J. and Evans, M. J. (1999). Mml, a mouse Mix-like gene expressed in the primitive streak. *Mech. Dev.* **87**, 189-192.

Schneider, C. A., Rasband, W. S. and Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671-675.

Scialdone, A., Tanaka, Y., Jawaid, W., Moignard, V., Wilson, N. K., Macaulay, I. C., Marioni, J. C. and Gottgens, B. (2016). Resolving early mesoderm diversification through single-cell expression profiling. *Nature* **535**, 289-293.

Scotti, M. and Kmita, M. (2012). Recruitment of 5′ Hoxa genes in the allantois is essential for proper extra-embryonic function in placental mammals. *Development* **139**, 731-739.

Sirbu, I. O. and Duester, G. (2006). Retinoic-acid signalling in node ectoderm and posterior neural plate directs left-right patterning of somitic mesoderm. *Nat. Cell Biol.* **8**, 271-277.

Smith, J. C. and Slack, J. M. (1983). Dorsalization and neural induction: properties of the organizer in Xenopus laevis. *J. Embryol. Exp. Morphol.* **78**, 299-317.

Smith, J. L., Gesteland, K. M. and Schoenwolf, G. C. (1994). Prospective fate map of the mouse primitive streak at 7.5 days of gestation. *Dev. Dyn.* **201**, 279-289.

Smyth, G. K. (2004). Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.* **3**, Article3.
Soshnikova, N. and Duboule, D. (2009). Epigenetic temporal control of mouse Hox genes in vivo. *Science* **324**, 1320-1323.
Spielman, S. and Folch, D. (2015). Social area analysis and self organizing maps. In *Geocomputation, A Practical Primer* (ed. C. Brunsdon and A. Singleton), pp. 152-168. SAGE Publications.
Szumska, D., Pieles, G., Essalmani, R., Bilski, M., Mesnard, D., Kaur, K., Franklyn, A., El Omari, K., Jefferis, J., Bentham, J. et al. (2008). VACTERL/caudal regression/Currarino syndrome-like malformations in mice with mutation in the proprotein convertase Pcsk5. *Genes Dev.* **22**, 1465-1477.
Taguchi, A., Kaku, Y., Ohmori, T., Sharmin, S., Ogawa, M., Sasaki, H. and Nishinakamura, R. (2014). Redefining the in vivo origin of metanephric nephron progenitors enables generation of complex kidney structures from pluripotent stem cells. *Cell Stem Cell* **14**, 53-67.
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.
Takahashi, K. and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663-676.
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.
Tam, P. P. (1986). A study of the pattern of prospective somites in the presomitic mesoderm of mouse embryos. *J. Embryol. Exp. Morphol.* **92**, 269-285.
Tamplin, O. J., Cox, B. J. and Rossant, J. (2011). Integrated microarray and ChIP analysis identifies multiple Foxa2 dependent target genes in the notochord. *Dev. Biol.* **360**, 415-425.
Tamplin, O. J., Kinzel, D., Cox, B. J., Bell, C. E., Rossant, J. and Lickert, H. (2008). Microarray analysis of Foxa2 mutant mouse embryos reveals novel gene expression and inductive roles for the gastrula organizer and its derivatives. *BMC Genomics* **9**, 511.
Tang, X., Falls, D. L., Li, X., Lane, T. and Luskin, M. B. (2007). Antigen-retrieval procedure for bromodeoxyuridine immunolabeling with concurrent labeling of nuclear DNA and antigens damaged by HCl pretreatment. *J. Neurosci.* **27**, 5837-5844.
Temple, S. (2001). The development of neural stem cells. *Nature* **414**, 112-117.
Thomas, P. Q., Brown, A. and Beddington, R. S. (1998). Hex: a homeobox gene revealing peri-implantation asymmetry in the mouse embryo and an early transient marker of endothelial cell precursors. *Development* **125**, 85-94.
Trainor, P. and Krumlauf, R. (2000). Plasticity in mouse neural crest cells reveals a new patterning role for cranial mesoderm. *Nat. Cell Biol.* **2**, 96-102.
Tsakiridis, A., Huang, Y., Blin, G., Skylaki, S., Wymeersch, F., Osorno, R., Economou, C., Karagianni, E., Zhao, S., Lowell, S. et al. (2014). Distinct Wnt-driven primitive streak-like populations reflect in vivo lineage precursors. *Development* **141**, 1209-1221.
Tschopp, P., Tarchini, B., Spitz, F., Zakany, J. and Duboule, D. (2009). Uncoupling time and space in the collinear regulation of Hox genes. *PLoS Genet.* **5**, e1000398.
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.
Ukita, K., Hirahara, S., Oshima, N., Imuta, Y., Yoshimoto, A., Jang, C.-W., Oginuma, M., Saga, Y., Behringer, R. R., Kondoh, H. et al. (2009). Wnt signaling maintains the notochord fate for progenitor cells and supports the posterior extension of the notochord. *Mech. Dev.* **126**, 791-803.
Wang, Z.-X., Kueh, J. L., Teh, C. H.-L., Rossbach, M., Lim, L., Li, P., Wong, K.-Y., Lufkin, T., Robson, P. and Stanton, L. W. (2007). Zfp206 is a transcription factor that controls pluripotency of embryonic stem cells. *Stem Cells* **25**, 2173-2182.
White, P. H. and Chapman, D. L. (2005). Dll1 is a downstream target of Tbx6 in the paraxial mesoderm. *Genesis* **42**, 193-202.
White, P. M., Morrison, S. J., Orimoto, K., Kubu, C. J., Verdi, J. M. and Anderson, D. J. (2001). Neural crest stem cells undergo cell-intrinsic developmental changes in sensitivity to instructive differentiation signals. *Neuron* **29**, 57-71.
White, P. H., Farkas, D. R. and Chapman, D. L. (2005). Regulation of Tbx6 expression by Notch signaling. *Genesis* **42**, 61-70.
Wilkinson, D. G. (1998). *In Situ Hybridization: A Practical Approach*: Oxford University Press.
Wilson, V. and Beddington, R. S. (1996). Cell fate and morphogenetic movement in the late mouse primitive streak. *Mech. Dev.* **55**, 79-89.
Wilson, V., Olivera-Martinez, I. and Storey, K. G. (2009). Stem cells, signals and vertebrate body axis extension. *Development* **136**, 1591-1604.
Wymeersch, F. J., Huang, Y., Blin, G., Cambray, N., Wilkie, R., Wong, F. C. and Wilson, V. (2016). Position-dependent plasticity of distinct progenitor types in the primitive streak. *Elife* **5**, e10042.
Yamaguchi, T. P., Bradley, A., McMahon, A. P. and Jones, S. (1999). A Wnt5a pathway underlies outgrowth of multiple structures in the vertebrate embryo. *Development* **126**, 1211-1223.
Young, T. and Deschamps, J. (2009). Hox, Cdx, and anteroposterior patterning in the mouse embryo. In *HOX Genes* (ed. O. Pourquie), pp. 235-256: Academic Press.
Young, T., Rowland, J. E., van de Ven, C., Bialecka, M., Novoa, A., Carapuco, M., van Nes, J., de Graaff, W., Duluc, I., Freund, J. N. et al. (2009). Cdx and Hox genes differentially regulate posterior axial growth in mammalian embryos. *Dev. Cell* **17**, 516-526.
Zhao, D., McCaffery, P., Ivins, K. J., Neve, R. L., Hogan, P., Chin, W. W. and Drager, U. C. (1996). Molecular identification of a major retinoic-acid-synthesizing enzyme, a retinaldehyde-specific dehydrogenase. *Eur. J. Biochem.* **240**, 15-22.

