

Supplementary Note: Supplementary Table legend

Supplementary Information Table 1. Embryo collection information. Columns contain: the sample number (“sample”); the relevant dataset (“dataset”); the cell genotype (“genotype”); whether the cells were injected (applicable only to chimera samples; “injected”); the embryo pool the sample was from (“pool_number”); the number of embryos in the pool (“num_embryos”); the estimated number of cells loaded (“loaded_cells”); the loading densities (in cells per μl ; “loading_density_cells_per_ul”); the expected number of cells recovered (“expected_cells”); the expected multiplet rate (estimated from 1; “expected_mutliplet_rate”); the number of cells called (“called_cells”); the number of cells passing QC (“qc_pass_cells”); the number of cells retained for downstream analysis (“non_doublet_stripped_cells”); and a column for sample notes (“notes”). Where samples share the same embryo pool number, this indicates that cells from a given pool of embryos were split between several 10X channels. Sample 11 entirely failed QC.

Supplementary Information Table 2. Dynamically expressed genes along the trajectory from visceral and embryonic endoderm to hindgut. Columns contain: trajectory evaluated (“trajectory”); Ensembl Gene ID (“ensembl”); cluster number (“cluster”); MGI gene name (“mgi”).

Supplementary Information Table 3. Cell numbers for chimera haemato-endothelial cell abundance analysis. Row names correspond to the haemato-endothelial sub-clusters. Columns indicate either the cell numbers (“*_count”) or cell fraction after removal

of blood-relevant cell types (see Methods; “*_frac”). Tomato positive and negative cell populations are indicated by “tom_plus” and “tom_minus”, respectively.

Supplementary Information Table 4. Metadata for all atlas cells that passed QC.

Columns contain cell names (“cell”); 10X cell barcode (“barcode”), cell sample number with respect to Supplementary Information Table 1 (“sample”); time-point collected (“stage”); sequencing batch (“sequencing.batch”); Theiler stage (“theiler”); cell doublet call (TRUE being a doublet; “doublet”); doublet density (“doub.density”); stripped nuclei (TRUE being a stripped nucleus; “stripped”); first general clustering (“cluster”; mentioned only in Methods); subclustering of each of the clusters (“cluster.sub”; mentioned in Methods); subclustering for all cells by time-point (“cluster.stage”; available on website); subclustering for all cells by Theiler stage (“cluster.theiler”; available on website); manually curated annotation, as in Fig. 1 (“celltype”); cell type colour (“colour”, as hexcode); UMAP coordinates (Fig. 1c; “umapX”, “umapY”); coordinates of haemato-endothelial force directed layout (Fig. 3a; “haem_gephiX”, “haem_gephiY”); haemato-endothelial subclustering (Fig. 3a, “haem_subclust”); coordinates of endoderm force directed layout (Fig. 2a-c, “endo_gephiX”, “endo_gephiY”); endoderm trajectory they belong to (“endo_trajectoryName”); endoderm DPT values for trajectories (“endo_trajectoryDPT”); gut diffusion maps coordinates (endo_gutX, endo_gutY); gut pseudospacial coordinates (“endo_gutDPT”, Fig 2f); gut cluster label (Fig. 2c, “endo_gutCluster”).

Supplementary Information Table 5. Metadata for chimera cells from *Tal1*^{-/-} into wildtype that passed QC. Columns contain cell names ("cell"); 10X cell barcodes ("barcode"); cell sample number ("sample"); time-point collected ("stage"); tdTomato status ("tomato"); mapped stage ("stage.mapped"); mapped cell type ("celltype.mapped"); closest cell in the atlas after batch correction ("closest.cell"); haemato-endothelial subcluster annotation "haem_subclust.mapped").

Supplementary Information Table 6. Metadata for chimera cells from wildtype into wildtype that passed QC. Columns contain cell names ("cell"); 10X cell barcodes ("barcode"); cell sample number ("sample"); time-point collected ("stage"); tdTomato status ("tomato"); mapped stage ("stage.mapped"); mapped cell type ("celltype.mapped"); and calculated doublet density ("doub.density").