

A

Age	Stage	Region	Major components							Minor components					replicates	no. regions per replicate				Replicate Naming			
			NotoP	NMP	LPMP	Nascent presomitic mesoderm	Nascent lateral mesoderm	Anterior neuroectoderm	Notochord	Endoderm	Ventral neuroectoderm	Surface ectoderm	Presomitic mesoderm	Somite		Blood vessel	9	13	25	7A1	7A2	7A3	
E7.5	HF	A							✓	✓	✓						3	9	13	25	7A1	7A2	7A3
		P	✓	(✓)	✓	✓	✓									2	9	13	-	7P1	7P2		
		PP			✓											2	12	15	-	7PP10	7PP15		
E8.5	2-5s	RN	✓								✓	✓			✓	2	45	45	-	RN-1	RN-2		
		NSB	✓	✓												3	45	45	57	Bord-1	Bord-2	8B3	
		CLE		✓								✓				3	12	43	36	8L1	8L2	8L3	
		St1					✓					✓				2	45	45	-	St1-1	St1-2		
		St5				✓		✓				✓	✓			2	45	45	-	St5-1	St5-2		
9.5	22-25s	CNH	✓	✓											2	6	10	-	9M1	9M2			
		P-CNH					✓							✓	2	6	11	-	9P1	9P2			
E10.5	32-35s	CNH	✓	✓											2	17	17	-	MCNH-1	MCNH-2			
		P-CNH					✓							✓	2	17	17	-	PCNH-1	PCNH-2			
11.5	40-45s	CNH	✓	✓											3	38	29	52	11M1	11M2	11M3		
		P-CNH					✓							✓	2	33	65	-	11P1	11P2			
12.5	50-55s	CNH	✓	✓											2	15	33	-	12M1	12M2			
13.5	60-63s	CNH	(✓)	(✓)											2	48	74	-	13M1	13M2			

Microarray 1

Microarray 2

Microarray 3

Microarray 4

Microarray 5

Microarray 6

Batch 1
Batch 2
Batch 3

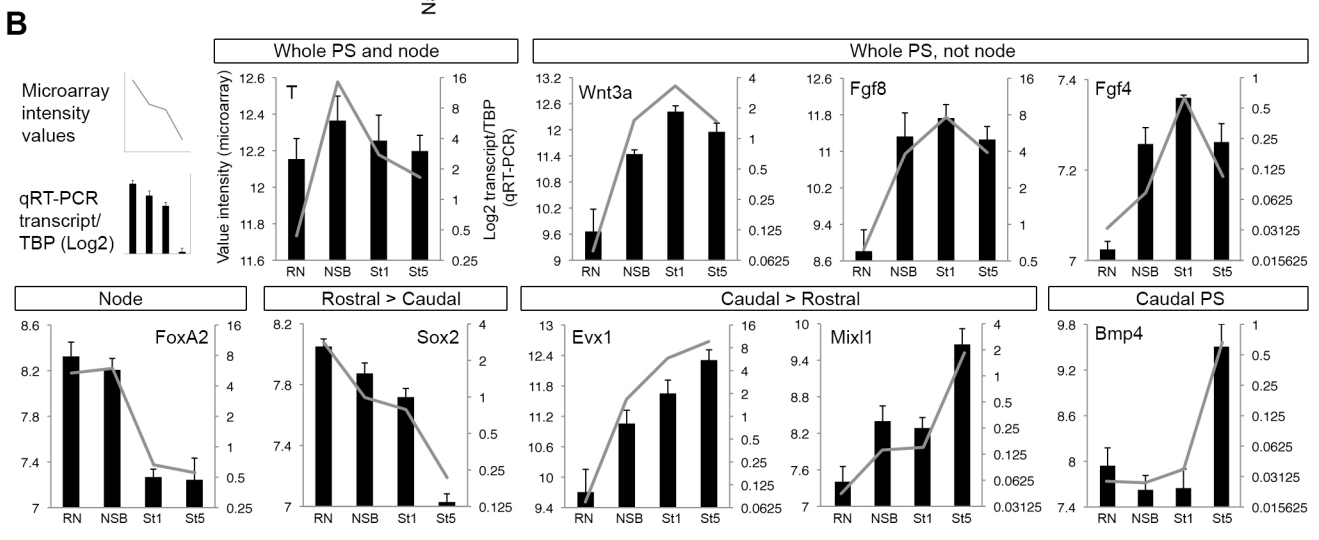


Figure S1. Microarray sample specifics and qRT-PCR validation.
 (A) Table of dissected samples, showing major and minor cell components included in the samples. A tick indicates presence, brackets indicate uncertainty about the proportion of progenitors in the sample. Replicates were named as shown and 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 \pm s.e.m. plotted as log2 values (black bars) and compared to the normalised intensity values of the microarray (grey line).

