Developmental Timing Establishes Hematopoietic Stem and Progenitor Cell Lineage Bias

Anastasia Nizhnik[1], Bianca A. Ulloa[1&], Kevyn Jackson[2], Deyou Zheng[2,3], Teresa V. Bowman[#1,4]

1. Albert Einstein College of Medicine, Department of Developmental and Molecular Biology; Gottesman Institute of Stem Cell Biology and Regenerative Medicine,, Bronx, NY, USA
2. Albert Einstein College of Medicine, Department of Genetics, Bronx, NY, USA
3. Albert Einstein College of Medicine and Montefiore Medical Center, The Saul R. Korey Department of Neurology and Department of Neuroscience, Bronx, NY, USA
4. Albert Einstein College of Medicine and Montefiore Medical Center, Department of Oncology; Department of Medicine, Bronx, NY, USA

[#] To whom correspondence should be addressed.

Address: 1300 Morris Park Avenue, Bronx, NY 10461
Email: teresa.bowman@einsteinmed.edu
Phone: (718) 430-4001
Fax: (718) 430-8567

& Current Address:
Weill Cornell Medical Center& New York Presbyterian

Word count: 8,204
Abstract/first paragraph word count: 226
Number of figures: 7 main, 7 supplemental figures, 7 supplemental tables, 2 supplemental videos
Number of references: 86

Keywords:
Hematopoietic stem cells, HSCs, HSPCs, zebrafish, differentiation, lineage tracing, single cell RNA sequencing, Innate Lymphoid Cells, ILCs, hematopoietic ontogeny

ABSTRACT

Embryonic development is a critical time window for the establishment of the hematopoietic and immune systems. During embryogenesis, hematopoietic stem and progenitor cells (HSPCs) that possess divergent differentiation preferences arise and sustain lifelong hematopoiesis. How the pace and differentiation repertoire of native hematopoiesis is set in development and how it evolves as the organism grows is incompletely elucidated. Here, we use temporal lineage tracing of the emerging hematopoietic system during zebrafish development coupled with single cell RNA sequencing to define the origins of larval and adult hematopoietic cells. In both larvae and adult tissues, HSPCs arising earlier during embryogenesis show a predilection for lymphoid lineage production while those arising later are skewed towards erythroid. Moreover, early arising HSPCs are the main contributors to adult tissue resident lymphocytes. Mechanistically, early and late arising HSPCs show divergent transcriptional signatures and differential sensitivities to the levels of the transcription factor Runx1. Additionally, we identified that young larvae possess a more heterogeneous set of lymphoid cells than previously recognized including diverse T-lymphocytes and Innate Lymphoid-like Cells (ILCs). We demonstrated that these newly described larval ILC-like cells reside in lymphoid and mucosal organs and are responsive to viral mimic immune stimulation, indicating their relevance in early vertebrate life. The work provides new fundamental knowledge on how the heterogeneous HSPC pool establishes early immune hierarchies during embryogenesis and its persistence in adulthood.

INTRODUCTION

The hematopoietic system is comprised of an array of blood cells carrying out a variety of functions ranging from oxygenation to pathogen defense. These differentiated mature blood cells are produced by Hematopoietic Stem and Progenitor Cells (HSPCs) that possess multilineage differentiation potential. Although multipotency is a defining feature of HSPCs, their differentiation output is not always balanced. HSPCs can have lineage preferences or biases in the type of mature cells they make with observed predilections toward lymphoid, myeloid, or erythroid lineages.[1-3] HSPC lineage preferences are evident and important in human health and seem to have an age dependency to their features. For example, age-dependent myeloid skewing seems to be attributed in part to shifts in the relative frequency and function of lymphoid

34 and myeloid biased HSPC.[4] Additionally, lymphoid leukemias are more common in childhood,
35 while myeloid leukemias are more common with age. Although the mechanisms guiding lineage
36 preference of HSPCs remains elusive, recent work suggests that it could originate embryonically
37 and persist throughout life.[5]

38 Embryonic development is a critical stage for the formation of the hematopoietic system,
39 where multiple sequential and overlapping 'waves' of hematopoiesis produce distinct HSPCs
40 and mature cell types.[6-8] The cells generated at each wave serve a particular purpose for the
41 proper establishment and function of lifelong hematopoietic and immune systems.[9] Currently,
42 the field refers to two main waves of hematopoiesis: primitive and definitive.[8] The 'primitive wave'
43 gives rise to macrophages, erythrocytes, and neutrophils.[10,11] The 'definitive wave' is subdivided
44 into an earlier pro-definitive phase and a definitive phase.[9] The pro-definitive phase progenitors
45 generate myeloid lineage cells, megakaryocytes, B-cells, and NK cells.[9,12-14] The definitive
46 phase produces Hematopoietic Stem Cells (HSCs) that can sustain lifelong hematopoiesis. The
47 classic paradigm was that pro-definitive wave cells were transient and only necessary for fetal
48 life, while definitive wave HSCs sustained the entirety of the adult hematopoietic and immune
49 systems. Work over the past decade using lineage tracing methods that permit dissection of
50 native hematopoiesis have challenged this model and illustrated that pro-definitive wave
51 progenitors can contribute to adult steady-state bone marrow hematopoiesis,[5] tissue-resident
52 macrophages,[15,16] and innate B1a cells.[17] These paradigm-challenging studies illuminated that
53 much remains unknown about the differentiation repertoire of the heterogeneous pool of
54 embryonically emerging HSPCs and the persistence of their progeny beyond development.