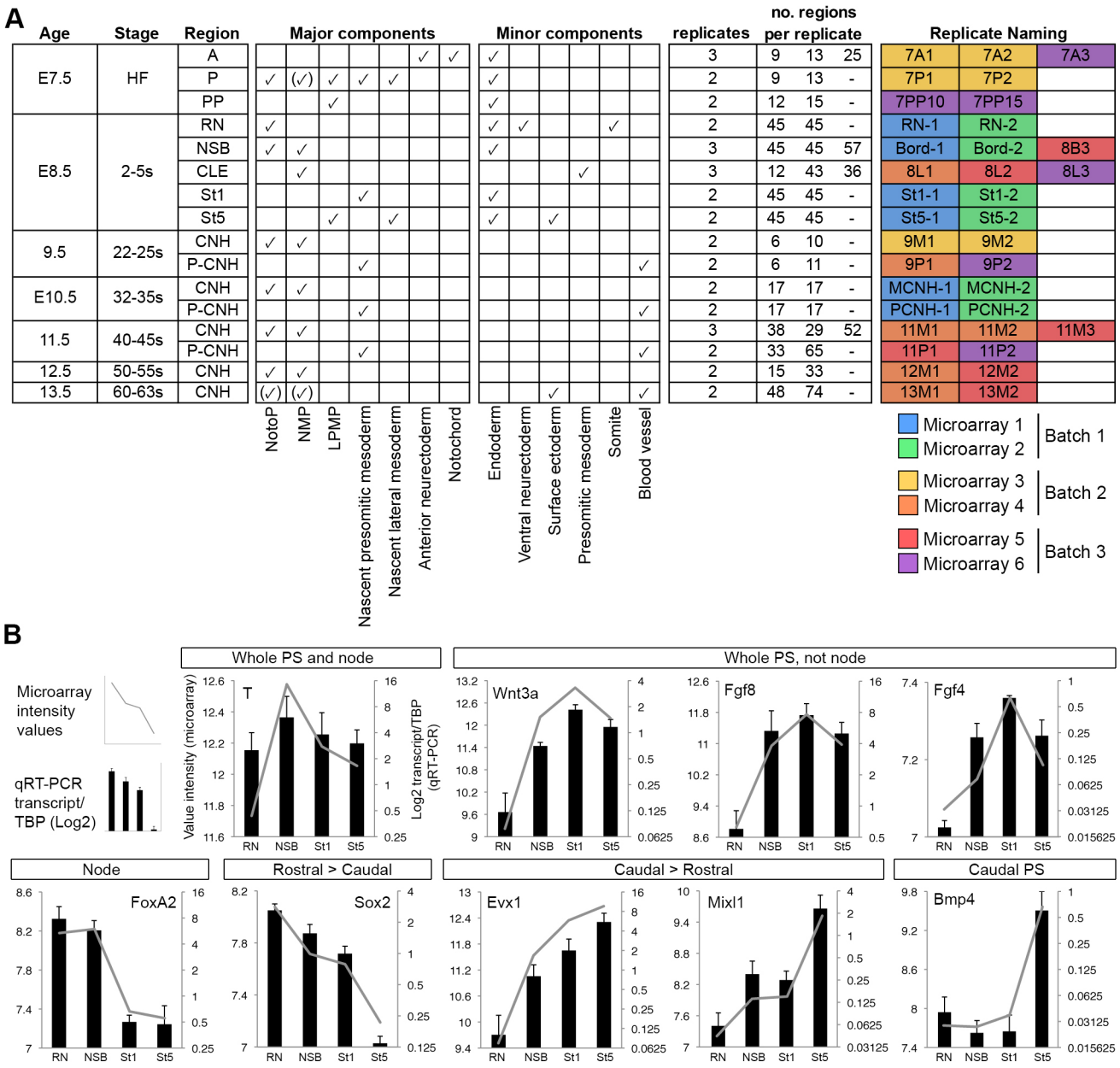


Figure S1. Microarray sample specifics and qRT-PCR validation.
(A) Table of dissected samples, showing major and minor cell components included in the samples. A tick indicates presence, brackets indicate uncertainty about the proportion of progenitors in the sample. Replicates were named as shown and run on six different arrays in three batches. Colour of the sample corresponds to the microarray and batch. (B) Microarray validation on independently-dissected samples consisting of ~10-15 pooled regions of the primitive streak (PS). qRT-PCR values were normalised to TBP, the mean ± s.e.m. plotted as log2 values (black bars) and compared to the normalised intensity values of the microarray (grey line).

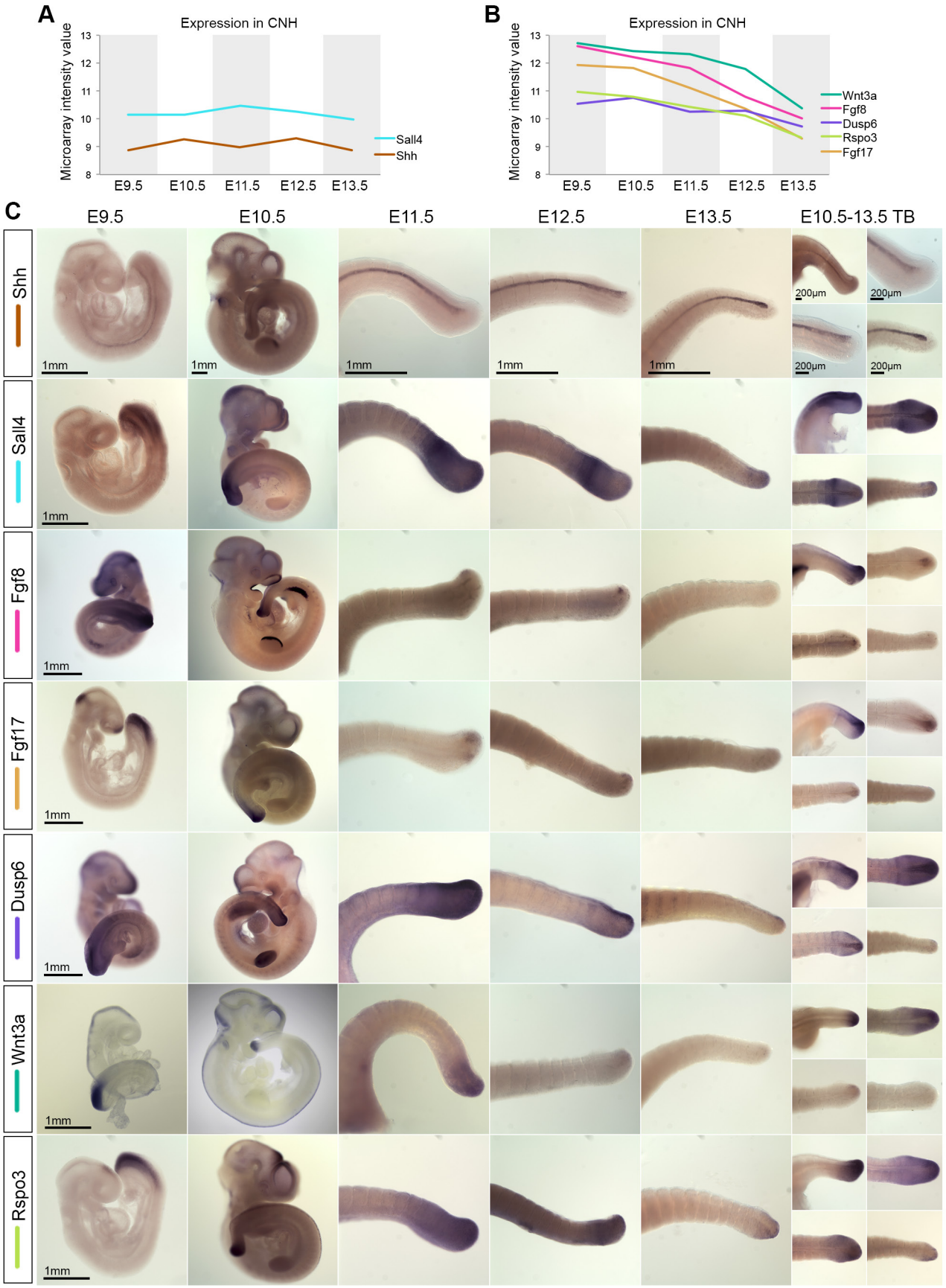
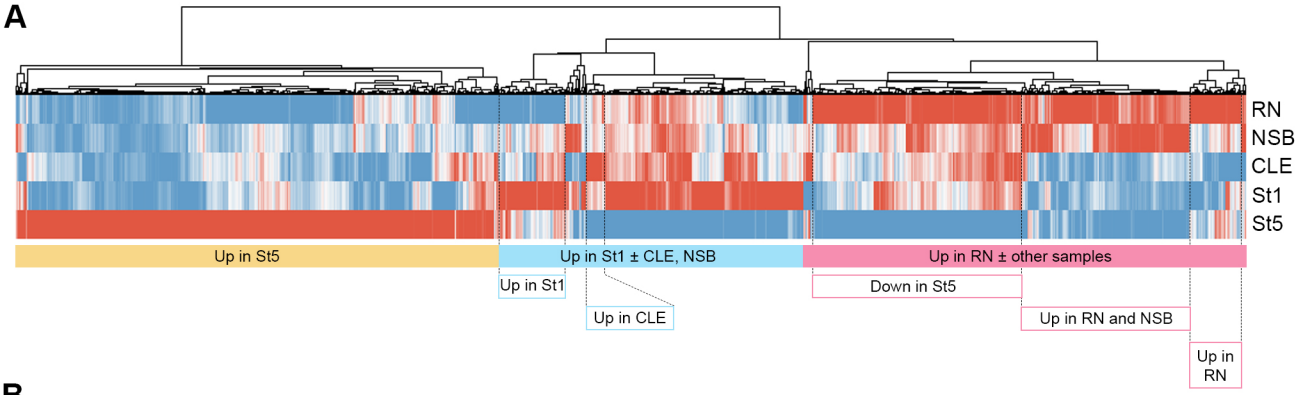


Figure S2. Validation of microarray by in situ hybridisation.
(A-C) Microarray expression values for selected marker genes in the CNH. (A) *Sall4* and *Shh* are stable over time, whereas *Wnt3a, Fgf8, Dusp6, Rspo3* and *Fgf17* decline during tail elongation (B). Microarray intensity values correlate well to *in situ* hybridisation data of the caudal progenitor area between E9.5 and E13.5 (C). CNH, chordoneural hinge; TB, tail bud.

