SUPPLEMENTAL DATA

SUPPLEMENTAL METHODS

Mice and tissue preparations
To obtain embryos, p57Kip2 heterozygous females (Zhang et al., 1997) were crossed with C57BL/6J or CD45.1 males. Wnt1-Cre (Lewis et al., 2013) and Gt(Rosa)26Sortm14(CAG-tdTomato)Hze (Madisen et al., 2010) mice were crossed to obtain embryos for lineage tracing. As p57Kip2 is imprinted, offspring receiving the deleted allele from the mother will have a complete knockout phenotype. The day of vaginal plug detection was considered E0. Embryos smaller than their littermates or lacking a heartbeat were excluded. Single cell suspensions were obtained by treating AGMs, placentas and yolk sacs with collagenase (Alfa Aesar; 0.125% in phosphate-buffered saline) at 37°C for 45-90 minutes. Fetal livers were dissociated by pipetting.

Transplantation Experiments
Dissociated cell suspensions of embryonic tissues were intravenously injected into recipients with different CD45 isoforms that had received a split dose of 9.2–9.5 Gy of γ-irradiation. After 1 and 4 months, donor contribution to the recipients’ peripheral blood was determined by flow cytometry, using anti-CD45.1-PE and anti-CD45.2-FITC antibodies (eBioscience). Mice were considered to be repopulated if the donor contribution was at or above 5%. Statistics were performed on GraphPad Prism and the Mann-Whitney U-test was used to determine significance levels.

Adrenergic blockers administration in vivo
Each adrenergic receptor was blocked from E8 of gestation. α adrenergic receptors (Adra1 and Adra2) were blocked by intra-peritoneal (i.p.) administration of phentolamine (5 mg/kg). β adrenergic receptors (Adrb1, Adrb2 and Adrb3) were blocked by oral administration of propranolol (0.5 g/l of drinking water) or, for specific β adrenergic receptor inhibition, i.p. administration of betaxolol (1mg/kg) for Adrb1, ICI 118,551 (1mg/kg) for Adrb2 and SR 59230A (5 mg/kg) for Adrb3, all from Sigma. All inhibitors were diluted in PBS, apart from propranolol which was diluted in drinking water and SR 59230A which was diluted in PBS containing 1% DMSO. All blockers were administered every 24h, apart from the β3 adrenergic receptor inhibitor which was administered every 12h.

Catecholamine detection by HPLC
Hematopoietic tissues were dissected and snap-frozen in liquid nitrogen. Perchloric acid (PCA; Fisher Chemicals) was added to the tissues and the samples were homogenized. The samples were spun and
the diluted supernatant (1:10 in 0.2 M PCA buffer) was auto-injected into the HPLC machine (234 Autoinjector, Gilson; Hyperclone, 5u, BDS, C18, 100A, 100 x 4.60mm column, Phenomenex). The HPLC system mobile phase consisted of 31.90 g citric acid, 2 g sodium acetate, 460 mg Octanesulfonic acid, 30 mg Ethylenediamine tetra acetic acid (EDTA), 150 ml Ethanol and distilled water up to 1 L; pH was adjusted to 3.6 with NaOH (all from Fisher Chemicals, HPLC grade). The following standards were used to calibrate the HPLC equipment prior to measurement acquisition: NA (noradrenaline), EP (epinephrine/adrenaline) and DA (dopamine), all from Sigma. The signal was detected by an electrochemical detector (ESA, Coulochem II; Parameters: E1 -200 mw, R1 10uA, E2 250 mw, R2 200 nA, flow rate: 0.750 ml per minutes, pressure: 129 bar) and data were collected by the Dionex Data System (Chromelieon). Statistics were performed on GraphPad Prism and unpaired t-test was used to determine significance levels, n=14 per genotype/stage; n=56 in total.