55 While the embryonic origins of tissue-resident macrophages and B-1a lymphocytes are
56 studied,[7,18] hematopoietic origins of other Innate Lymphoid Cells (ILCs) are poorly resolved. ILCs
57 belong to the lymphoid lineage and resemble T-helper lymphocytes, however ILCs lack the
58 adaptive antigen-specific receptors characteristic of T lymphocytes.[19] ILCs are grouped into
59 three categories: ILC1, ILC2, and ILC3, paralleling the T-helper groups Th1, Th2, and Th17,
60 respectively.[20] In adult tissues, ILCs can secrete cytokines and recruit other immune cells in
61 response to infections but their behaviors in perinatal life is poorly resolved.[21] Moreover, it is
62 unknown currently if the lineage biases of embryonic HSPCs is relevant to ILC ontogeny.

63 While substantial work in the field dissected the differences between extraembryonic-
64 derived HSC-independent progenitors and intraembryonic-derived HSCs, there is less research

on the influence of developmental timings on HSPC properties. To address this gap, we sought to resolve the cell types generated by early *versus* late emerging HSPCs during zebrafish development using a temporally inducible lineage tracing system followed by single cell RNA sequencing (scRNA-seq). By lineage-tracing HSPC progeny in larval and adult life, we find that embryonic time imparts discrete and long-lasting differentiation functions onto HSPCs: early arising HSPCs produce more lymphoid lineage cells and later arising HSPCs show a preference towards the erythroid lineage. The lymphoid bias of early arising HSPCs in adult organs was evident not only in marrow but also in barrier tissues like the skin and intestine. Mechanistically, early- and late-emerging HSPCs exhibit differential sensitivity to the levels of the transcription factor Runx1. Furthermore, we uncovered that HSPCs make a diverse set of previously undescribed immune cells in larval zebrafish, including various groups of ILCs and several other novel immune cells that are responsive to immune stimulation in early life. Overall, the study presents new insights on the embryonic regulation of HSPC lineage bias across a broad swath of blood and immune tissues suggesting that embryonic environment could have long lasting impacts on hematopoietic aging and chronic inflammatory disease.

RESULTS

Lineage tracing and scRNA-seq resolve early, mid and late HSPCs in zebrafish.

During embryonic and larval development, long-lived HSCs and HSC-independent progenitors display distinct lymphoid-myeloid differentiation kinetics.[22-24] HSC-independent progenitors are the first to differentiate into mature immune cells, while HSC-derived progeny appear later.[22-24] While prior studies established this new paradigm, they were limited in the cell types they could examine due to technical limitations. The hematopoietic system is diverse, with numerous immune cell types that serve an array of functions and are found localized in many tissues. In a previous zebrafish lineage tracing study[22], neutrophils and thymic cells were examined, but it is unclear if the differentiation repertoire and kinetics that distinguish HSC and HSC-independent progenitors in larvae is unique to these cell types or more generalizable throughout the hematopoietic system. To address this question, we used an inducible Cre-Lox-based genetic lineage tracing approach that fluorescently labels HSPCs and their progeny, coupled with scRNA-seq to delineate the identities of labeled cells (**Fig. 1A**). The Cre-Lox system is comprised of two components: *Tg(draculin:creER^{T2})*, which allows hematopoietic-enriched