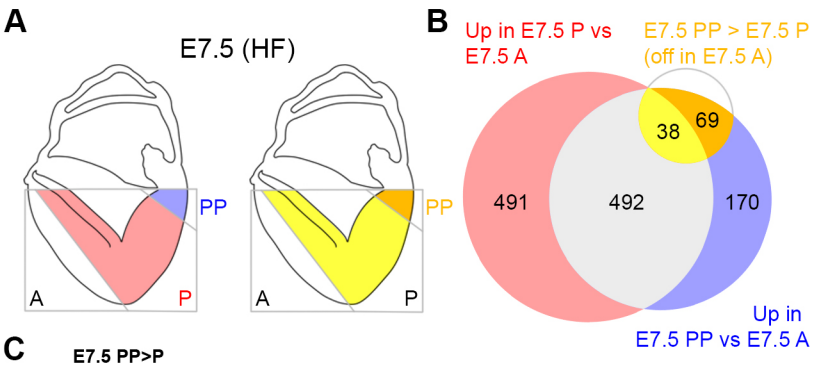
A



B

	GO-term	Description	p-value	KEGG ID	Description	p-value	Genes
Up in RN ± other samples	GO.0030182	neuron differentiation	9.79E-13	5217	Basal cell carcinoma	0.0112	Fzd2, Fzd6, Fzd7, Gli2, Ptch1, Shh, Stk36, Tcf7l2, Wnt8a
	GO.0044782	cilium organization	4.13E-14				
	GO.0098609	cell-cell adhesion	0.0195	100	Steroid biosynthesis	0.0302	Dhcr24, Dhcr7, Lss, Sc4mol, Sqle
	GO.0030336	negative regulation of cell migration	1.65E-05	4512	ECM-receptor interaction	0.033	Agrn, Cd47, Col2a1, Col4a2, Dag1, Itga3, Itga6, Reln, Thbs3, Vtn
	GO.0030178	negative regulation of Wnt signaling pathway	8.40E-06				
	GO.0007368	determination of left/right symmetry	1.16E-06				
	GO.0006694	steroid biosynthetic process	0.000554				
	GO.0016125	sterol metabolic process	0.00179				
	GO.0007219	Notch signaling pathway	0.00344				
	GO.0007492	endoderm development	0.000428				
	GO.0007224	smoothened signaling pathway	0.000674				
	GO.0008203	cholesterol metabolic process	0.00402				
	GO.0021517	ventral spinal cord development	0.000397				
	GO.0030903	notochord development	0.000633				
	GO.0035567	non-canonical Wnt signaling pathway	0.012				
	GO.0060071	Wnt signaling pathway, planar cell polarity pathway	0.00793				
	GO.1903053	regulation of extracellular matrix organization	0.00793				
	GO.0033504	floor plate development	0.0011				
	GO.0010873	positive regulation of cholesterol esterification	0.0224				
Up in St1 ± CLE, NSB	GO.0009893	positive regulation of metabolic process	1.30E-06	4310	Wnt signaling pathway	0.00398	Ctbp2, Dkk4, Dvl1, Frat2, Fzd10, Fzd5, Gpc4, Sox17, Tcf7, Trp53, Wnt10a,
	GO.0016055	Wnt signaling pathway	1.76E-05				
	GO.0001843	neural tube closure	0.000412	4330	Notch signaling pathway	0.0217	Aph1a, Ctbp2, Dll1, Dvl1, Kat2a, Ncstn
	GO.0036342	post-anal tail morphogenesis	0.00463	4014	Ras signaling pathway	0.0301	Angpt2, Epha2, Ets2, Fgf15, Fgf17, Fgf8, Hras1, Plcg2, Rapgef5, Rasa1, Rasa3, Rgl2, Vegfc
	GO.0007155	cell adhesion	0.0181				
	GO.0016477	cell migration	0.0198				
	GO.0030155	regulation of cell adhesion	0.0259	4015	Rap1 signaling pathway	0.047	Angpt2, Epha2, Fgf15, Fgf17, Fgf8, Gnai2, Hras1, Mapk11, Mapk12, Rapgef5, Src, Vegfc
	GO.0001839	neural plate morphogenesis	0.0261				
	GO.0072331	signal transduction by p53 class mediator	0.0316				
	GO.0046328	regulation of JNK cascade	0.0372				
Up in St5	GO.0001944	vasculature development	4.87E-24	4390	Hippo signaling pathway	7.94E-07	Actb, Axin2, Bmp4, Bmp7, Bmpr2, Csnk1d, Frmd6, Id1, Id2, Lef1, Nkd1, Ppp1ca, Ppp2r1a, Snai2, Tgfb1, Tgfb2, Tgfbr2, Wnt4, Wnt5a, Wnt5b, Wnt6, Wtip
	GO.0042981	regulation of apoptotic process	2.41E-13				
	GO.0043066	negative regulation of apoptotic process	3.84E-13				
	GO.0007155	cell adhesion	3.27E-09				
	GO.0060173	limb development	1.23E-06	4350	TGF-beta signaling pathway	2.82E-05	Bmp4, Bmp7, Bmpr2, Id1, Id2, Id3, Id4, Ltbp1, Pitx2, Ppp2r1a, Smad6, Tgfb1, Tgfb2, Tgfbr2
	GO.0043408	regulation of MAPK cascade	0.000307				
	GO.0090090	negative regulation of canonical Wnt signaling pathway	0.00115				
	GO.0038084	vascular endothelial growth factor signaling pathway	0.00155	4015	Rap1 signaling pathway	0.000111	Actb, Adora2b, Akt3, Arap3, Ctnnd1, Efna1, Efna4, F2r, Fgf10, Flt1, Id1, Kit, Ngfr, Pdgfra, Pfn1, Prkd2, Prkd3, Ralb, Rasgrp3, Tek, Tiam1, Vegfa
	GO.0045746	negative regulation of Notch signaling pathway	0.00198				
	GO.0003401	axis elongation	0.00455				
	GO.0030509	BMP signaling pathway	0.0365				
				4151	PI3K-Akt signaling pathway	0.000214	Akt3, Ccne2, Cdkn1a, Col24a1, Efna1, Efna4, F2r, Fgf10, Flt1, Fn1, Ghr, Gnb4, Gng12, Gng2, Hsp90ab1, Itga11, Itga5, Itga9, Itgb5, Kit, Lamc1, Ngfr, Pdgfra, Ppp2r1a, Ppp2r5c, Sgk3, Tek, Vegfa, Vwf
				4014	Ras signaling pathway	0.00392	Akt3, Efna1, Efna4, Ets1, Fgf10, Flt1, Gab1, Gnb4, Gng12, Gng2, Kit, Ngfr, Pdgfra, Ptpn11, Ralb, Rasgrp3, Tek, Tiam1, Vegfa
				4115	p53 signaling pathway	0.0217	Bid, Ccne2, Cdk1, Cdkn1a, Gadd45g, Igfbp3, Sesn1, Shisa5
				4512	ECM-receptor interaction	0.0292	Col24a1, Fn1, Hspg2, Itga11, Itga5, Itga9, Itgb5, Lamc1, Vwf
				4010	MAPK signaling pathway	0.0449	Akt3, Dusp1, Dusp4, Dusp5, Fgf10, Flnc, Gadd45g, Gng12, Map4k4, Nfatc1, Pdgfra, Ppp3ca, Rasgrp3, Rps6ka1, Tgfb1, Tgfb2, Tgfbr2

Figure S3. Gene expression patterns in the E8.5 primitive streak.
(A) Unsupervised hierarchical clustering of all genes that showed ≥1.5 fold change across E8.5 samples. Three major clusters were identified: up in RN ± other samples (pink), up in St1 ± CLE, NSB (blue) and up in St5 (yellow). In the first two clusters, we further distinguish several subpatterns (see Table S1). (B) Overview of enriched GO-terms and KEGG pathways from patterns in A.



C

E7.5 PP>P

Included in E7.5 P (≥BMP4 levels)	Apoa1, Apoa2, Apob, Apom, Bmp4, Clic6, Cotl1, Ctsh, Dab2, Dlk1, Dusp4, Efna1, Ets1, Gal, Ifitm1, Igfbp3, Lin28, Lxn, Mixl1, Mogat2, Mrg1, Msx1, Nefm, Nfkbiz, Nkx1-2, Nme7, Pabpc1, Pmp22, Ppic, Rap2c, Rn18s, Sct, Slc13a4, Supt16h, Tcf7, Tmpo, Ttr, Vps35
Excluded from E7.5 P (<BMP4 levels)	2700060E02Rik, 2810026P18Rik, 2810043G13Rik, 6330419J24Rik, Abca3, Acbd3, Apoe, Atf4, Brd7, C1qbp, Car4, Cdc16, Comt, Crlf1, Ctsb, Cubn, Dkk3, Dnajb11, Dtnbp1, E430033B07Rik, Eef1e1, EG665369, EG666756, Egr1, Etsrp71, Fbxl14, Ghr, Gng12, Hadhb, Hist1h2ah, Hoxd8, Hoxd9, Hsp90aa1, Klhl6, Lamp2, LOC100039786, LOC100042514, LOC100046298, LOC100048436, LOC212399, LOC245892, LOC380665, LOC666053, Mrps6, Myo5b, Ndufs4, Npm3, Pcgf5, Plac1, Psma4, Rbp4, Rdh11, Reep6, Rhox5, S100a11, Sfrs11, Skil, Slc11a1, Slpi, Soat1, Spcs1, Spin2, Spink3, Tbx4, Tcof1, Tdh, Tubb6, Uap1l1, Uchl3

Figure S4. Comparison of LPMPs at E7.5 and E8.5.
(A) Schematic overview of dissected and compared regions at E7.5 headfold stage (HF). Since E7.5 P overlaps with PP gene expression, samples were compared as follows: E7.5 P versus A (red), E7.5 PP versus A (blue) and E7.5 PP versus P (shades of yellow). Since PP is included in the P sample and thus no gene will be ON in PP and OFF in P, we used the values of a known marker of the PP region, *Bmp4* (Lawson and Wilson, 2016), as the minimum cut-off for transcripts enriched in PP versus P. (B) Overlap of E7.5 sample comparisons with genes specific for E7.5 PP shown in (C) 'Included' genes, i.e. ≥BMP4 values, are most likely upregulated in PP, while some of the remainder may also be upregulated there, although less strongly.

Lawson, K. A. and Wilson, V. (2016). A Revised Staging of Mouse Development Before Organogenesis. In Kaufman's Atlas of Mouse Development Supplement (ed. B. R., B. J., D. D. & M.-K. G.), pp. 51–64: Elsevier.