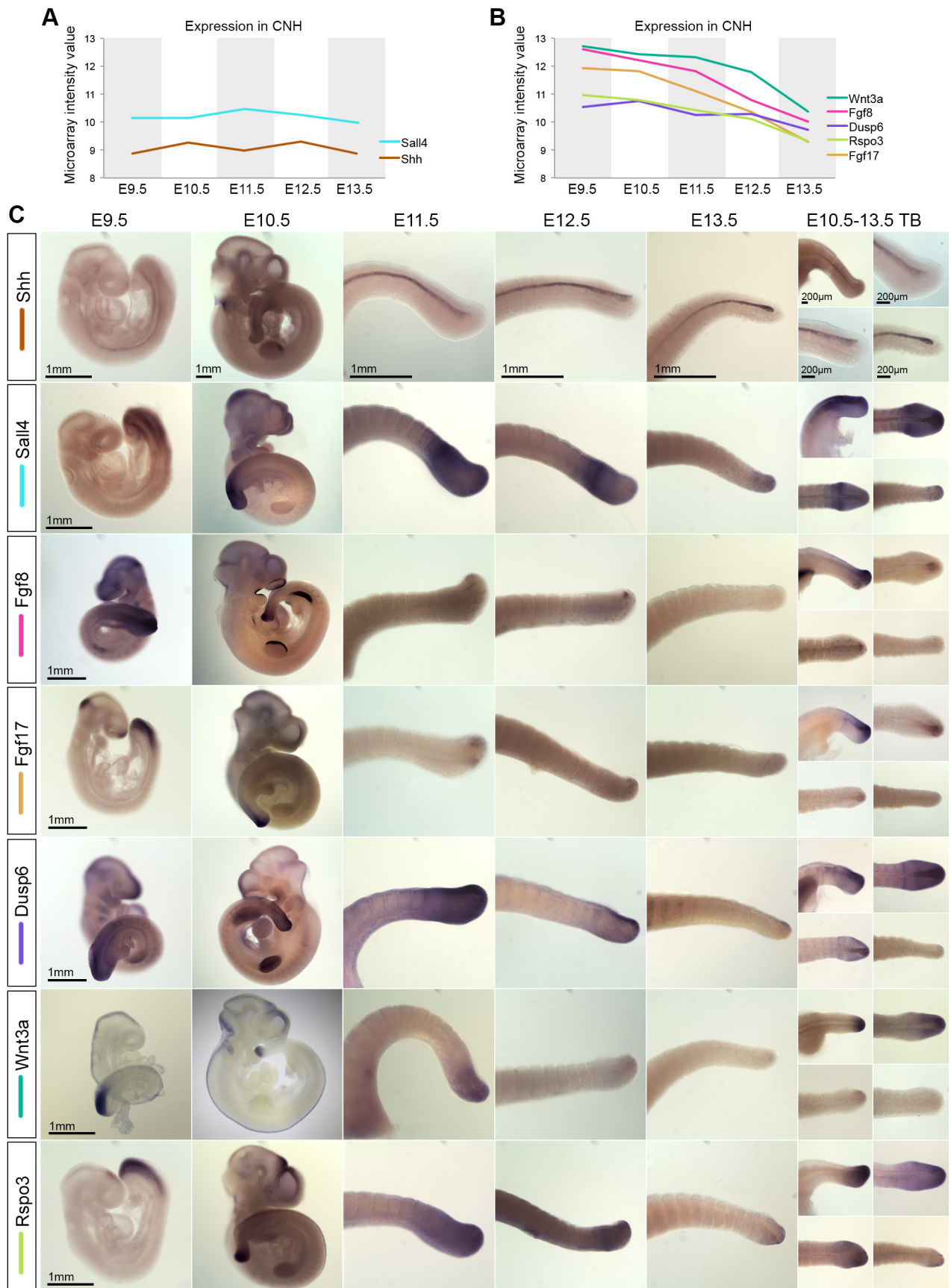


Figure S2. Validation of microarray by in situ hybridisation.

(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.



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 \pm other samples (pink), up in St1 \pm 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.

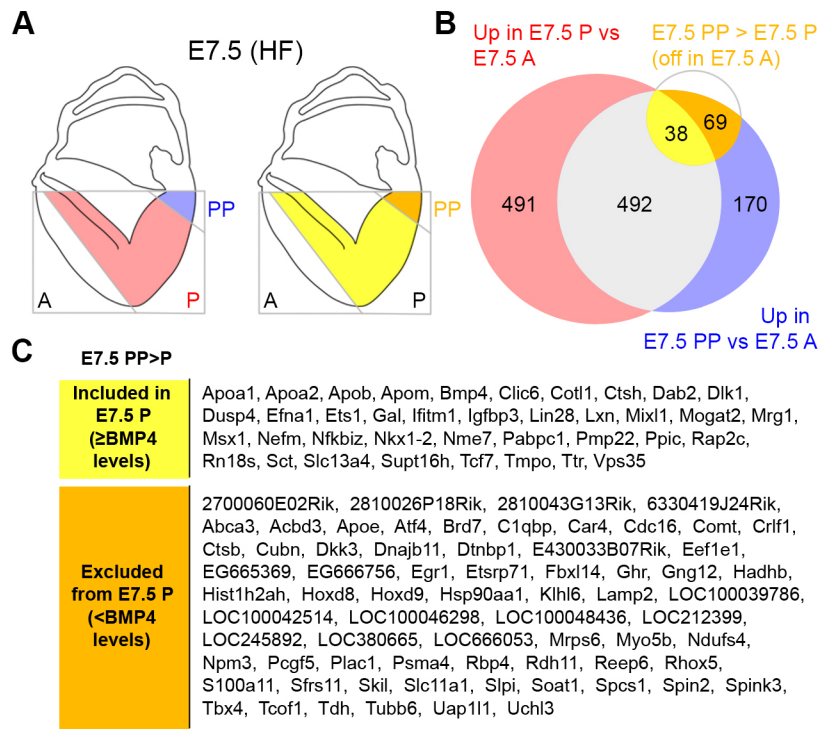


Figure S4. Comparison of LPMPs at E7.5 and E8.5.

(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.