**Flow Cytometry and Cell Sorting**

The staining in all the experiments was performed on ice in the dark for 30 minutes. After washing, cells were resuspended in buffer (PBS/2% FCS/ 1% penicillin/streptomycin) containing a viability dye from the Sytox series (Life Technologies) or DAPI (Sigma). Intracellular staining was performed with the Cytofix/Cytoperm kit (BD), where a permeabilization and fixation step is inserted after viability staining with a fixable viability dye (eBioscience), followed by intracellular staining with primary antibodies for Gata3 (KT63; Absea Biotechnology) and Th (LNC1; Merck), followed by an anti-rat IgG secondary antibody (ThermoFisher). All antibodies were purchased from eBioscience, BD and BioLegend, unless otherwise stated. The following antibodies were used: CD41 (MWReg30), CD43 (S7), CD45.1 (A20), CD45.2 (104), CD34 (RAM34), Pdgfrβ (APB5), CD45 (104), VE-Cadherin/VEC (11D4.1), Ngfr (polyclonal; ANT-007-AO; Alomone Labs Ltd). Cell cycle analysis was performed with DAPI staining. Data were acquired on a Fortessa instrument or cells sorted on Influx or ARIA instruments (all from BD Biosciences). Data were analyzed using FlowJo software (vX.0.6, Tree Star, Inc.).

**Immunohistochemistry**

Embryos were fixed in 4% PFA (Sigma) in PBS for 1.5 hours at 4°C and equilibrated in 30% sucrose (Fisher Scientific) at 4°C overnight. The following day, the embryos were frozen in OCT (Sakura-Finetek) on dry ice and kept at -80°C until they were sectioned on a cryostat (OTF-5000, Bright Instruments and CM1900, Leica). Blocking solution (2% serum from the animal that the secondary antibody was raised in and 0.4% Triton-X 100 in PBS or 0.1% Tween (all from Sigma)) was added for 1 hour at room temperature (RT). Primary antibodies were then added in blocking solution and left at
4°C overnight. Primary antibodies were: p57Kip2 (rabbit H-91; 1:500, Santa Cruz Biotechnology), Th (mouse LNC1, 1:300, Millipore), Gata3 (goat polyclonal, 1:200, R&D Systems), CD34 (FITC RAM34, 1:200, BD), CD34-biotin (rat RAM34, 1:200, eBioscience), ckit (goat polyclonal, 1:100, R&D Systems), Phox2b (guinea pig, 1:200, kind gift from J.F. Brunet), Ngfr (goat C-20, 1:300, Santa Cruz Biotechnology), Pdgfra (rat APA5, 1:200, eBioscience) and Sox10 (goat N-20, 1:250, Santa Cruz Biotechnology). The next day, the secondary antibodies were added for 1 - 1.5 hours at RT. Secondary antibodies were: Alexa555 goat anti-rabbit (1:500), Alexa647 chicken anti-mouse (1:2000), Alexa647 goat anti-rat (1:500), Alexa488 donkey anti-goat (1:300), guinea pig (Life Technologies) and Cy™2 Streptavidin (Jackson ImmunoResearch, 1:50). Subsequently, DAPI solution was added to each slide (1:5000 dilution from 5 mg/μl stocks), which were incubated for 5 minutes. Slides were mounted with 150 μl of Mounting Medium Fluoromount-G (Southern BioTech). The images were acquired on an Axioimager Z2 upright Microscope (Zeiss) (Camera: ORCA Flash 4 v.2, Objectives: 40x Oil and 63x Oil). Extended Depth of Focus images were created from Z-stacks of separate tiles that were then stitched together. Acquisition and processing was performed using the imaging software ZEN 2011 (Zeiss).