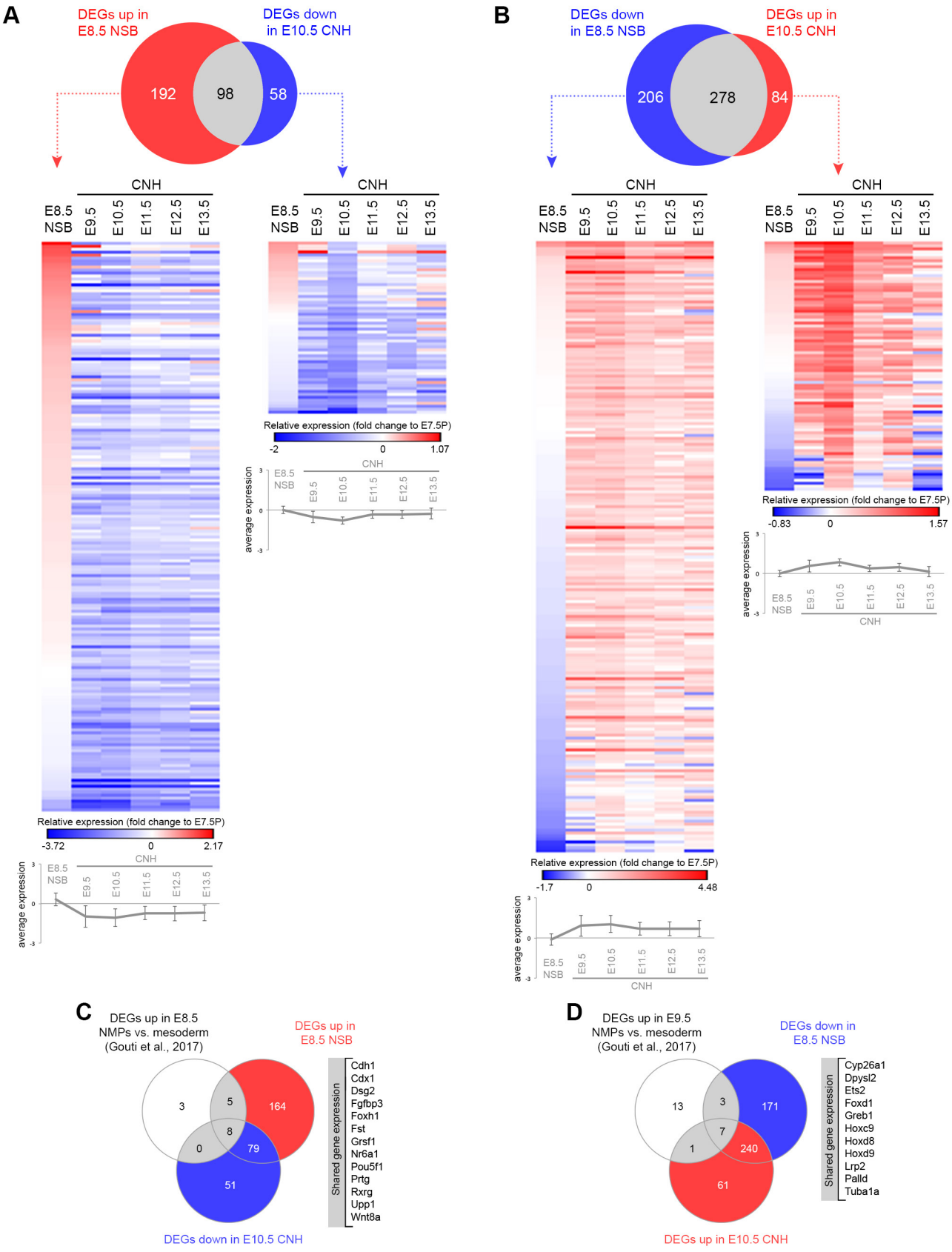
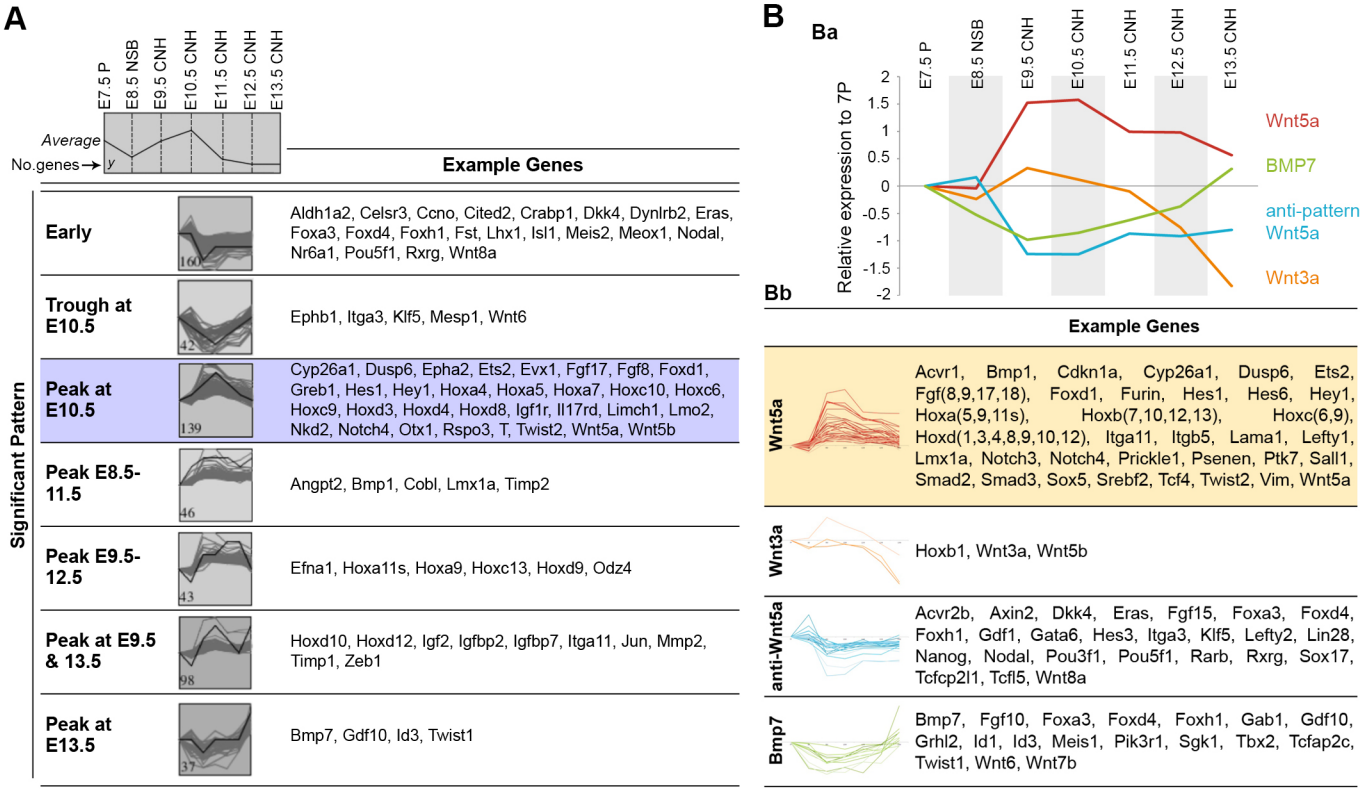
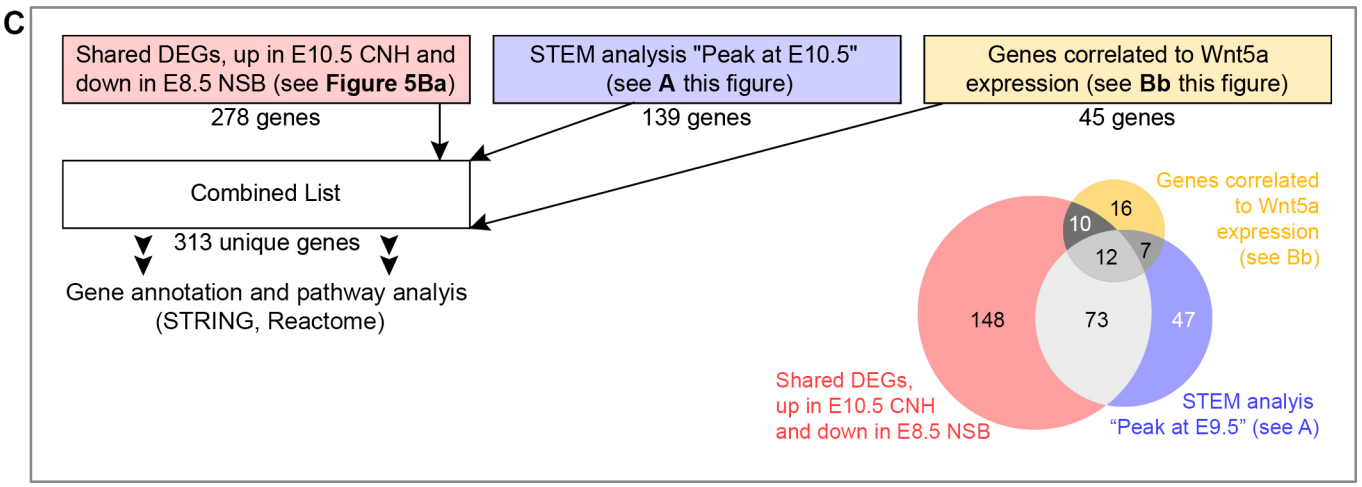


Figure S5. Transcriptional changes at mid-trunk formation.
(A) DEGs between the E8.5 NSB and the E10.5 CNH (intersections of these sets are shown in Fig. 5B). (B) Expression heatmaps of NMP-containing regions over time show a similar peak-decline pattern as those in Fig. 5Bb-c. The mean (± s.d.) is shown below each heatmap (see Table S5 for gene lists). (C-D) Up- or downregulated DEGs at E8.5 and E10.5, shown in A and B compared with enriched NMP genes (versus nascent mesoderm), obtained in a parallel single cell analysis (Gouti et al., 2017).





D

GO-term	Description	p-value	KEGG ID	Description	p-value	Genes
GO.0009893	positive regulation of metabolic process	3.23E-05	4151	PI3K-Akt signaling pathway	0.00197	Cdk6, Cdkn1a, Epha2, Fgf8/9/17/18, Ghr, Gng12, Igf1r, Insr, Itga11, Itgb5, Lama1, Vegfc
GO.0042127	regulation of cell proliferation	0.00953				
GO.0007155	cell adhesion	0.0146				
GO.0001568	blood vessel development	4.02E-06	4014	Ras signaling pathway	0.00616	Epha2, Ets2, Fgf8/9/17/18, Gng12, Igf1r, Insr, Rasa3, Vegfc
GO.0043408	regulation of MAPK cascade	0.00797				
GO.0030334	regulation of cell migration	0.0399	650	Butanoate metabolism	0.0115	Aacs, Abat, Acat2, Oxct1
GO.0045785	positive regulation of cell adhesion	0.00508	4015	Rap1 signaling pathway	0.0402	Epha2, Fgf8/9/17/18, Igf1r, Insr, Pard6g, Vegfc
GO.0050769	positive regulation of neurogenesis	0.0153				
GO.0048545	response to steroid hormone	0.0353	100	Steroid biosynthesis	0.0452	Cyp51, Hsd17b7, Sc4mol
GO.1901342	regulation of vasculature development	0.00256				
GO.0001558	regulation of cell growth	0.0415				
GO.0016055	Wnt signaling pathway	0.0326				
GO.0007219	Notch signaling pathway	0.00427				
GO.0036342	post-anal tail morphogenesis	0.00773				
GO.0030947	regulation of vascular endothelial growth factor receptor signaling pathway	0.0199				

Figure S6. Upregulated genes at mid-trunk formation.
(A) STEM analysis shows significant patterns in NMPs over time, relative to sample E7.5 P. Grey boxes show the number of genes in each pattern; the black line, their average (significant patterns defined as p≤0.05 by permutation test in STEM). (B) Correlated expression to selected genes: (Ba) chosen gene or pattern. (Bb) table of example genes that show a similar expression profile (Pearson's correlation factor ≥7). A full gene list can be found in Table S6. (C) A combined list of upregulated genes during mid-trunk formation, with the overlap showing in (D). Red, DEGs up in E10.5 CNH and down in E8.5 NSB. Blue, genes that peak at E10.5 (taken from the STEM analysis pattern in A). Yellow, genes with similar pattern to *Wnt5a* expression (taken from correlation analysis in Bb). (E) Overview of GO terms and KEGG signalling pathways in this combined list (see also Table S7).