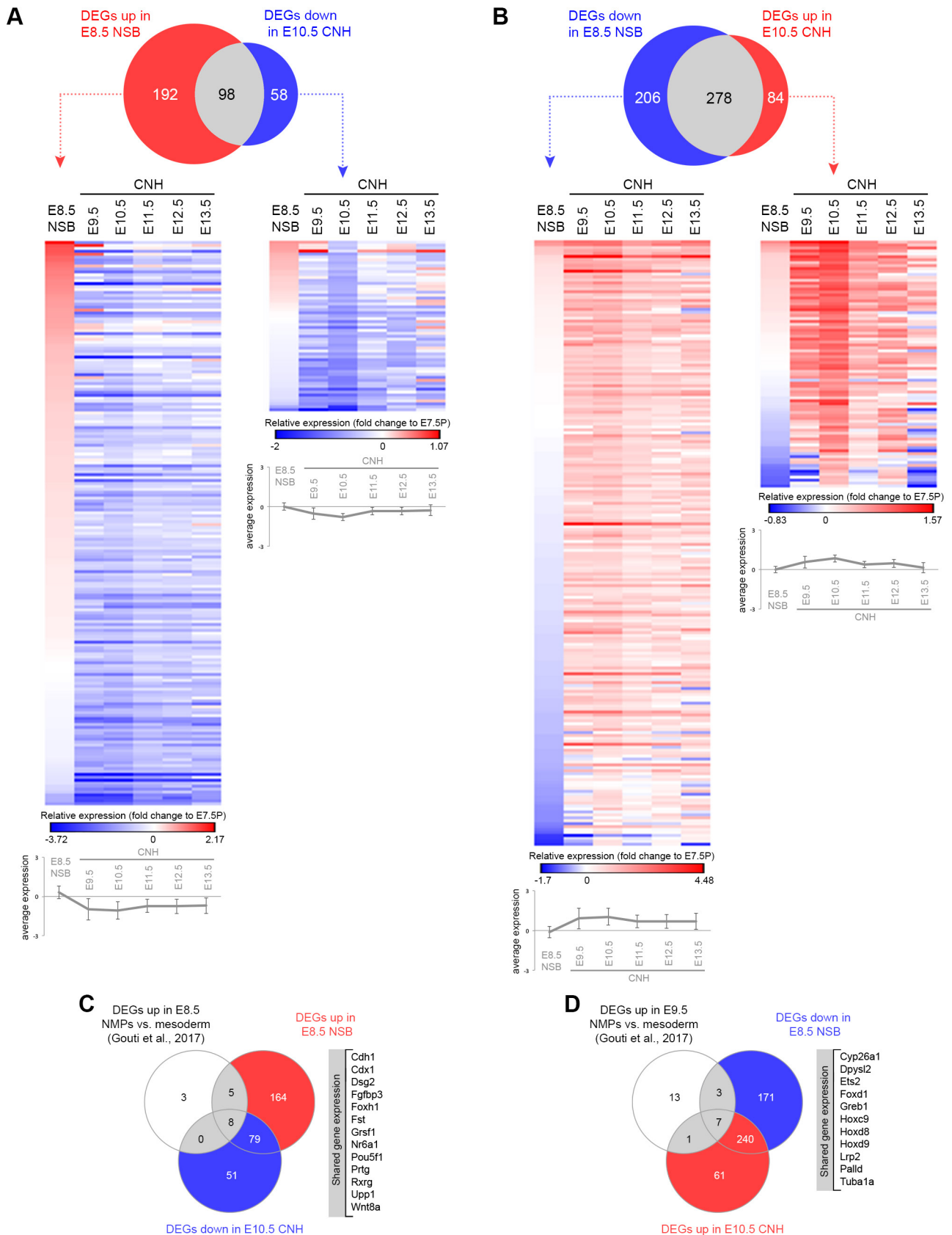


Figure S5. Transcriptional changes at mid-trunk formation.

(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 (\pm 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).

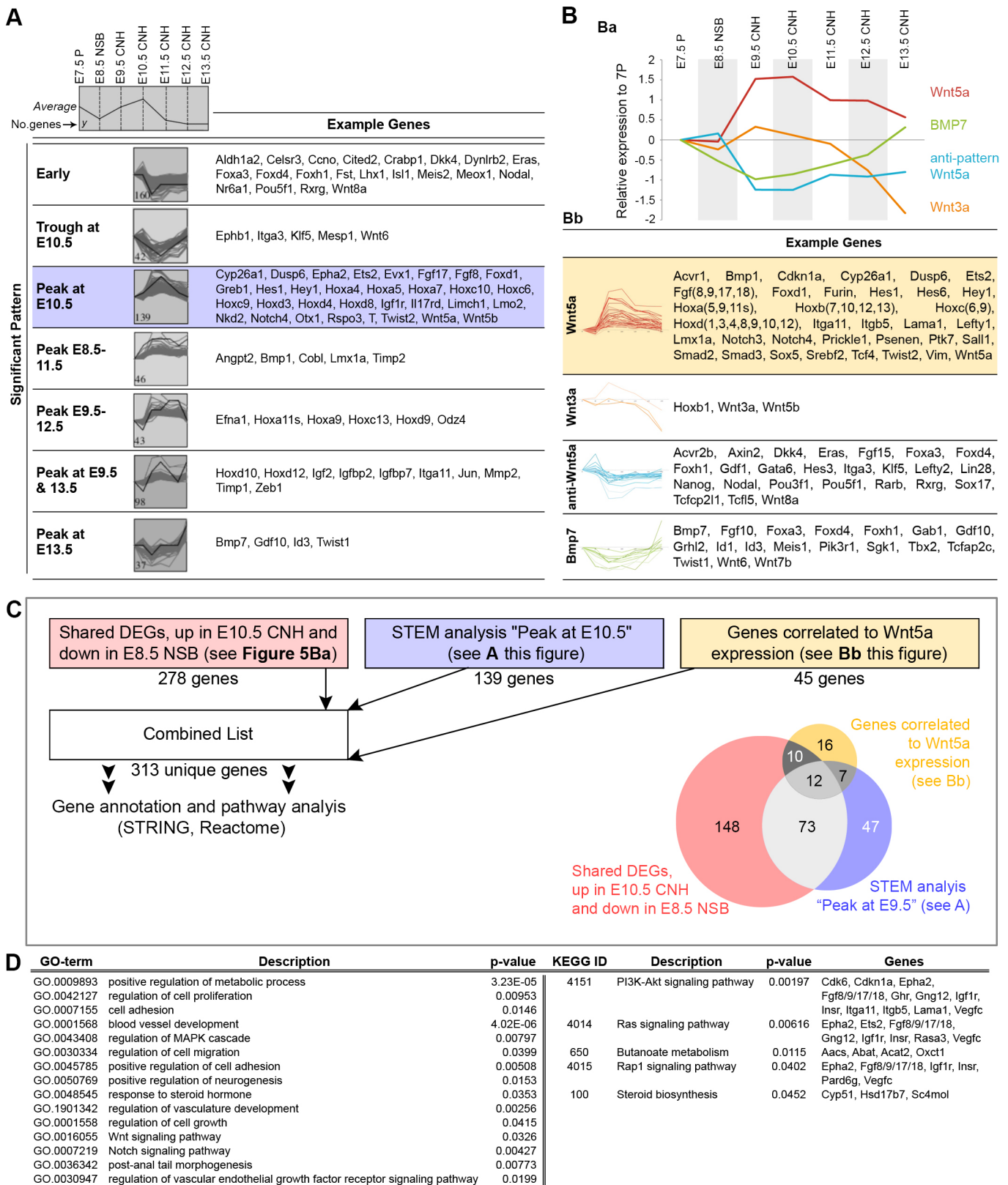


Figure S6. Upregulated genes at mid-trunk formation.