**In situ hybridization**

Cryosections from frozen embryos prepared as above were air-dried for 20 minutes and fixed in 4% PFA/PBS for ten minutes at RT, washed three times with PBS and then acetylated (1.3% v/v triethanolamine, 0.175% v/v hydrochloric acid and 0.25% v/v acetic anhydride in nuclease-free water) for 10 minutes at room temperature. After three washes with PBS, slides were incubated in hybridization buffer (50% formamide, 5xSSC, 1xDenhardts, 0.02% polyvinylpyrrolidone, 0.02% bovine serum albumin), 0.1% Tween-20, 250μg/ml MRE 600 tRNA, 500μg/ml salmon sperm DNA) for one hour at RT. Digoxigenin-labelled Cxcl12 riboprobe was diluted 100pg in 100μl hybridization buffer, denatured at 80°C for five minutes and then added to the slides. Slides were covered in parafilm and hybridized overnight at 65°C in a humidified box. The following day, slides were placed in 5xSSC at 65°C to allow the coverslips to detach before four washes in 0.2xSSC at 65°C for 25 minutes each and a final wash for five minutes at RT. Slides were then washed in buffer 1 (0.1M Tris at pH 7.5, 0.15M NaCl) for five minutes and pre-blocked in buffer 2 (10% heat-inactivated sheep serum in buffer 1) for at least one hour at RT. They were incubated overnight at 4°C with 1:4000 alkaline phosphatase-conjugated anti-digoxigenin antibody (Roche) in buffer 2. On the next day, slides were washed in buffer 1, equilibrated in buffer 3 (0.1M Tris at pH 9.5, 0.1M NaCl, 50mM MgCl2, 0.24mg/ml levamisole) for five minutes at room temperature before being incubated in the dark with 2% Nitro-Blue Tetrazolium Chloride/5-Bromo-4-Chloro-3"-Indolyphosphate p- Toluidine Salt (NBT/BCIP; Roche) in buffer 3 at RT until staining developed. To stop the staining, slides were rinsed with TE solution (10mM
Tris, pH 7.5, 1mM EDTA), then water, fixed for 30 minutes in 4% PFA/PBS, then rinsed again with TE followed by water before being mounted in Hydromount mounting medium (National Diagnostics). Images were taken with a Zeiss AxioSkop2 microscope.

Cxcl12 and p57Kip2 fragments for riboprobe synthesis were amplified from E11 AGM cDNA by RT-PCR using the following primers: Cxcl12 fwd: TTTCACTCTCGGTCCACCTC, Cxcl12 rev: TAATTTCGGGTCAATGCACA; p57Kip2 fwd: CTGACCTCAGACCCAATTCC, p57Kip2 rev: GATGCCCACGAAGTTCTCTC. The gel-purified fragment was cloned into the p-GEM-T Easy vector (Promega) and digoxygenin-labelled probes generated by in vitro transcription using a DIG RNA labeling kit (Roche).

**Gene Expression Analysis by real-time PCR**

RNA was extracted using the miRNAeasy Micro kit (Qiagen) according to manufacturer instructions, and RNA quality was assessed by the High Sensitivity RNA Assay (Agilent Technologies) on an Agilent Tapestation instrument. The iScript Advanced cDNA Synthesis Kit for RT-qPCR (Bio-rad) was used according to manufacturer instructions for cDNA synthesis. 2x Sybr (Brilliant III Ultra-Fast SYBR QPCR; Agilent Technologies, UK) was used for qPCR on a LightCycler 480 system (Roche Diagnostics, UK). The program was set as follows: 95°C for 5 minutes, 55 cycles of 95°C for 10 seconds, 60°C for 10 seconds and 72°C for 5 seconds, followed by 95°C for 5 seconds and 65°C for 1 minute. At the end of the qPCR program, a melting curve was run at continuous acquisition mode (97°C, 5 seconds/°C), followed by a cooling step (40°C for 10 seconds). The Ct values were retrieved at the end of the run and data were analyzed in Microsoft Excel according to the ΔCt method. All samples were run in at least three biological replicates, each of which was run in technical triplicates. Graphs were made on GraphPad Prism and the student t-test was used to determine significance levels. Primer sequences were: p57Kip2: F: 5’-CAGCGGACGATGGGAAGACT-3’, R: 5’-CTCCCGTGCTCTGCTACATGAA-3’; Gata3: F: 5’-CGAAACCGGAAGATGTCTAGC-3’, R: 5’-AGGAACTCTTCGCACACTTGG-3’; Th: F: 5’-TATGGAGAGCTCCTGCACTC-3’, R: 5’-TTCTCGAGCTTGCTTGGC-3’; β-Actin: F: 5’-GGCTGTATTTCCCCCTCATCG-3’, R: 5’-CCAGTTGGTAAACCATGCCATGT-3’.