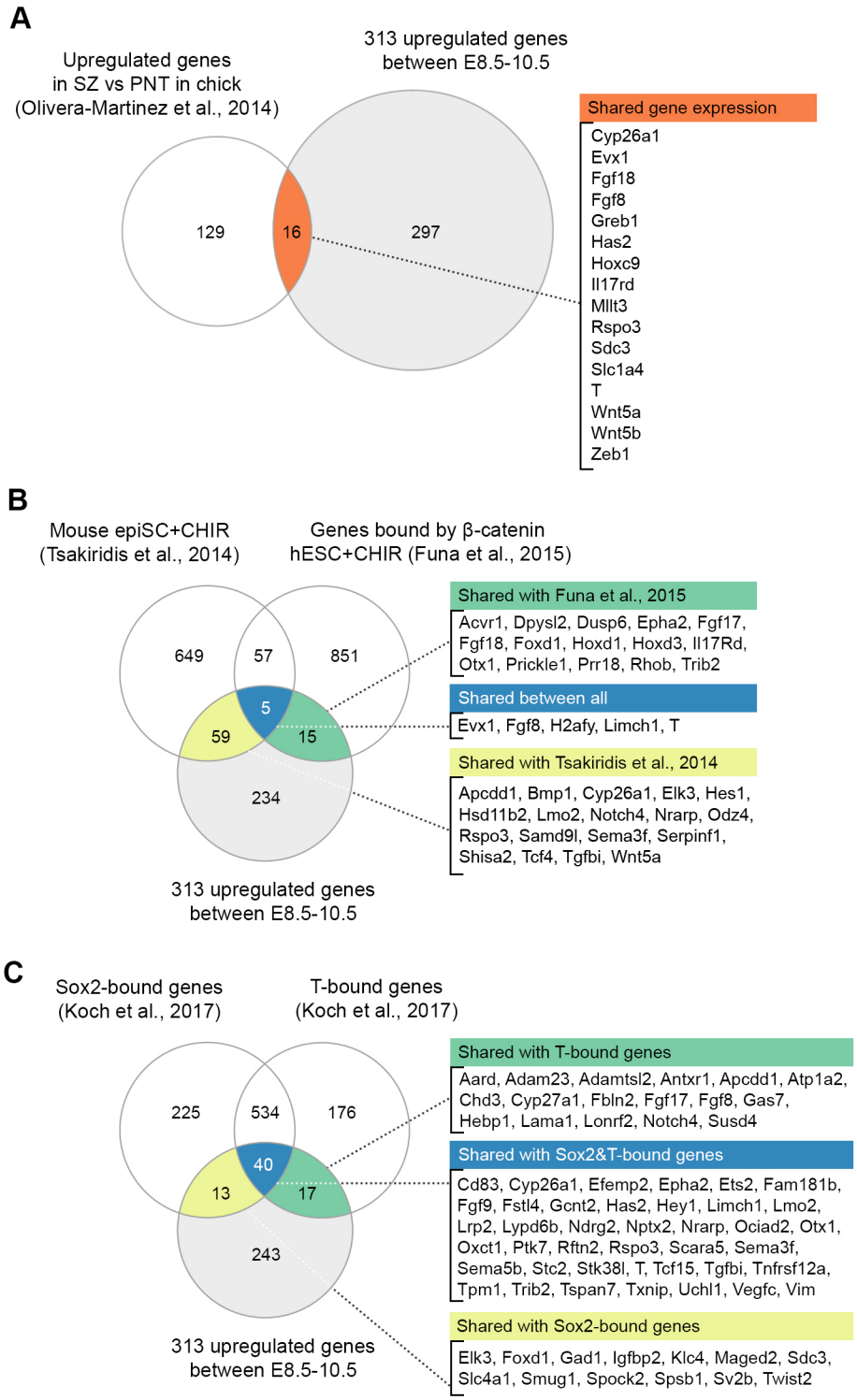


Figure S7. Comparison of upregulated genes at mid-trunk formation to other available datasets.
(A) Overlap between the 313 upregulated genes at mid-trunk formation with upregulated genes in the chick stem zone (versus the pre-neural tube) (Olivera-Martinez et al., 2014). (B) Overlap of 313-gene list with β-catenin bound genes in CHIR-treated human ESC (Funa et al., 2015) or mouse EpiSCs (Tsakiridis et al., 2014). (C) Overlap of 313-gene list with Sox2-, T- or Sox2&T- bound genes (Koch et al., 2017) (see also Table S7).

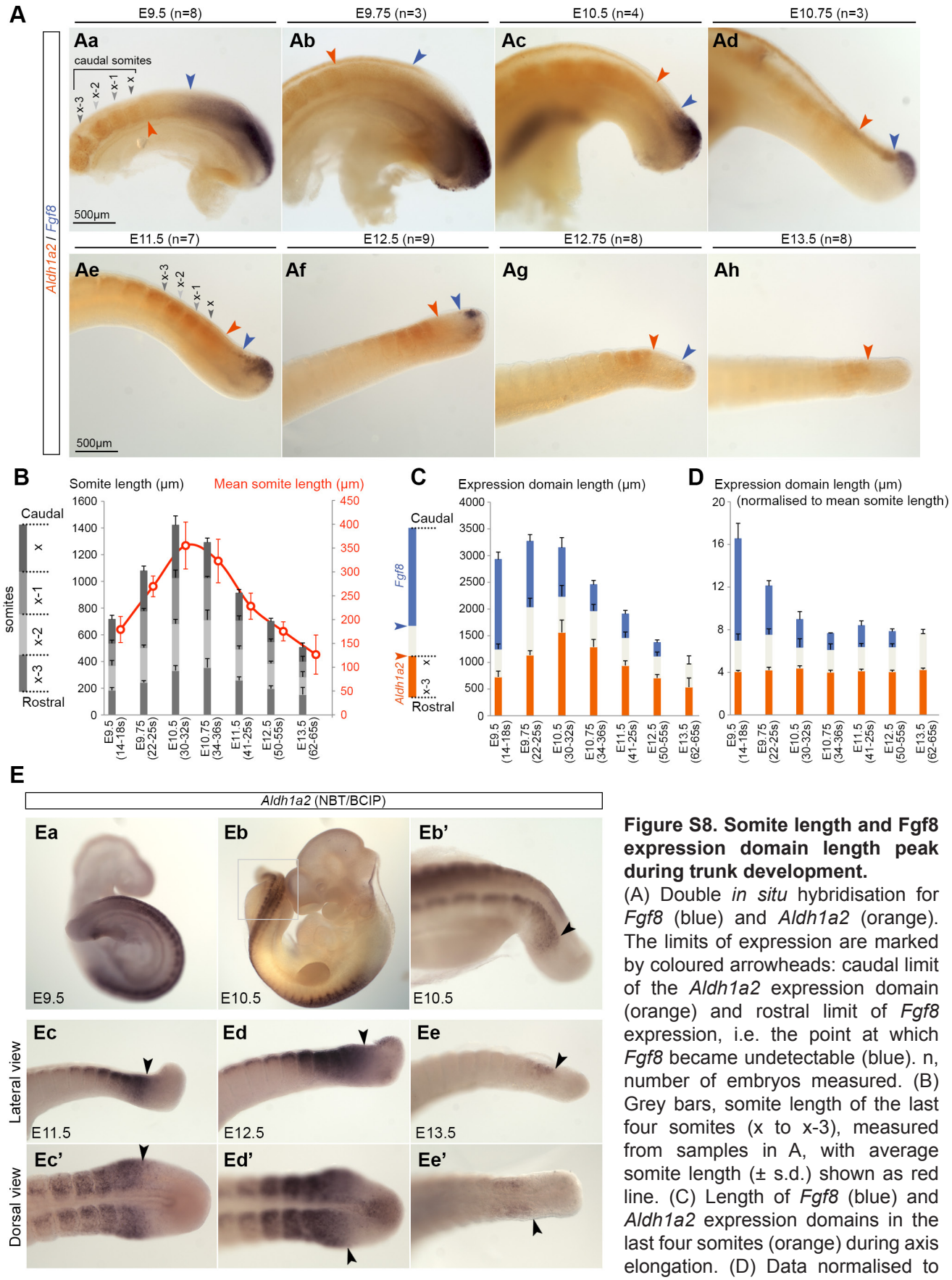


Figure S8. Somite length and Fgf8 expression domain length peak during trunk development.
(A) Double *in situ* hybridisation for *Fgf8* (blue) and *Aldh1a2* (orange). The limits of expression are marked by coloured arrowheads: caudal limit of the *Aldh1a2* expression domain (orange) and rostral limit of *Fgf8* expression, i.e. the point at which *Fgf8* became undetectable (blue). n, number of embryos measured. (B) Grey bars, somite length of the last four somites (x to x-3), measured from samples in A, with average somite length (± s.d.) shown as red line. (C) Length of *Fgf8* (blue) and *Aldh1a2* expression domains in the last four somites (orange) during axis elongation. (D) Data normalised to the mean somite length. (E) Single *in situ* hybridisation controls for *Aldh1a2* transcripts at a series of developmental stages, using NBT/BCIP (5-bromo-4-chloro-3-indolyl-phosphate/nitro blue tetrazolium) instead of Fast Red as development substrate (used in A).