97 expression of a tamoxifen-inducible Cre recombinase, and *Tg(ubi:loxP-GFP-loxP-stop-*
98 *mCherry),* which allows for ubiquitous expression of fluorescent proteins.[25,26] Upon addition of
99 tamoxifen, Cre recombinase excises the loxP-GFP-loxP cassette, allowing for the expression of
100 mCherry. The result is permanent mCherry labeling of *draculin+* cells, including HSPCs and their
101 differentiated progeny. From prior studies employing the *drl:CreER^{T2}* system, it will label HSPCs
102 contributing to larval hematopoiesis including early thymic-seeding cells[22] as well as long-lived
103 HSPCs that sustain all blood lineages within the adult kidney marrow (mammalian bone marrow
104 equivalent),[27] suggesting it encompasses both HSC-independent progenitor and HSC
105 populations. We think this could include the previously described erythroid-myeloid progenitors
106 based on expression of *drl*-driven transgenes within the microenvironment where these cells
107 reside during development.[22] Previously, we determined that varying the window of tamoxifen
108 exposure during embryonic development helped to distinguish HSPC differentiation kinetics into
109 thymic cells and neutrophils.[22] In the current study, tamoxifen-mediated Cre induction of
110 mCherry labelling is begun at either one, two, or three days post fertilization (dpf), with 4-
111 hydroxytamoxifen added for 20 hours (**Fig. 1A**). We refer to the labeling at one dpf as "early-
112 HSPCs", at two dpf as "mid-HSPCs" and at three dpf as "late-HSPCs". These time points were
113 chosen as they represent early, middle, and late time points during HSPC specification,
114 maturation, and expansion during zebrafish development.[6] Based on prior studies,[22-24] we
115 hypothesized that early-HSPCs would be enriched for cells with rapid lymphoid differentiation
116 properties and mid/late-HSPCs would be enriched for long-lived HSCs.[22] To identify what cells
117 are labeled at early, mid, and late induction time points, we performed scRNA-seq on mCherry-
118 expressing cells isolated from larvae one-day-post labelling, so at two, three, and four dpf,
119 respectively. To assess the subsequent differentiated hematopoietic progeny, mCherry-
120 expressing cells were sequenced from 6 and 10 dpf larvae (**Fig. 1A**). After filtering out low quality
121 cells based on standard quality control parameters in the field,[28] we clustered cells in Seurat.[29]
122 In total, about 75% of mCherry-expressing cells were identified as endothelial or hematopoietic,
123 while the remaining 25% were non-hematopoietic cells like mesenchyme, neurons, cardiac,
124 muscle, and epidermis and thus were filtered out for subsequent analyses (**Supplementary Fig.**
125 **1A-D, Table S1**). As a result, 25 hematopoietic and immune clusters were detected across all
126 timepoints and traces, including HSCs, progenitors, and other cells of the major mature
127 hematopoietic lineages, including lymphoid, myeloid, and erythroid cells (**Fig. 1B,**

128 **Supplementary Fig. 1E,F & 2A-C, Table S1**). To facilitate ease of access for the community,
129 this full dataset and all others described below are available to search via a shinyapp webtool
130 made with ShinyCell[30] at https://sites.google.com/view/tvbowmanlab/developmental-
131 hematopoiesis-atlas.

132 The HSPC clusters were identified based on the expression of *myb*, which is a marker
133 gene commonly used to assess HSPCs by *in situ* hybridization.[31,32] Expression of *myb* across
134 all time points was localized to four stem and progenitor clusters identified as Myeloid
135 Progenitors (MyPs), Erythroid Progenitors (EryPs), Lymphoid Progenitors (LyPs), and HSCs
136 (**Fig. 1C,D**). The number of all HSPCs increased with developmental time (**Fig. 1E**). Differential
137 expression analysis comparing HSCs to EryPs, MyPs and LyPs showed distinct transcriptomes
138 across the four clusters (**Fig. 1F, Table S2**). Consistent with the lineage-enriched expression,
139 EryPs express higher levels of erythroid genes such as *klf17, alas2*, and *hemgn*, MyPs express
140 higher levels of genes critical for the myeloid lineage such as *spi2, csf2rb* and *mmp9*, and LyPs
141 express higher levels of genes associated with lymphopoiesis such as *runx3, tcf7,* and *il7r* (Table
142 S2). HSCs were enriched for genes encoding stem cell factors like *runx1*, *drl*, *apln,* and *itga2b*.

143 To identify transcriptional changes occurring in HSPCs as they mature during
144 development, we did gene enrichment analysis comparing each cluster over developmental time
145 (**Supplementary Fig. 2D, Table S2**). We observed that the HSC transcriptome was enriched
146 for genes involved in RNA splicing and metabolism at 1-day-post labeling and 10 dpf, with
147 elevated expression of genes involved with hypoxia at 6 dpf. In contrast, LyPs and MyPs showed
148 more similar changes with elevated expression of genes involved with proteostasis, protein
149 translation, and protein folding at 1-day-post labeling and 10 dpf. These observations are in line
150 with the rapid growth and protein production typically associated with hematopoietic
151 progenitors.[33]

152 To visualize the early-, mid-, and late-HSPCs *in vivo*, we did confocal imaging of the larval
153 zebrafish HSPC niche termed the Caudal Hematopoietic Tissue (CHT) at 1-day-post labelling
154 (**Fig. 1G**). The CHT contained the highest number of mCherry-labeled cells in the early-HSPC
155 trace and the lowest number in the late-HSPC trace. To identify if there were differences in the
156 composition of HSPCs labeled by each trace, we analyzed the HSPC subsets at 1-day-post
157 labeling in the scRNA-seq dataset (**Fig. 1H**). Although all four populations, MyPs, LyPs, EryPs
158 and HSCs, were detected in the early-, mid-, and late-traces, the frequency of the populations