**Library preparation for scRNA-Seq**

scRNA-seq analysis was performed using the Smart-seq2 protocol as described previously (Picelli et al., 2014). Single Ngfr+Pdgfrb- cells were index-sorted by FACS directly into individual wells of a 96-well plate containing lysis buffer (0.2% RNase inhibitor (Ambion, Thermo Fisher Scientific) in Triton X-100 (Sigma)), and libraries were prepared using the Illumina Nextera XT DNA preparation kit. Pooled libraries were run on the Illumina Hi-Seq 2500.
**Read alignment**

Reads were aligned using Kallisto (Linux v0.43.0, (Bray et al., 2016)) using parameters `kallisto quant --plaintext --bias --single --fragment-length=200 --fragment-length=200 --sd=20`. Samples were mapped to Mus_musculus.GRCm38.cdna.all.fa version of the transcriptome downloaded from Ensembl (www.ensembl.org) Feb. 2017. This library was appended with the ERCC92 spikes (ERCC92.fa) downloaded from www.thermofisher.com and appended to the transcript library. Read quality from individual libraries was assessed using FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc) and MultiQC, (Ewels et al., 2016)). Transcripts were trimmed using Cutadapt (Martin, 2011)) using the command `cutadapt -a CTGTCTCTTATA -f fastq -e 0.1 -O 3 -q 20 -m 20`.

**Data preprocessing**

Data was pre-processed following the single-cell transcriptomics workflow of Lun et al. (Lun et al., 2016) using R version 4.0.0 under 64bit Windows >= 8. Data were imported into R using the scater Bioconductor package (scater_1.18.6, (McCarthy et al., 2017)) function readKallistoResults. Transcript level quantitation was also collapsed to gene at this stage using a transcript to gene mapping table generated from Ensembl Transcript Stable ID and Gene Stable ID. All downstream analysis was carried out at gene level. Ensembl IDs were mapped to MGI Gene Symbols using Bioconductor mouse gene annotation files (org.Mm.eg.db_3.6.0).

Following the Lun et al. (Lun et al., 2016) workflow, cell quality was assessed based on library size, number of expressed features in each library, percentage of reads aligned to mitochondrial genes (obtained from Ensembl) and percentage of reads aligned to spike-in transcripts. We used the median absolute deviation (MAD) definition of outliers to remove putative low-quality cells from subsequent analysis: cells with any quality measurement more extreme than 3 MADs from the median were removed. Similarly, cells which had outlying proportions of mitochondrial genes or samples with low numbers of unique genes were also removed. Removing cells with a high proportion of reads mapping to spike-in RNAs left only one cell in the HSC1 sample group after filtering. This cell was removed from the data and the HSC1 group eliminated from further analysis. The HSC2 group was referred to as HSC in the rest of the analysis. Overall, 335 cells consisting of 65 HSC2, 93 Dorsal, 89 Ventral and 88 Mesenchymal cells passed initial filtering. We filtered genes with low abundance (those with mean
count across all cells less than 1), since these genes are likely to be dominated by drop-out events, limiting their usefulness in further analysis. 14,109 genes passed this filtering step.