(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 $p \leq 0.05$ 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).

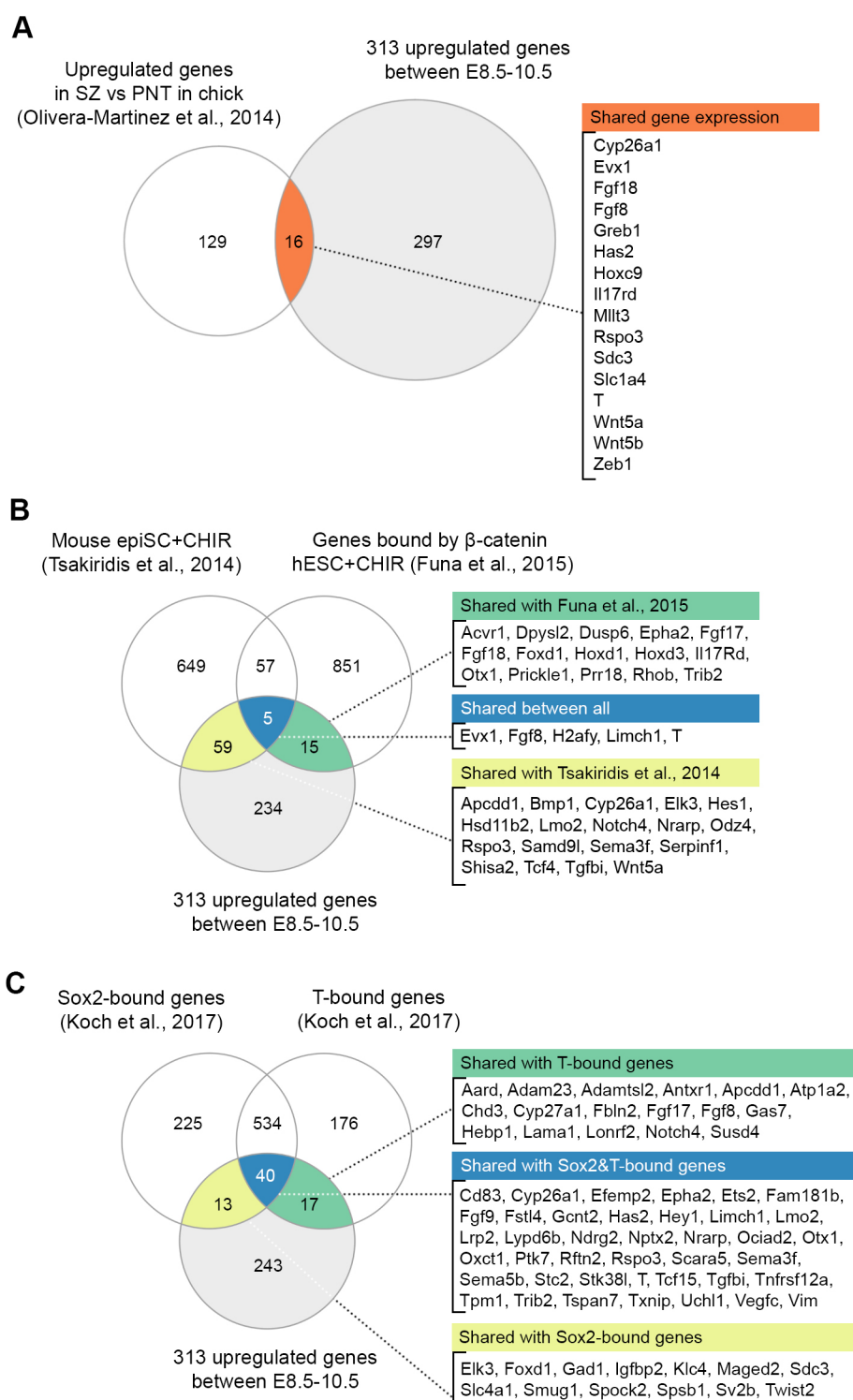


Figure S7. Comparison of upregulated genes at mid-trunk formation to other available datasets.

(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).