159 was distinct. LyPs were most abundant in the early-HSPC trace, HSCs were enriched in the mid-
160 HSPC trace, while EryPs were most frequent in the late-HSPC trace. Comparison of differentially
161 expressed genes within the *myb+* HSC cluster at 1-day-post labeling from early-, mid-, and late-
162 HSPC traces uncovered differences related to lineage and signaling (**Fig. 1I**). These data
163 highlight differences within the HSPC pools across the window of HSPC specification,
164 maturation, and expansion during zebrafish development.
165

166 **HSPCs differentiate into a large repertoire of immune and hematopoietic cell types**
167 **during larval life.**

168 By 6 and 10 dpf, HSPCs differentiate into various immune and hematopoietic lineages.
169 Consistent with our prior study,[22] there is a substantial contribution of the early-HSPCs to the
170 CHT and thymus at 6 dpf, with minimal contributions from mid- and late-HSPC traces, as
171 observed by confocal imaging (**Fig. 2A,B**). The precise identity of the hematopoietic cells arising
172 in the hematopoietic tissues was unclear, so we leveraged the scRNA-seq data to identify the
173 immune cells derived from early-, mid-, and late-HSPC lineage traces. Lymphoid and myeloid
174 lineage clusters were subset and analyzed by developmental timepoint (**Fig. 2C, Table S3**).
175 When examined as all lymphoid clusters in aggregate, we observed a burst of differentiation
176 from the early-HSPCs, and a concomitant decrease in the proportion of LyPs, potentially due to
177 ongoing differentiation (**Fig. 2D**). Compared to mid- and late-HSPCs, the early-HSPCs showed
178 a stronger propensity for differentiation into lymphoid cluster 3 identified as ILC3/il2rb+Tcells.
179 This observation parallels confocal imaging, where early-HSPCs contributed significantly more
180 to the thymus than mid- or late-HSPCs (**Fig. 2A**). Not all lymphoid clusters behaved the same.
181 For example, cluster 13 rag+ T cells arose at similar frequency from early- and mid-HSPCs,
182 while late-HSPCs did not generate rag+ T cells by 10 dpf (**Fig. 2D**). In contrast to the burst of
183 lymphoid production, analysis of all myeloid lineage clusters in aggregate showed a steady level
184 of differentiation from all three HSPC traces (**Fig. 2E**). However, the production of myeloid
185 lineage cells by the three HSPC pools was not entirely homogenous. Late-HSPCs showed a
186 bias toward making more MyPs at 6 and 10 dpf compared to early- or mid-HSPCs, suggesting
187 that HSPCs maybe skewed toward myeloid lineage differentiation during later larval
188 development. Furthermore, macrophages and Antigen Presenting Cells (APCs) showed
189 different enrichments among the three traces (**Fig. 2E**). In sum, the confocal imaging data along

190 with single cell RNA sequencing illustrate that early HSPCs are preferentially differentiating to
191 lymphocytes (**Fig. 2F**).

192 Immune cells derived from distinct developmental sources can have transcriptomes that
193 reflect different functions.[34] To determine if the immune cells derived from early-, mid- and late-
194 HSPCs are transcriptionally distinct, we did differential gene expression and pathway enrichment
195 analyses (**Fig. 2G**). The analysis was focused on neutrophils because there are established
196 gene signatures in zebrafish that can allow us to assess if transcriptional differences are due to
197 neutrophil maturation stage or reflect potential functional differences related to the originating
198 HSPC. First, we examined maturation stage based on M1 (immature) to M3 (most mature) gene
199 module scores defined by Kirchberger et al.[35] We compared the neutrophils from all three traces
200 at 10dpf as we posited these would be at similar maturation stages. The module score analysis
201 confirmed that hypothesis with no major differences in the distribution of neutrophil maturation
202 stages across the three traces with the majority of cells in the M1 immature stage (**Fig. 2H, Table**
203 **S3**). After excluding significant differences in neutrophil maturation state across the traces, we
204 next compared the transcriptomes for signatures that distinguish those derived from different
205 HSPC traces. The results show that neutrophils derived from early- and mid-HSPCs upregulate
206 metabolic processing of peroxide and NFkB-signaling, while those from late-HSPCs upregulated
207 actin polymerization (**Fig. 2I, Table S3**). To further identify factors that may be driving distinct
208 differentiation of neutrophils from the early-, mid- and late-HSPCs, we did pseudotime analysis
209 using psupertime, a method that takes into account the real biological timepoints in pseudotime
210 inference.[36] The cytokine *il1b* appeared to be a driver of neutrophils differentiation from early-
211 HSPCs but not from mid or late HPSCs, while *mmp9* was strongly associated with late-HSPCs
212 (**Supplementary Fig. 3**). The finding further indicates that neutrophils arising from HSPCs at
213 different developmental windows are transcriptionally distinct.