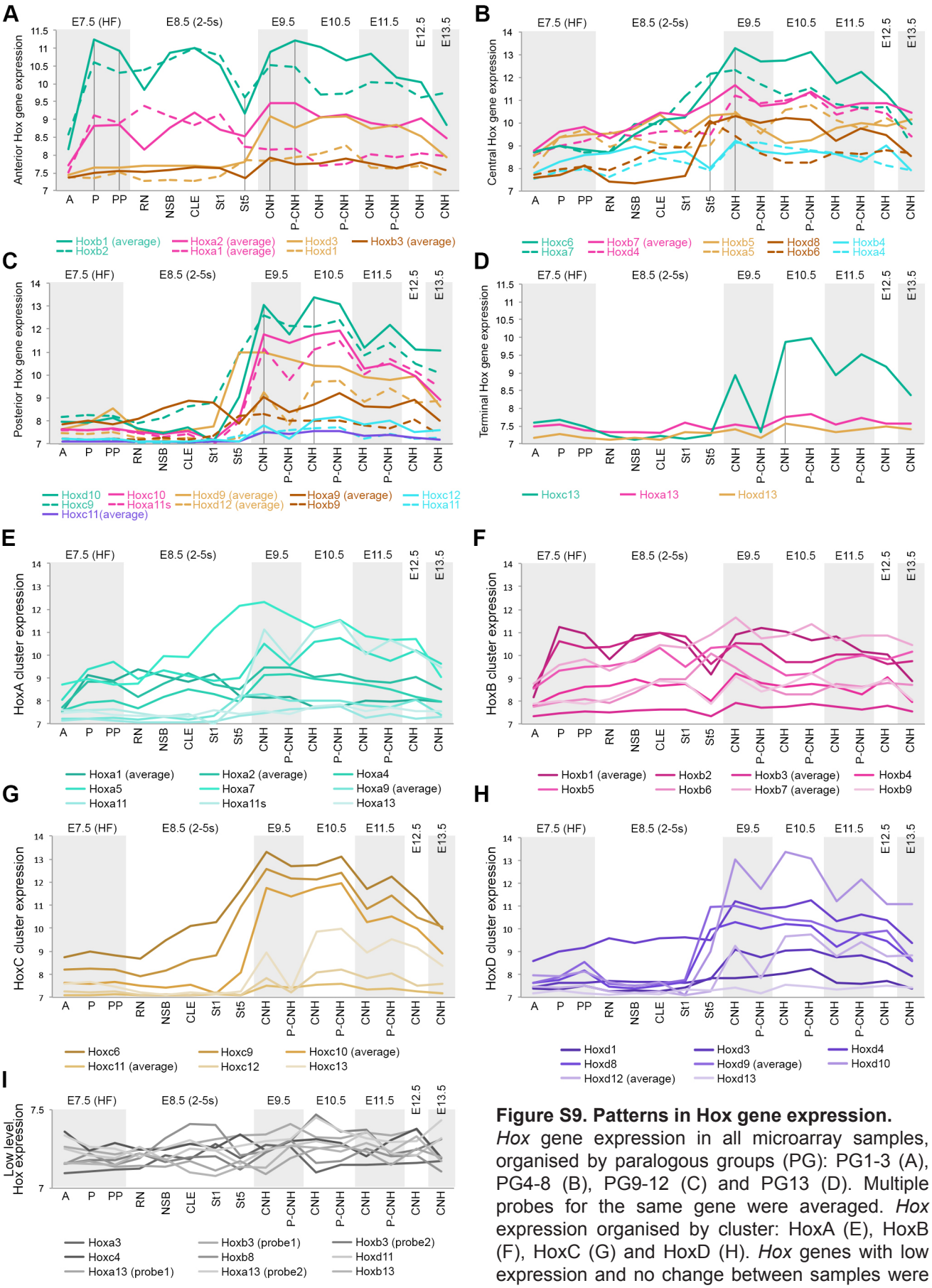
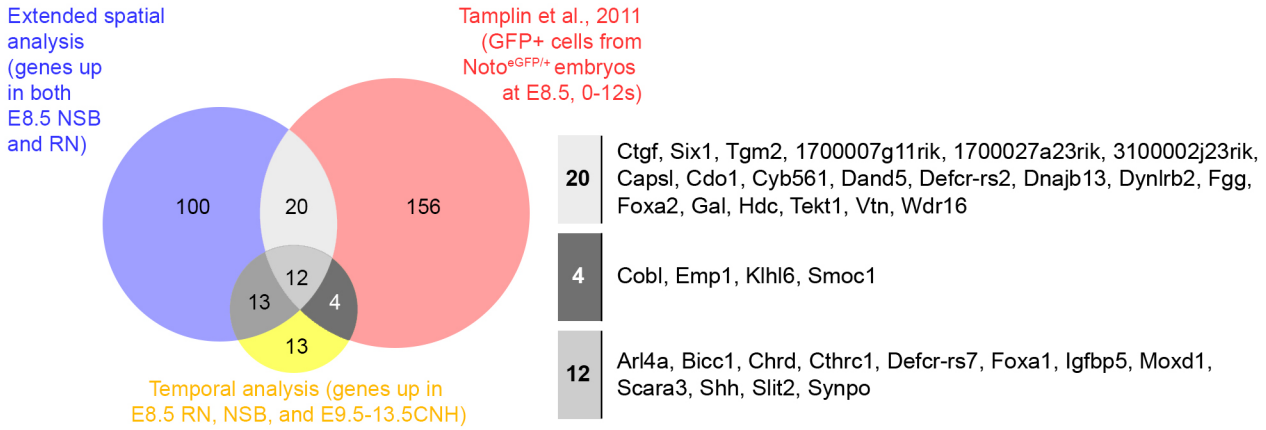
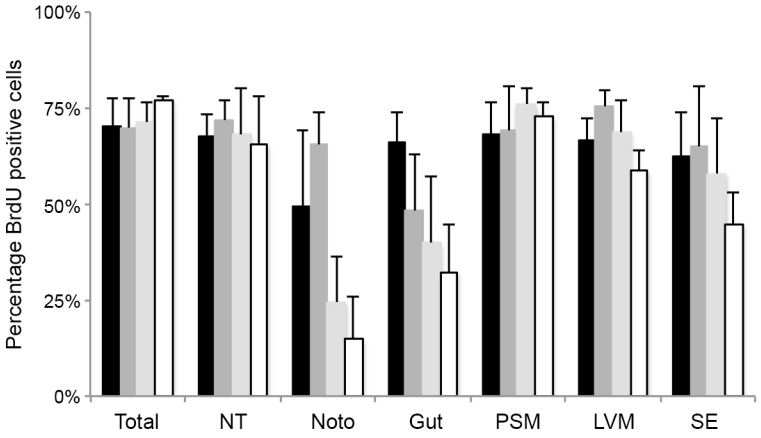


Figure S9. Patterns in Hox gene expression. *Hox* gene expression in all microarray samples, organised by paralogous groups (PG): PG1-3 (A), PG4-8 (B), PG9-12 (C) and PG13 (D). Multiple probes for the same gene were averaged. *Hox* expression organised by cluster: HoxA (E), HoxB (F), HoxC (G) and HoxD (H). *Hox* genes with low expression and no change between samples were excluded from the analysis (I).

A



B



	No. tails	No. TB sections	No. cells counted
E10.5	5	12	16.831
E11.5	6	14	20.509
E12.5	6	18	31.717
E13.5	4	5	17.102

Percentage BrdU positive cells in tail bud tissues (mean ± s.d.)

	Total	NT	Noto	Gut	PSM	LVM	SE
E10.5	70.4 ± 7.3	67.7 ± 5.5	49.6 ± 19.7	66.2 ± 7.8	67.9 ± 8.4	66.4 ± 5.8	62.7 ± 11.3
E11.5	69.9 ± 7.5	71.9 ± 5.4	65.5 ± 8.6	48.2 ± 14.9	69.2 ± 11.2	75.4 ± 4.0	65.0 ± 15.8
E12.5	71.4 ± 5.3	68.2 ± 11.8	24.5 ± 11.6	39.9 ± 17.5	76.1 ± 4.1	68.8 ± 8.1	58.0 ± 14.3
E13.5	76.9 ± 1.3	65.3 ± 12.6	15.1 ± 10.9	32.2 ± 12.5	73.1 ± 3.6	58.7 ± 5.3	44.5 ± 8.4

Figure S10. Experimental controls and comparison of NotoP markers.
(A) Comparison of E8.5 spatial analysis (DEGs shared in both RN and NSB), E8.5-13.5 temporal analysis (DEGs shared across E8.5 RN, NSB and E9.5-13.5 CNH samples) gene expression in sorted GFP$^+$ cells from E8.5 Noto$^{eGFP/+}$ embryos (0-12s) (Tamplin et al., 2011) (Table S8). (B) Percentage of BrdU$^+$ cells counted from immunostained tail sections after 4h *ex vivo* culture in the presence of BrdU. No. tails, no. tail bud (TB) sections and total no. cells counted as shown. Data is organised by tissue type: NT, neural tube, Noto, notochord; PSM, presomitic mesoderm; LVM, lateral and ventral mesoderm; SE, surface ectoderm.