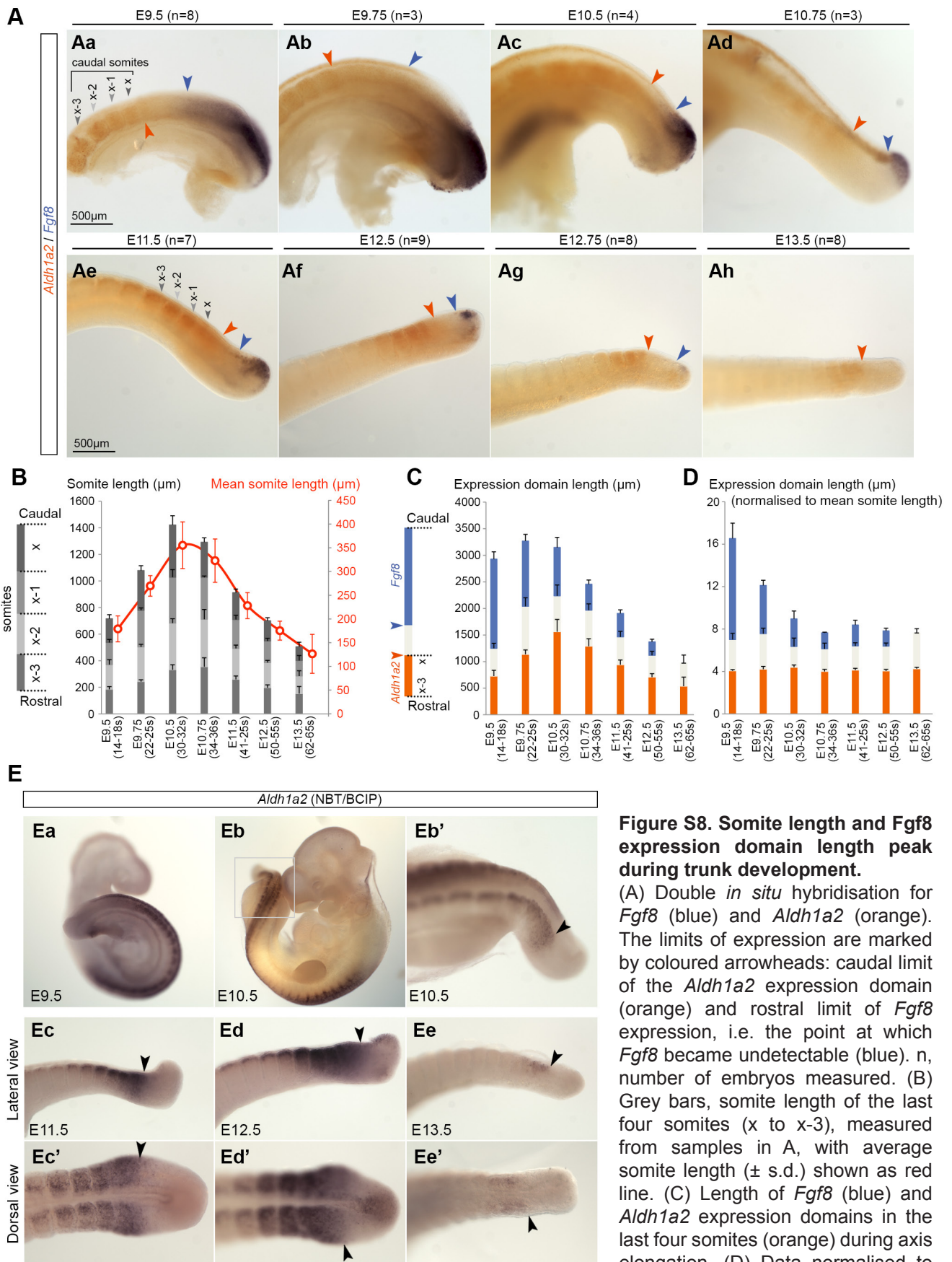


Figure S8. Somite length and *Fgf8* expression domain length peak during trunk development.

(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 (\pm 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).

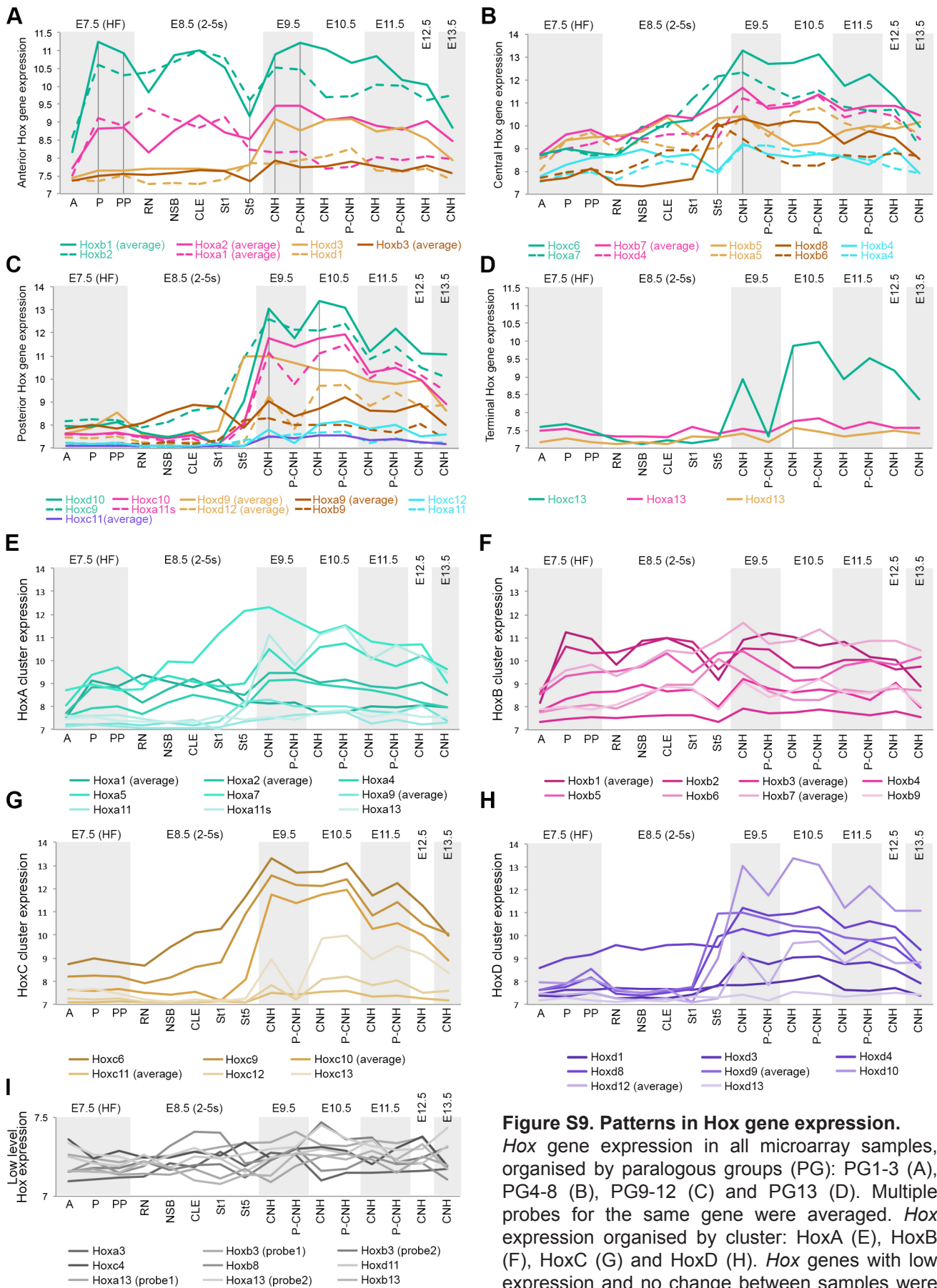
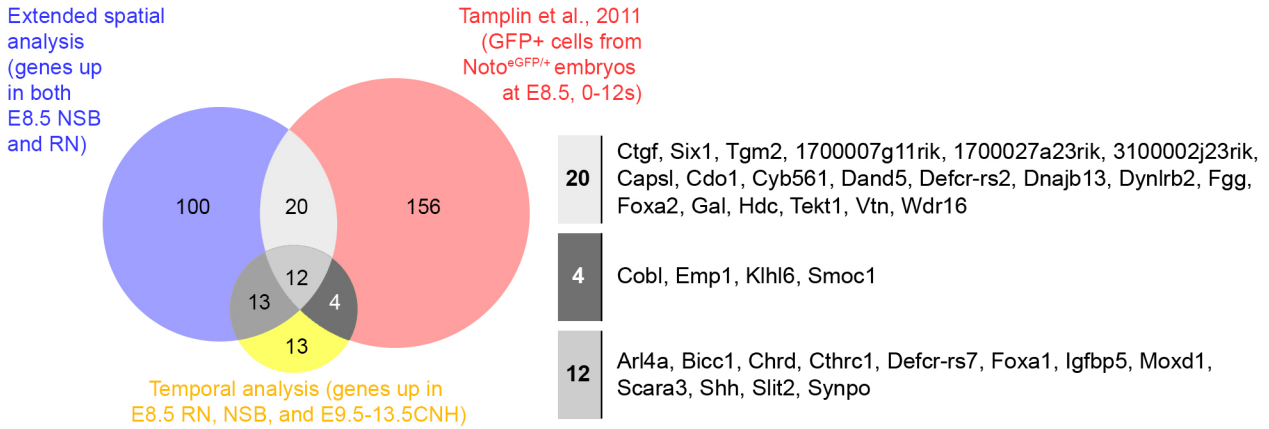