214

215 **HSPCs differentiate to ILC-like and Natural Killer-like cells.**

216 Initial analysis of the lymphoid subset of clusters in the scRNA-seq data revealed a highly
217 heterogeneous pool (**Fig. 1B**). To date, rag+ T-cells are one of the only defined lymphocytes in
218 young larval zebrafish, so the identity of the other lymphoid cells detected in the lineage traced
219 scRNA-seq is unclear.[37,38] Non-T lymphocytes like NK cells and ILCs were detected in juvenile
220 and adult zebrafish but have not been detected in larval developmental stages.[28,39] Mice and

221 humans have a highly diverse repertoire of lymphoid cells emerging during embryogenesis and
222 fetal life including many innate lymphocytes such as NK cells, ILC2s ILC3s, and B1A B-cells,[40-
223 42] thus we expected that some of the new populations could represent these cells. We posit that
224 the enrichment of hematopoietic cells via *drl*-based lineage tracing and the isolation of mCherry+
225 cells found throughout the entire larvae could permit detection of rare immune cells, like ILCs,
226 missed in prior larval zebrafish publications.[43,44]

227 Consistent with previous zebrafish work,[38] we detected rag+ T lymphocytes expressing
228 *lck, cd4-1, cd8b, tcra,* and *tcrb* in cluster 9 of the lymphoid lineage subset of cells in the scRNA-
229 seq data (**Fig. 3A,B, Table S4**). Other less defined populations in larval zebrafish expressed top
230 marker genes associated with mammalian ILCs and NK cells. Cluster 0 expresses a mix of ILC3-
231 like markers such as *il7r, il23r, il17ra1a* and *tnfa*, as well as T-cell marker genes like *tcra* and
232 *lck*. Cluster 2 expresses *eomesa*, *nkl.1, tbx21 (t-bet)*, and *gzm3.3* reminiscent of ILC1/NK-like
233 cells. Cluster 3 expresses *il13, il4, gata3, mafa (MAF)* consistent with an ILC2-like identity.
234 Cluster 7 highly expresses *gata3,* suggesting these could be ILC progenitors. Differential
235 expression analysis between clusters 0, 2, 3, and 7, representative of the three groups of ILCs
236 and potential progenitor, further highlights their unique transcriptional identities (**Fig. 3C, Table
237 S4**).

238 Aside from ILC-like identity, the lymphoid subset contains a variety of T-cells, progenitors,
239 and other previously uncharacterized immune cells in larval zebrafish (**Fig. 3A,B**). Cluster 4
240 appears to be immature T cells, expressing *tcf7, mxd, cd28,* and *zap70*. Cluster 8 expresses
241 *tcrg* and *trdc,* indicating a possible presence of gamma-delta T-cells. The putative identity of
242 other clusters is less clear. For example, cluster 6 expresses *blnk*, *ccl36.1,* and *il12rb2* possibly
243 indicating presence of a blnk+ B-cell progenitor although classic markers such as *pax5* are not
244 detected (Table S4). Cluster 5 expresses genes like *il12bb, tox, cd59*, and *rorb*, that are
245 expressed in many immune cell types making their identity less clear thus we refer to them as
246 cd59+ Immune Cells.

247 The various ILC types have prototypic tissue locations. For example, mammalian ILC2s
248 can be found residing in barrier tissues like lungs, intestine and skin.[45] To determine if larval
249 zebrafish ILC-like cells are in tissues that align with their hypothesized identity, we did whole
250 mount *in situ* hybridization for top marker genes of ILC-like cells in 6 and 8 dpf larvae. We
251 selected *eomesa* for the ILC1/NK-like cluster and *il7r* for the ILC3/il2rb+Tcells. To detect ILC2-

252 like cells, we selected the gene *si:dkey21e2.16*, which is a predicted serine-type endopeptidase,
253 and named it *swift* (**Fig. 3D**). The expression of *eomesa* puncta is detectable around the thymic
254 area, consistent with thymic NK cell localization[28] (**Fig. 3E**). The expression of *swift* is detectable
255 along the pharyngeal arches and intestines, which is consistent with mucosal localization of
256 mammalian ILC2s. Lastly, *il7r* expression was localized to the thymus, but some puncta were
257 present outside the thymus, along the pharyngeal arches. The two different locations are
258 consistent with *il7r* marking multiple pots of cells, possibly ILC3s and il2rb+Tcells.

259 To determine if zebrafish ILCs are similar to human ILCs, we did gene module scoring
260 using available datasets from human fetal stage ILCs.[46] The gene sets enriched in human fetal
261 ILCs were converted to zebrafish orthologs and then used to compute a score in the zebrafish
262 lymphoid cell subset (**Fig. 3F, Table S4**). For each ILC subtype, the module scores were highest
263 in the respective zebrafish ILC-like clusters. The result indicates that zebrafish ILCs are
264 transcriptionally similar to the human counterparts. These data illustrate a previously
265 unappreciated variety of lymphoid lineage immune cells that are detectable in hematopoietic and
266 non-hematopoietic zebrafish tissues similar with their mammalian counterparts.