We used the computeSumFactors function in the scran R package (Lun et al., 2016) to normalize within each plate and then scale plate samples to each other. Read counts mapping to ERCC spikes were normalized separately to the expression data and not used for normalization of expression samples. In this analysis we cannot completely exclude cofounding plate-effects. However, internal control marker and cell profiles were used to assess the normalization and are consistent with the expected biology.

**Cell cycle correction**

Following the workflow of Lun et al. (Lun et al., 2016), we used the Cyclone prediction method (Scialdone et al., 2015) to assign cells into cell cycle phases based on the gene expression data. Cells were assigned to a cell cycle phase if the score for that phase was greater than 0.5. Cells were considered of ambiguous cell cycle phase otherwise. Cells were initially visualized using their cell cycle phase before removal of this effect using limma (v3.46.0, (Ritchie et al., 2015)) by fitting a linear model (~ G1+G2M+S) and decomposing the variance. The trendVar function was used to fit a loess smoothed line to the gene specific variances and the decomposeVar function was used to decompose the variances into biological and technical variances. Genes were selected with the biological component of the variance >=0.5 and FDR <0.2 against the null hypothesis that biological variance equals 0. This generated a set of highly variable genes that was used as a set of genes for dimensional reduction with t-SNE.

**Cell clustering and dimensional reduction**

t-distributed Stochastic Neighbour Embedding (t-SNE; (van der Maaten and Hinton, 2008)) was used for dimensionality reduction. t-SNE visualizations were generated with varying settings of the perplexity parameter (5, 10 and 20) and the clusters obtained found to be robust. A t-SNE perplexity value of 10 was chosen as representative. t-SNE was performed across all the data and colored by group: HSC2, Dorsal, Mesenchymal and Ventral cells. A set of predetermined marker genes for Dorsal/Ventral (Phox2b, Ngfr, Sox10, Cdkn1c, Th, Gata3), HSC (Ptprc, Kit, Procr, Runx1, Cd34) and Mesenchymal (Pdgfra, Col1a2, Dlk1, Cspg4, Acta2, Pdgfrb and Cxcl12) were used to confirm the biological identify of each cluster.

A group of cells that were from Dorsal/Ventral origin were selected from the data and used to identify highly variable genes, dimensionally reduced using t-SNE and split into six clusters using k-means clustering. Robustness of the computed clusters was confirmed using the Dynamic Tree Cut method.
Analysis of the computed clusters revealed that the cells in one cluster expressed a range of known macrophage markers; cells in this cluster were removed from downstream analysis. Further analysis of the remaining clusters identified one cluster that was purely defined by cell cycle identity; following batch correction for cell cycle using limma and reclustering, cells in this cluster were found to be redistributed among the remaining four clusters.

**Differential expression analysis**
The edgeR Bioconductor package (v3.32.1; (Robinson et al., 2010)) was used to identify significantly differentially expressed genes between cell clusters. Differential expression results were filtered by fold change and an FDR threshold of 0.2. The top 40 differentially expressed genes for each cluster were computed using fold change ranking and a fixed FDR threshold of 0.2.

**Cell lineage inference analysis**
The Slingshot Bioconductor package (v.1.8.0; (Street et al., 2018)) was used to perform cell lineage inference analysis. The expression data and cell cluster labels were used as input to infer the global lineage structure. Based on this structure, two smooth cell lineages were constructed and pseudotime variables were inferred for both smooth lineages.

**Gene set enrichment analysis**
The GSEAPreranked tool (GSEA 4.1.0, build 27; (Subramanian et al., 2005)) was used for gene set enrichment analysis. To create the ranked gene list, genes were filtered to remove genes with logCPM \leq 0; remaining genes were ranked based on log fold change. Gene sets were obtained from the ConsensusPathDB-mouse interaction database (Kamburov et al., 2013); gene sets containing less than 15 genes or more than 500 genes were excluded from further analysis. The default GSEA enrichment statistic (weighted) was used, corresponding to a value of p=1 in the GSEA enrichment score calculation (Subramanian et al., 2005). GSEA’s meandiv method (default) was used to normalize enrichment scores to account for differences in gene set size and allow analysis over gene sets. Multiple hypothesis testing was corrected using sample permutation (n=1000).