Figure S9. Patterns in Hox gene expression.

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

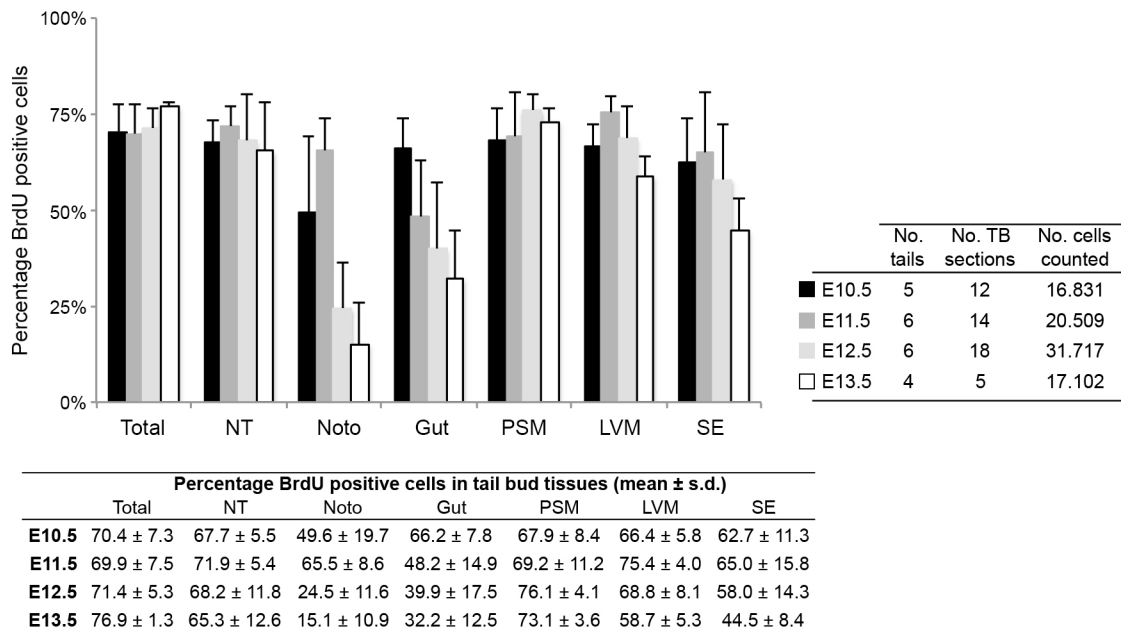


Figure S10. Experimental controls and comparison of *NotoP* markers.

(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.