267

268 **Zebrafish ILC-like cells show similar dependencies on Il2rγ and Rag1 as mammalian ILCs.**

269 ILCs and adaptive T-lymphocytes, specifically *cd4+* T-helper cells, express similar
270 genes.[19] Therefore, identification of ILCs solely based on gene expression is inadequate to
271 distinguish ILCs from T-helper cells. In humans and mice, both ILCs and T-lymphocytes depend
272 on the function of interleukin 2 common gamma chain (IL2Rγ) because the immune cells express
273 various interleukins that require IL2RG as a co-receptor.[47] Unlike T-lymphocytes, ILCs lack
274 antigen-specific receptors, and therefore do not require the function of Rag1 and Rag2
275 recombinase enzymes for their development.[20] To better parse out ILCs from T-cells in larval
276 zebrafish, we assessed the cells in *il2rga*^{Y91fs} homozygous mutants[48] and *rag1* mutagenized
277 zebrafish. We anticipated that zebrafish ILC-like cells would be diminished in an *il2rga* mutant
278 and unaffected by mutagenesis of *rag1*.

279 In zebrafish, Il2rγ is encoded by the duplicated genes *il2rga* and *il2rgb*, however only
280 *il2rga* functions during developmental lymphopoiesis.[49] We used the *il2rga*^{Y91fs} zebrafish mutants
281 to determine if larval ILCs require the function of Il2rγ. Heterozygous *il2rga*^{Y91fs} fish containing
282 either *Tg(drl:creER*^{T2}*)* or *Tg(ubi:loxP-GFP-loxP-stop-mCherry)* were crossed to obtain *il2rga*

283 wild-type, heterozygous, and homozygous embryos with both the tamoxifen-inducible Cre
284 enzyme and the switch cassette (**Fig. 4A**). Early-HSPCs were labeled at 1 dpf and the mCherry
285 expressing progeny were analyzed at 8 dpf using confocal imaging. In homozygous *il2rga*^{Y91fs}
286 mutants, mCherry-expressing cells were significantly reduced in the thymus and intestine
287 compared with the wild-type and heterozygous siblings (**Fig. 4B**). However, mCherry-expressing
288 cells in the CHT were unaffected by the loss of *il2rga,* consistent with the CHT mostly containing
289 HSPC, myeloid, and erythroid cells[50] which are not dependent on Il2rga. To further parse out if
290 ILC-like cells are dependent on *il2rga*, we did *in situ* hybridization for *eomesa, swift,* and *il7r* at
291 6 dpf and showed reduced expression of all three markers in *il2rga*^{Y91fs} homozygous mutants
292 compared to wild-type and heterozygous siblings (**Fig. 4C**).

293 To distinguish ILCs from T-lymphocytes, we mutagenized the *rag1* locus using
294 CRISPR/Cas9 in *Tg(drl:creER*^{T2}*,ubi:Switch)* zebrafish larvae and then assessed T-lymphocytes
295 and ILCs (**Supplementary Fig. 4A**). To control for off-target effects from mutagenesis, we
296 compared *rag1* crispants to larvae mutagenized at the *slc24a5 golden* locus, a non-essential
297 gene which is required for melanin pigmentation in zebrafish.[51] In *rag1* crispants, T-cells were
298 decreased when assessed by imaging of the T-cell fluorescent reporter *Tg(rag2:mCherry)* and
299 by *in situ* hybridization for *rag1* and *tcra*, confirming decreased Rag1 functionality
300 (**Supplementary Fig. 4B,C**).

301 To assess all lymphopoiesis in *rag1* and control *slc24a5* crispants, we performed scRNA-
302 seq on the early-HSPC trace-derived mCherry-expressing cells isolated from 8 dpf larvae (**Fig.**
303 **4D, Supplementary Fig. 4D,E**). We hypothesized that Rag1-independent cells would show no
304 change in abundance in the *rag1* crispants, while Rag1-dependent cells like T-lymphocytes
305 would be diminished. As anticipated, cluster 8 identified as *rag*-expressing T-lymphocytes was
306 diminished in *rag1* crispants (**Fig. 4D,E, Table S5**). Based on gene expression, clusters 1 and
307 7 appear to be the ILC3s and T-cells observed as cluster 0 in the original HSPC trace dataset
308 (**Fig. 3A**), but due to the *rag1* genetic perturbation, the two cell types separated in this dataset
309 (**Fig. 4D**). The T-cell identity of the newly discovered il2rb+ T cells (cluster 1) was confirmed by
310 their decreased abundance in *rag1* crispants. Clusters 2, 6, 7, 9, and 11 were identified as NK
311 or ILC-like based on top marker gene expression and were still present in high abundance in
312 *rag1* crispants in accordance with ILCs not requiring the Rag enzyme for their development (**Fig.**
313 **4D,E, Supplementary Fig. 4D,E**). ILC2-like cells are represented by two clusters in this dataset,