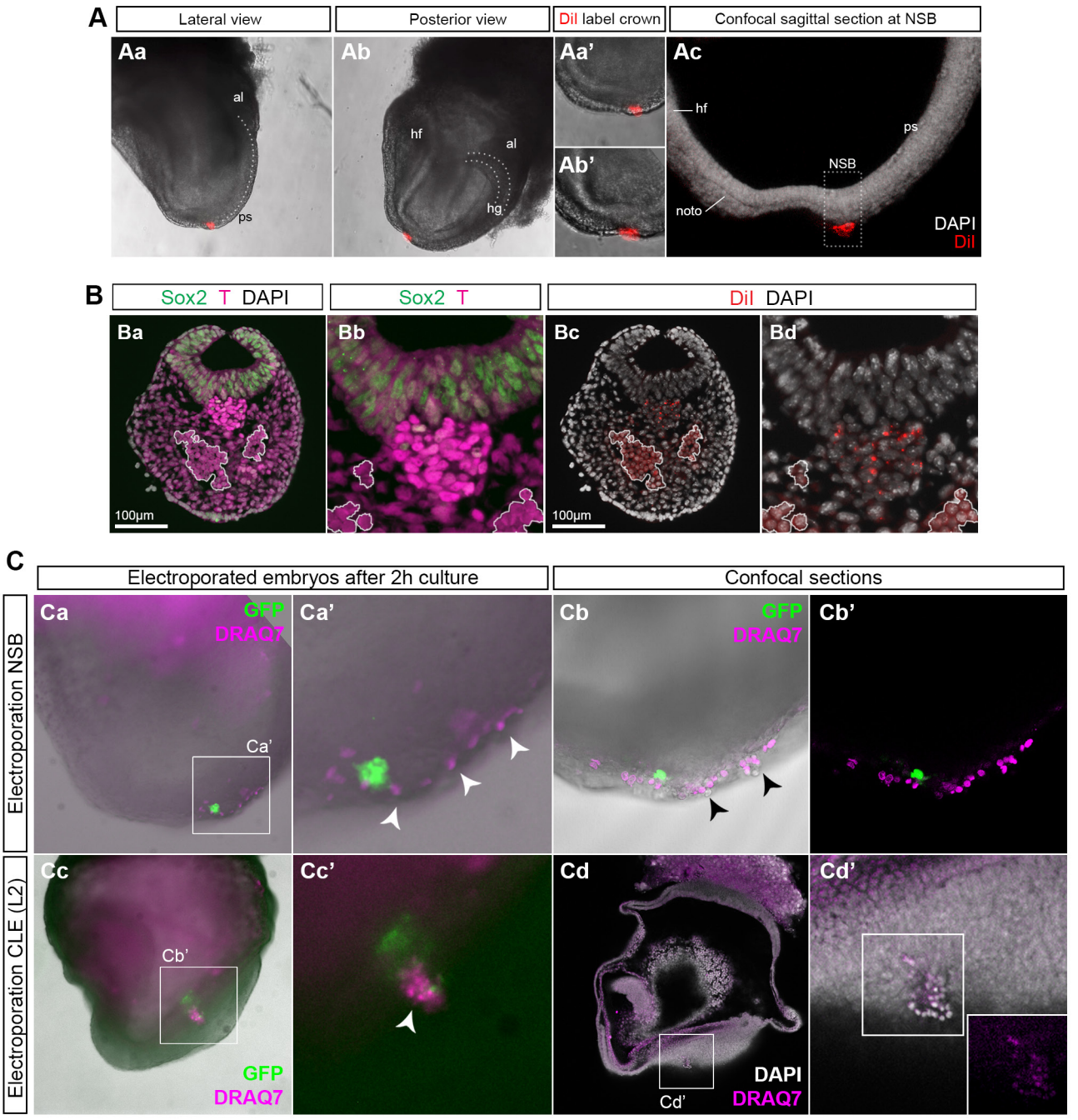
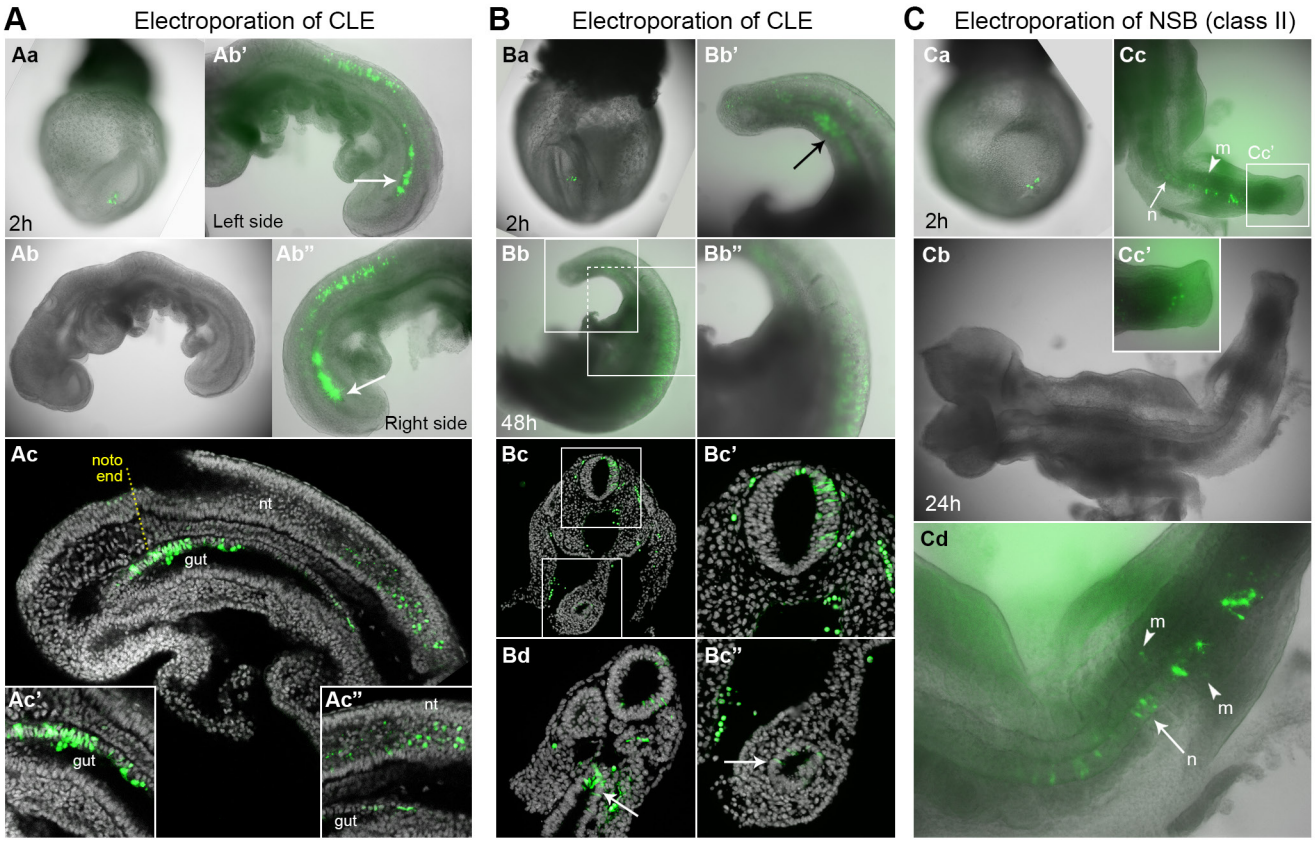


Figure S11. DiI labelling and electroporation controls.
(A) DiI labelling of the crown cells (red) in an E8.5 embryo: wholemount lateral (Aa), posterior view (Ab) validate the accuracy of initial labelling ($n_{embryos}$=3). (B) Sox2/T immunostained section of embryo labelled with DiI in the ventral node layer after 48h culture ($n_{embryos}$=2). White lines surround autofluorescent blood cells. (C) Cell death in NSB (Ca-Cb) or CLE (L2) (Cc-Cd) in electroporated embryos. After two hours *ex vivo* culture, GFP^+ cells (green) were observed at the electroporation site. Cell death, shown by DRAQ7 dye uptake (magenta), is located primarily at the ventral side of the embryo (arrowheads).



D

Electroporation site	No. embryos labelled in		Total no. electroporated
	Axis	Tail bud	
PS	2	2	2
CLE	10	2	10
NSB	14	6	20

Figure S12. Fate of electroporated cells.
Cell fate was examined in CLE-electroporated embryos after 24 (A) or 48h (B) *ex vivo* culture. In addition to the expected fate, the hindgut was labelled in all examined embryos (arrows; $n_{embryos}$=4), suggesting that also hindgut progenitors were electroporated. nt, neural tube. (C) Fate in a class II NSB-electroporated embryo. After 24h, GFP^+ cells were found in the neural tube (n) and paraxial mesoderm (m) of this embryo. (D) Contribution to the axial tissues and tail bud in electroporated samples.