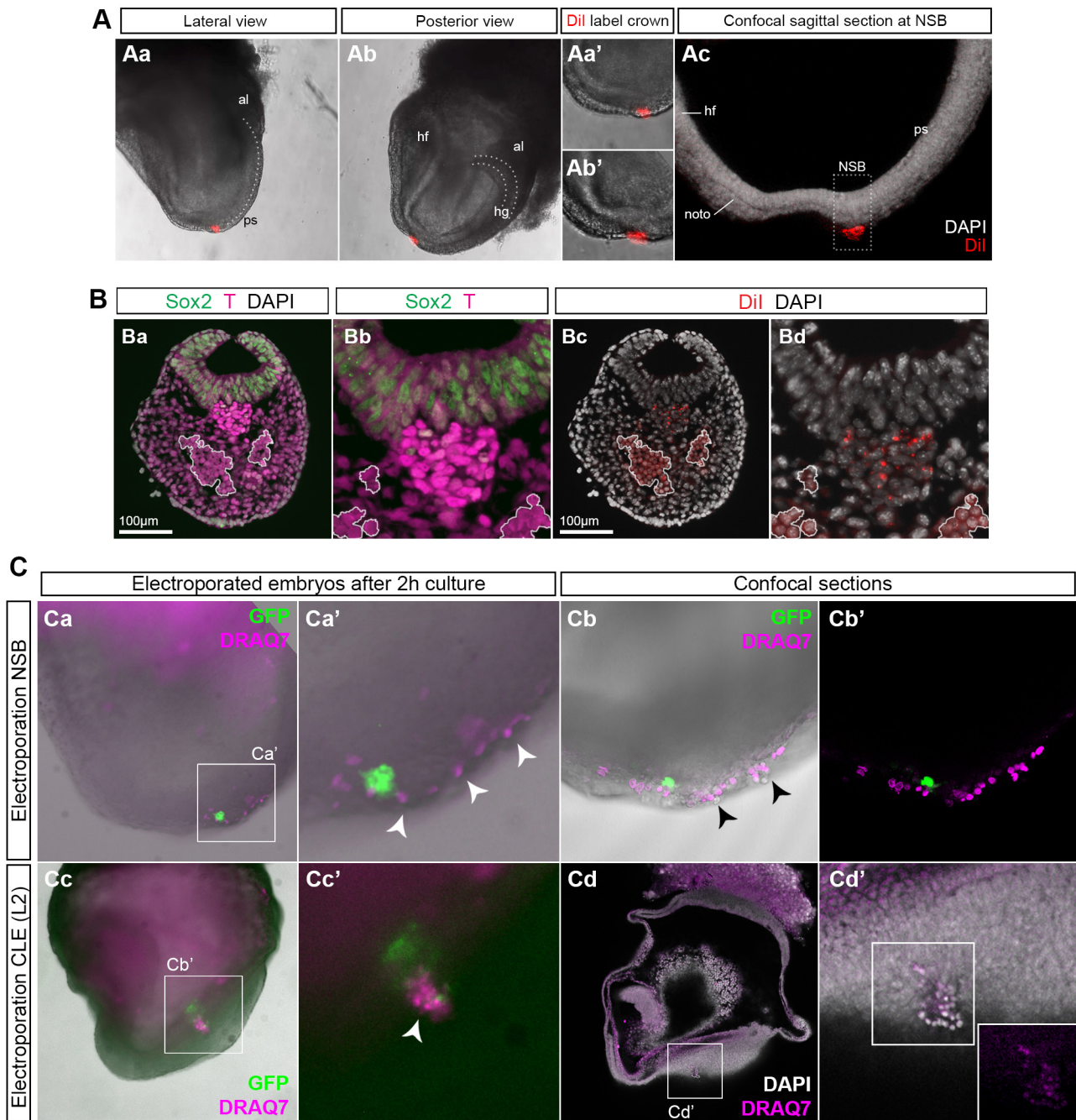


Figure S11. Dil labelling and electroporation controls.

(A) Dil labelling of the crown cells (red) in an E8.5 embryo: wholemount lateral (Aa), posterior view (Ab) validate the accuracy of initial labelling ($n_{\text{embryos}}=3$). (B) Sox2/T immunostained section of embryo labelled with Dil in the ventral node layer after 48h culture ($n_{\text{embryos}}=2$). 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).

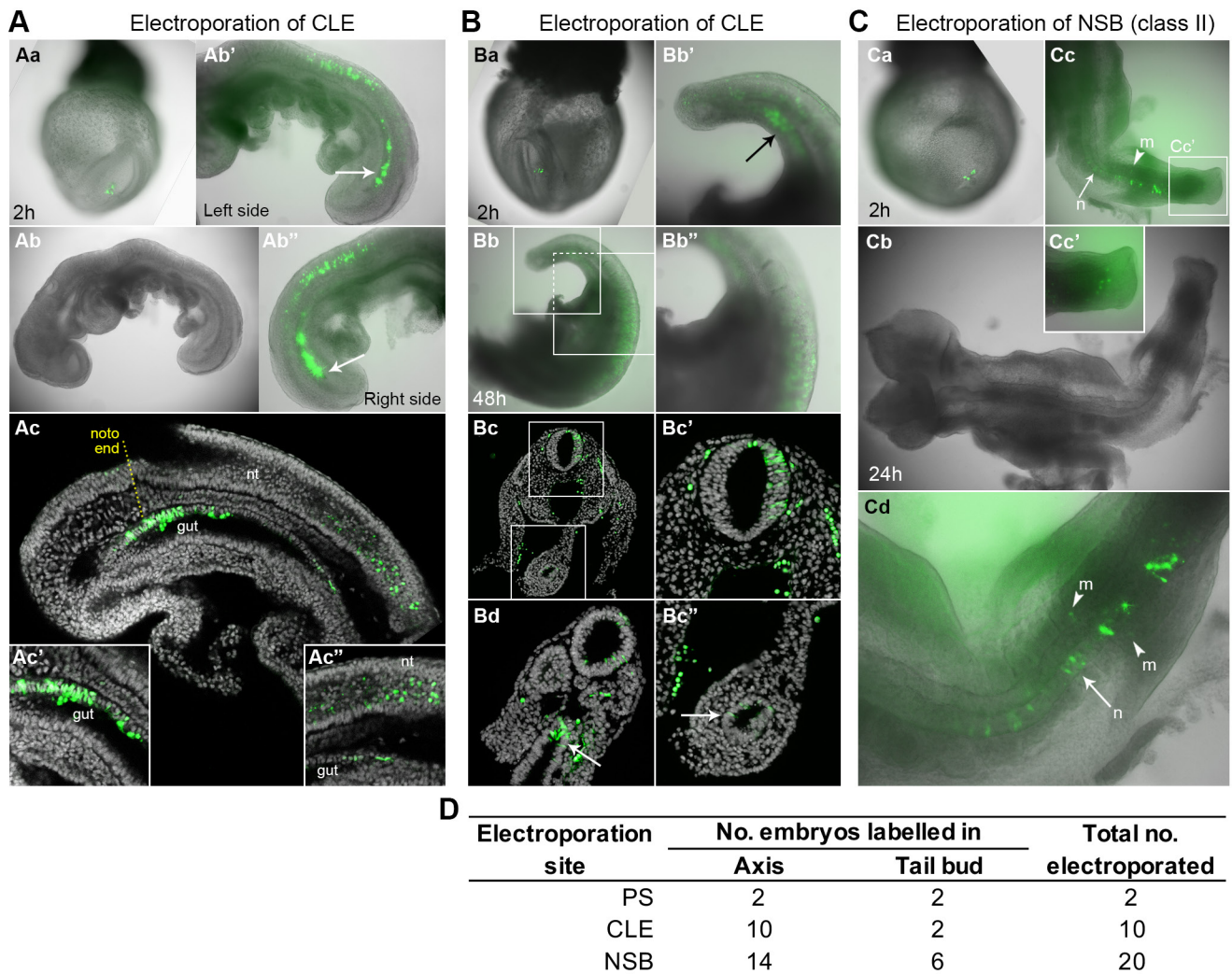


Figure S12. Fate of electroporated cells.