314 clusters 6 and 9. Cumulatively, the ILC2 levels were not significantly altered indicating Rag1
315 independence. Clusters 3 and 4 identified as cd59+ and blnk+ immune cells, respectively, were
316 also Rag1-independent. *In situ* hybridization for markers of ILC1/NK-like and ILC2-like cells,
317 *eomesa* and *swift* respectively, confirmed Rag1 independence (**Fig. 4F**). However, expression
318 of *il7r* decreased in *rag1* crispants, indicating that the cells marked by *il7r* include T-lymphocytes
319 consistent with the expression of *il7r* in both il2rb+ T-cell cluster 1 and ILC3-like cluster 7 (**Table**
320 **S5**). In summary, transcriptionally detected ILC-like cells in zebrafish show genetic
321 dependencies consistent with their proposed identity, and in agreement with mammalian ILCs.
322

323 **Zebrafish larval ILCs respond to viral mimicry.**

324 In adults, innate lymphocytes participate in the immune system response to pathogenic
325 insults by viruses, bacteria, and helminths.[21] Fetuses and newborns can also be exposed to
326 these pathogens or their immunogenic products, yet the full immune system response during
327 these life stages, especially by innate lymphocytes, is largely unknown. The identification of ILCs
328 and the ability to look organism wide in larval zebrafish provides an opportunity to explore their
329 early life immune responses. In humans, perinatal NK cells can respond to infectious agents
330 such as single-stranded RNA viruses like influenza and SARS-CoV2.[52-54] To test the
331 responsiveness of the larval immune system to a relevant perinatal infection known to stimulate
332 an innate lymphocyte population, we exposed larval zebrafish to the TLR7/8 agonist R848
333 (Resiquimod) as a means to mimic the presence of a ssRNA virus.[55]

334 To assess the global larval hematopoietic and immune response to viral ssRNA mimicry,
335 we did whole body confocal imaging of early-HSPC-derived mCherry+ cells at 8 dpf following a
336 two-day exposure to R848 (**Fig. 5A**). The treatment window of 6 – 8 dpf was selected as
337 representative of a perinatal equivalent time point in zebrafish. At 8 dpf, the R848-treated larvae
338 had a significant increase in mCherry signal in the thymus and CHT as well as an increased
339 number of puncta in the intestine as compared to untreated controls (**Fig. 5B**).

340 To determine which immune cells respond to the viral mimic, we did scRNA-seq on early-
341 HSPC-derived mCherry-expressing cells from 8 dpf *slc24a5* crispant controls and *rag1* crispants
342 treated with R848 (**Supplementary Fig. 5A,B**). The data were then used to decipher changes
343 in cell number and gene expression as readouts of the response to viral ssRNA mimicry.
344 Macrophages are known to respond strongly to TLR7/8 stimulation, so their response served as

345 a control.[56] In R848-treated larvae, there is a significant expansion of macrophage clusters 5, 9,
346 and 18 (**Fig. 5C, Supplementary Fig. 5C**). Gene enrichment analysis of R848-induced
347 transcriptional changes in these clusters revealed upregulation of genes that are concordant
348 with metabolic changes and defense response activation (**Supplementary Fig. 5D, Table S5**).

349 Consistent with the work in human perinatal NK cells, there was an expansion of cluster
350 12 ILC1/NK-like cells in both *slc24a5* controls and in *rag1* crispants treated with R848 (**Fig. 5C,D,**
351 **Supplementary Fig. 5C**). The cluster 6 il2rb+ T cells were also expanded in all R848-treated
352 samples, but there was no change in the number of ILC2-like and ILC3-like cells. Upregulation
353 in cytokine and immune signaling are another metric of an immune cell response, so we
354 examined transcriptomic changes in T-cell and ILC populations in response to R848 exposure.
355 In the cluster 6 il2rb+ T cells, there was an increase in genes associated with IFN-γ, STAT, and
356 NFκB signaling post R848 treatment (**Fig. 5E**). The cells also upregulated TNFα, several
357 cytokines, chemokines and granzymes, all consistent with an anti-viral immune response. The
358 ILC1/NK-like cells showed an increase in *myd88,* IFNγ related genes like *ifng1 ifi46, ifi45, irf9*,
359 STAT signaling components like *stat1b, stat3, il6st*, and higher expression of transcripts
360 encoding cytokines like IL-10 and IL-21. Although we did not observe substantial changes in the
361 number of ILC2-like and ILC3-like cells, these clusters also upregulated genes associated with
362 IFN-γ*,* NFkB and STAT signaling implying they were responsive to the R848 stimulation (**Fig.**
363 **5E**). There was also some cell-type specific immune signaling upregulation, like ILC2-like cluster
364 24 had an increase in IL-4 (*il4*) and ILC3-like cluster 20 had an increase in IL-7R, IL-17A/F (*il7r,*
365 *il17a/f3*), *nitr7a,* and *nitr3a* genes.

366 Gene set enrichment analysis (GSEA) using the hallmark gene set showed an Interferon
367 response in ILC1/NK-like and ILC3-like cells treated with R848, and a STAT3 signaling
368 enrichment in ILC1/NK-like cells (**Fig. 5F**). Enrichment analysis for GO terms (EnrichGO)
369 revealed an upregulation of actin polymerization and motility GO terms in ILC1/NK-like cells
370 which is consistent with a study showing increased migration of NK cells in responses to
371 R848[57](**Supplementary Fig. 5E, Table S5**). In agreement with this finding, we noted movement
372 of mCherry+ lineage traced cells out of the thymus of R848-treated animals (**Supplementary**
373 **Videos**).