**Ligand-receptor interactions**
scRNAseq data for E11-E11.5 HSCs was obtained from the counts matrix provided by Vink et al. 2020 GEO accession GSE143637 (Vink et al., 2020). Receptor-ligand interactions between cluster 3 (mesenchymal) and HSCs, as well as clusters 1, 2 and 4 (SNS) and HSCs were predicted using the liana package (v.0.0.8) (Dimitrov et al., 2021), which provides access to cell-cell communication databases,
including CellPhoneDB (Efremova et al, 2020). As the liana package is created for human data, the gene IDs in our gene expression matrix were converted from murine to human IDs using the biomaRt package (v.2.50.3.) (Durinck et al, 2009). The search was specifically focused on obtaining soluble factor information, so ligand and receptor information were defined using the import_omnipath_intercell() function within the OmnipathR package(v.3.3.20) (Turei et al., 2021). This custom resource was then fed to the liana_wrap() function, followed by liana_aggregate() to rank the results before visualization using liana_dotplot().
**Figure S1**

Sorting gates and fluorescence-minus one (FMO) controls for isolating endothelial cells (EC), hematopoietic stem and progenitor cells (HSPC), sympathoadrenal cells (SA) and mesenchymal cells (ME) from the AGM.
Figure S2

(A) Real-time PCR analysis of *Gata3* and *Th* expression in E11 wild-type (WT) and p57Kip2 knockout (KO) AGMs. Percentage of Gata3-Th+ (B) and Gata3+Th+ (C) cells within live cells of E11 wild-type (WT) and p57Kip2 knockout (KO) AGMs as determined by intracellular flow cytometry. Percentage of CD34+CD45+ (D), CD34+ (E) and Ngfr-Pdgfrb+ (F) cells within live cells of E11 wild-type (WT) and p57Kip2 knockout (KO) AGMs as determined by flow cytometry. A two-tailed t test was performed and error bars represent SEM. n=3. *Th* expression by real-time PCR in the AGM, fetal liver (FL), yolk sac (YS) and placenta (Pla) from E11 (G) and E12 (H) wild-type embryos. (I) Analysis of catecholamine levels by HPLC in the adult bone marrow (BM), AGM, FL, YS and Pla.
Figure S3

(A) t-SNE plots colored for clusters identified by k-means clustering (left) and for ventral and dorsal origin (right). Violin plots for expression levels of \textit{Ngfr} (B) and hematopoiesis-associated genes \textit{Ptprc/CD45} (C), \textit{Csf1r} (D), \textit{Cx3cr1} (E), \textit{Runx1} (F), \textit{Procr/EPCR} (G), \textit{Kit} (H) and \textit{Cd34} (I). (J) heatmap showing the top 40 differentially expressed genes in the macrophage cluster 3. (K) t-SNE colored for predicted cell cycle phases (top) and the identified clusters before (middle) and after (bottom) cell cycle correction.
Figure S4

(A-F) Heatmaps showing the top 40 differentially expressed genes in the remaining four clusters after cell cycle adjustments.
Figure S5

(A) Heatmaps showing the top 40 differentially expressed genes in the two alternative differentiation pathways of neural crest cells. (B) t-SNE plot of Ngfr+ cells (left) and Ngfr+ cells with Pdgfra+Ngfr-mesenchymal cells sorted from the ventral E11 AGM colored for the expression of Meox1.
Figure S6

Top Receptor-ligand interactions predicted between cluster 3 (mesenchymal) and AGM HSCs (A) and clusters 1, 2 and 4 (SNS) and AGM HSCs (B), with the soluble ligands derived from clusters 1-4, while the receptors are expressed in HSCs.