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; $n_{\text{embryos}}=4$), 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.

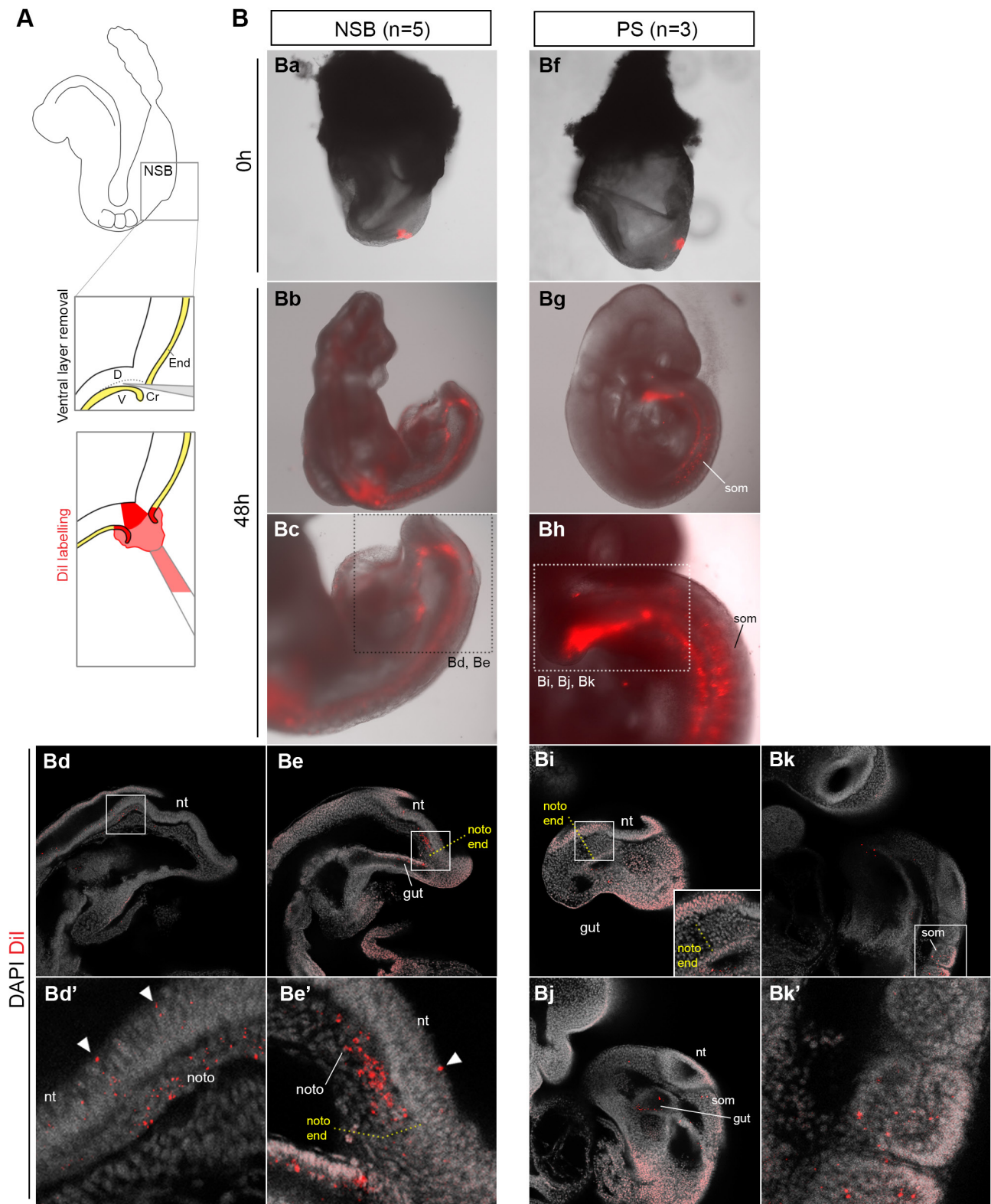


Figure S13. Effect of ventral cell layer removal at the NSB or St3.

(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 Dil 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, $n_{\text{embryos}}=5$, Fig. 8C). Dil-labelled cells were found in the notochord and neural tube (Bd-Be). Control embryos grew normally ($n=3$) and Dil was found in somite and gut tissues (Bg-Bh). Red dots, Dil (pink, background due to high exposure). nt, neural tube; noto, notochord; som, somites; white arrowheads, Dil 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	caaccaattatgggctggc
	Rev	ccacaatccaatcattccagc
Evx1	Fwd	gtttcaagaccgcgagat
	Rev	tgacgcttgctccttcatgc
Fgf4	Fwd	ccggttcttcgtggctatga
	Rev	cttactgagggccatgaacatacc
Fgf8	Fwd	atggcagaagacggagacc
	Rev	ttgttcatgcagatgtagagacc
Foxa2	Fwd	catccgactggagcagcta
	Rev	gcgcccacataggatgac
Mixl1	Fwd	agttgctggagctcgtcttc
	Rev	agggcaatggaggaaaactc
Sox2	Fwd	ggcggcaaccagaagaacag
	Rev	gcttggcctcgtc gatgaac
T(Bra)	Fwd	ccaaggacagagagacggct
	Rev	agtaggcatgtccaagggc
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.