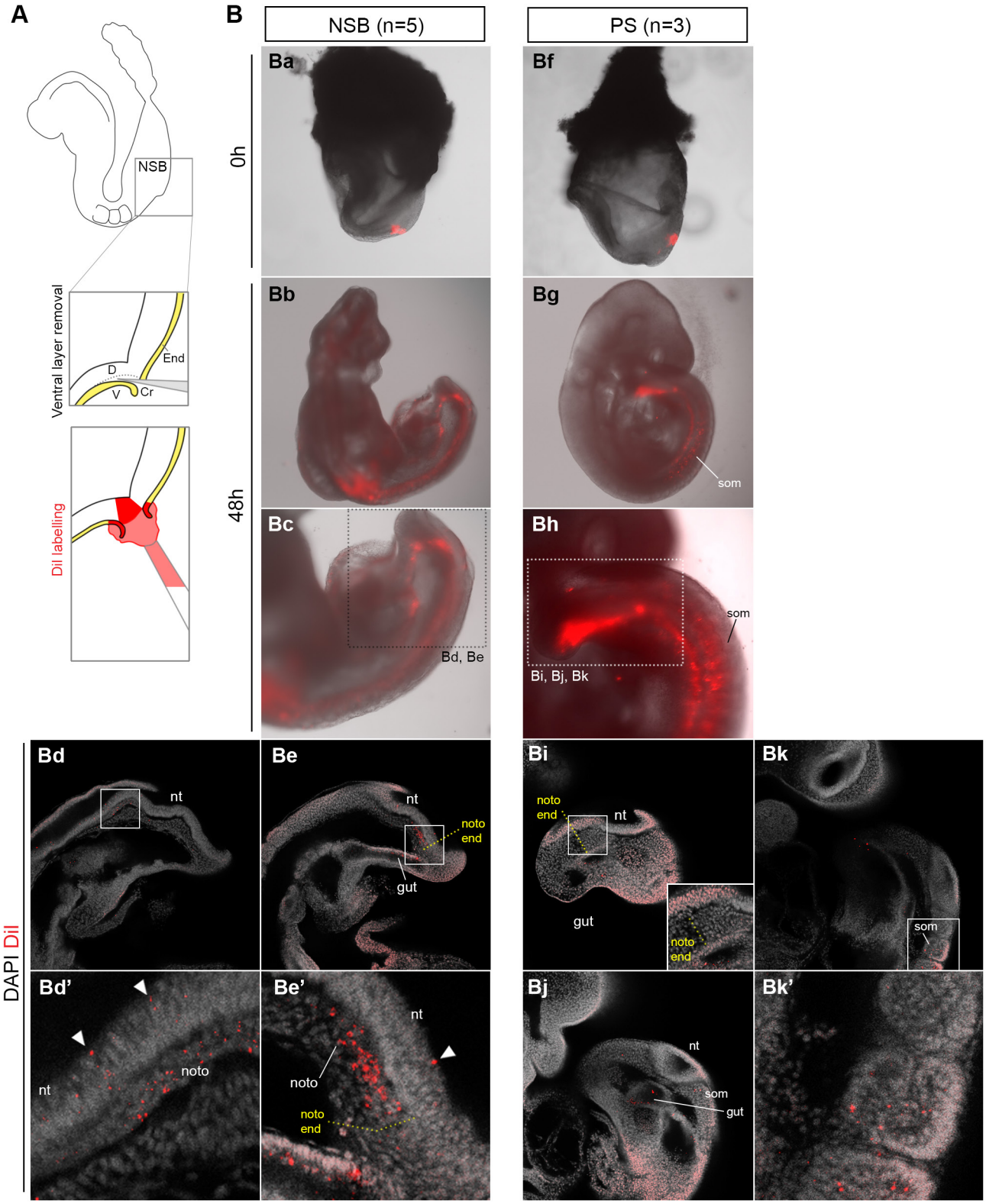


Figure S13. Effect of ventral cell layer removal at the NSB or St3.
(A) Experimental procedure schema illustrating the removal of the endoderm layer. A sharp glass needle was inserted posteriorly and pushed anteriorly to peel away the ventral endodermal cells. This ventral layer was further trimmed, after which DiI was poured on the damaged site, labelling the cell layer above as well as the endodermal cells. (B) Fate and phenotype after ventral layer removal in the NSB (Ba-Be) or St3 (Bf-Bk) after 48h *ex vivo* culture. (Bb-Bc) Removal of the ventral node layer resulted in abnormal growth (similar to class II phenotypes in electroporation, $n_{embryos}$=5, Fig. 8C). DiI-labelled cells were found in the notochord and neural tube (Bd-Be). Control embryos grew normally (n=3) and DiI was found in somite and gut tissues (Bg-Bh). Red dots, DiI (pink, background due to high exposure). nt, neural tube; noto, notochord; som, somites; white arrowheads, DiI label in ventral neural tube.

SUPPLEMENTARY TABLES

Table S1. Full list of DEGs in the primitive streak and St5 (supplementary to Figure 2).
(Tab 1) Unique DEGs for St5, not shown in Figure 2A. (Tab 2) List of unique DEGs for each region at E8.5. Columns allow for pairwise comparison between two regions (1, ≥1.5 fold upregulated; -1, ≥1.5 fold downregulated; 0, no significant change).

Click here to download Table S1

Table S2. Hierarchical clustering of DEGs ≥1.5 fold change across the primitive streak (supplementary to Figure S3).
Tab 1 shows all DEGs ≥1.5 fold changed across the E8.5 sample set that were used in hierarchical clustering of Figure S3 and analysed for GO-terms/KEGG pathways. Three major clusters were identified and several sub-clusters could be distinguished. Subsequent tabs show all GO-terms and KEGG pathways for each pattern obtained from the STRING online database (accessed 10th April 2017).

Click here to download Table S2

Table S3. STEM analysis across different sample sets (supplementary to Figure 4 and S6).
STEM analysis shows significant patterns in two datasets: (1) in the E7.5 embryo (A, P and PP samples) and (2) in comparable NMP-containing regions (E8.5 NSB to E9.5-E13.5 CNH, normalised to E7.5 P). Tables shows significant patterns, and the genes and enriched GO-terms within each pattern ($p \leq 0.05$, defined by permutation test in STEM).

Click here to download Table S3

Table S4. DEGs unique to and shared between LPMPs at E7.5/8.5 (supplementary to Figure 4).
Full list of DEGs not shown in Figure 4C.

Click here to download Table S4

Table S5. Change in NMP gene expression during axis elongation (supplementary to Figure 5 and Figure S5).
Lists of DEGs not shown in Figure 5B (Tab 1) and 5C (Tab 2). (Tab 3) Shared gene expression between DEGs uniquely changed at E8.5, E10.5 and DEGs upregulated in E8.5 or E9.5 single NMPs (compared to mesoderm) (Gouti et al., 2017).

Click here to download Table S5

Table S6. Correlated gene expression to selected genes or expression pattern (supplementary to Figure S6).
Genes that correlate to the selected gene or profile in NMP-containing regions during axis elongation (E8.5 NSB, E9.5-E13.5 CNH). Selected genes are *Bmp1, Bmp7, Wnt3a, Wnt5a*, and anti-correlated pattern to *Wnt5a*. Right column, genes with a similar pattern to the selected gene (Pearson's correlation factor ≥0.7). Left, curated genes and their temporal expression (≥1.5 fold change, normalised to 7P expression).

Click here to download Table S6

Table S7. Genes upregulated between E8.5 and E10.5 NMPs: construction of list, its annotation and comparison to other datasets (supplementary to Figure S6 and Figure S7).
(Tab 1) Genes upregulated during mid-trunk formation in NMPs, with the overlap between different analyses shown. Red, DEGs that were upregulated in the E10.5 CNH, and down at E8.5 NSB. Blue, genes that peak at E9.5 (from STEM analysis). Yellow, genes with similar temporal expression pattern to *Wnt5a*. (Tab 2) Manual annotation of the combined list shown in Tab 1. STRING annotation (Tab 3) and GO-terms associated with the combined list (Tab 4). Tab 5: Comparison with Olivera-Martinez et al., 2014. Tab 6: Comparison with Funa et al., 2015 and Tsakiridis et al., 2014. Tab 7: Comparison with Koch et al., 2017.

Click here to download Table S7

Table S8. Markers of notochord and node during axis elongation (supplementary to Figure S10).
(Tab 1) Common gene expression in E8.5 RN, NSB and E9.5 to E13.5 CNH versus all other samples (dataset normalised to E7.5P). Data is separated into known and potential novel markers for the node/notochord. Red: averaged expression from different probes on the microarray. (Tab 2) Comparison between Tamplin et al., 2011. and samples containing NotoPs: enriched in E8.5 RN and NSB versus E8.5 samples ('spatial analysis'), and in E8.5 RN, NSB and E9.5-13.5 CNH ('temporal analysis', see Tab 1).

Click here to download Table S8

	Embryo	Mouse line	Plasmid	GFP location at 2h	Culture period	Phenotype class
CLE	1	MF1	pCAG-GFP	L1	24h	I
	2	MF1	pCAG-Cre:GFP	L1	24h	N
	3	MF1	pCAG-GFP	L1	24h	N
	4	sGFP	pCAG-Cre:GFP	L1	48h	N
	5	sGFP	pCAG-Cre:GFP	L1	48h	N
	6	sGFP	pCAG-Cre:GFP	LL1	48h	N
	7	sGFP	pCAG-Cre:GFP	L1	48h	N
	8	sGFP	pCAG-Cre:GFP	L1	48h	N
	9	sGFP	pCAG-Cre:GFP	L1	48h	N
	10	sGFP	pCAG-Cre:GFP	L1	48h	N
PS	11	sGFP	pCAG-Cre:GFP	St1	48h	N
	12	MF1	pCAG-GFP	St3	24h	N
NSB	13	sGFPxMF1	pCAG-Cre:GFP	RN+NSB	24h	I
	14	sGFPxMF1	pCAG-Cre:GFP	NSB	24h	I
	15	MF1	pCAG-GFP	NSB	24h	II
	16	MF1	pCAG-GFP	NSB	24h	II
	17	MF1	pCAG-GFP	NSB	48h	II
	18	MF1	pCAG-GFP	NSB	24h	II
	19	MF1	pCAG-GFP	NSB	24h	II
	20	MF1	pCAG-GFP	NSB	24h	II
	21	MF1	pCAG-GFP	NSB	24h	II
	22	MF1	pCAG-GFP	NSB	24h	I
	23	MF1	pCAG-GFP	NSB	24h	I
	24	MF1	pCAG-GFP	NSB	24h	I
	25	MF1	pCAG-GFP	NSB	48h	N
	26	MF1	pCAG-GFP	NSB	48h	II
	27	MF1	pCAG-GFP	NSB	48h	II
	28	MF1	pCAG-GFP	RN+NSB	24h	I
	29	MF1	pCAG-GFP	NSB	24h	II
	30	MF1	pCAG-GFP	NSB	24h	II
	31	MF1	pCAG-GFP	NSB	24h	II
	32	MF1	pCAG-GFP	NSB	24h	II
Control	33	MF1	pCAG-GFP	EP control	24h	N
	34	MF1	pCAG-GFP	EP control	24h	N
	35	MF1	pCAG-GFP	EP control	24h	N
	36	MF1	pCAG-GFP	EP control	24h	II

Table S9. Overview of electroporated embryos.
Overview of electroporated embryos of CLE, primitive streak or NSB-targeted regions in E8.5 (2-5s) embryos. Table shows embryo number, corresponding genotype, plasmid and culture period. After 24/48h *ex vivo* development all embryos were assessed on their developmental features and assigned a phenotype class: N, normal; I, class I; II, class II.

Real-time qPCR primer list		
Gene		Primer sequence
Bmp4	Fwd	caacccaattatgggctggc
	Rev	ccacaatccaatcattccagc
Evx1	Fwd	gtttcaagaccgcggagat
	Rev	tgacgcttgtccttcatgc
Fgf4	Fwd	ccggttcttcgtggctatga
	Rev	cttactgagggccatgaacatacc
Fgf8	Fwd	atggcagaagacggagacc
	Rev	ttgttcatgcagatgtagagacc
Foxa2	Fwd	catccgactggagcagcta
	Rev	gcgcccacataggatgac
Mixl1	Fwd	agttgctggagctcgtcttc
	Rev	agggcaatggaggaaaactc
Sox2	Fwd	ggcggcaaccagaagaacag
	Rev	gcttggcctcgtcgatgaac
T(Bra)	Fwd	ccaaggacagagagacggct
	Rev	agtaggcatgttccaagggc
TBP	Fwd	ggggagctgtgatgtgaagt
	Rev	ccaggaaataattctggctca
Wnt3a	Fwd	aatggtctctcgggagtttg
	Rev	cttgaggtgcatgtgactgg

Table S10. Real-time qPCR primer list. Overview of the primers used in Fig. S1B.