374 As the scRNA-seq data suggested changes in cell number as well as potential mobility of
375 ILC1/NK cells upon R848 exposure, we did *in situ* hybridization in whole larvae at 8 dpf for the

376 ILC1/NK cell marker gene *eomesa*. (**Fig. 5G**). Concordant with the scRNA-seq result, *eomesa*$^{+}$
377 puncta were significantly increased by R848 treatment with more puncta noted outside the
378 thymus along the pharyngeal arches in treated larvae compared to controls, suggesting that
379 ILC1/NK-like cells are mobilized from the thymus. We also examined the expression of the ILC2
380 marker gene *swift* which was unchanged between controls and R848 treated (**Fig. 5G**). In
381 agreement with the observed expansion of cluster 6 il2rb+Tcells in the sequencing data,
382 expression of *il7r* was significantly increased in the thymus and appeared somewhat dispersed
383 from this tissue in response to R848, suggestive of activation (**Fig. 5G**). Taken together, the
384 evidence shows that R848 treatment elicited an immune response from the newly characterized
385 ILC-like cells in larval zebrafish.
386
387 **Early-HSPCs preferentially contribute to adult tissue-resident lymphocytes**
388 Temporal lineage tracing of larval hematopoiesis indicated that early HSPCs are the
389 major contributors to larval lymphopoiesis, including ILCs (**Fig. 2**). Based on the strong
390 preference of early HSPCs to generate lymphoid cells that seeded many tissues in larvae, we
391 posited that these cells might be a significant source of adult tissue-resident lymphocytes. To
392 address this question, we labeled early and late HSPCs using the *drl:CreERT;ubi:switch* system
393 and then performed scRNA-seq on mCherry+ cells from the kidney marrow (equivalent to
394 mammalian bone marrow), thymus, intestine, and tessellated lymphoid network (TLN, equivalent
395 to skin immune cells)[58] (**Fig. 6A**). The tissues were dissected from 4-month-old zebrafish and
396 hematopoietic cells were isolated by fluorescent-activated cell sorting enriching for mCherry+
397 cells. The sequenced cells were analyzed in Seurat and clusters were identified based on top
398 marker genes and the literature in the field. Across the tissues, we identified clusters
399 representative of all expected hematopoietic cell types such as thrombocytes, HSCs, erythroid
400 progenitors, neutrophils, macrophages, T-cells, B-cells, ILC1/NK-like, ILC2-like, ILC3-like, and
401 dendritic cell populations (**Fig. 6B, Supplementary Fig. 6, Table 6**). In the lymphoid and barrier
402 tissues, we also detected the new populations that we had found in larvae such as *blnk+* immune
403 cells and *cd59+* immune cells. As expected, based on cell type distributions in mammalian
404 counterparts and prior zebrafish studies,[28] the kidney was enriched for myeloid lineage cells,
405 while thymus, intestine and TLN contained more lymphoid lineage cells (**Fig. 6C**). The
406 distribution of lymphoid cell types was not identical across the tissues. Consistent with

407 mammalian bone marrow, the kidney marrow had mostly B-cells and T-cells with a lower
408 abundance of ILC1/NKs, some ILC2-like and no ILC3-like cells[19] (**Fig. 6B**). The intestine was
409 enriched for T lymphocytes and ILC3-like cells but did not contain ILC1/NK-like populations. The
410 TLN and the thymus both contained all three groups of ILC-like clusters.

411 To facilitate cross trace comparisons of early *versus* late HSPC contributions to adult
412 hematopoietic and immune landscapes, we normalized the data across samples by organ. Each
413 organ has a unique abundance of immune cells, and our goal is to define which fraction of those
414 are derived from early *versus* late HSPCs. For similar comparisons in mammalian studies, the
415 absolute number of hematopoietic and immune cells within a tissue of interest are quantified first
416 using pan-hematopoietic antibodies and then the relative fraction arising from a given lineage
417 trace is calculated. Due to the lack of antibodies or transgenic lines in zebrafish that label all
418 hematopoietic and immune cells, we adopted a slightly different approach. As the number of
419 immune cells detected in all tissues analyzed was highest in the early HSPC trace samples, we
420 normalized the abundance of immune cells for both samples to the early HSPC trace for each
421 organ. The analysis revealed early HSPCs contribute about two-fold more to the hematopoietic
422 and immune compartments in the thymus and skin-TLN (**Fig. 6C,D**). There is also a greater
423 fraction of early HSPC-derived cells in the kidney marrow and intestine, although to a lesser
424 extent.

425 Examination of the relative distribution of hematopoietic and immune cells in these adult
426 tissues derived from the early *versus* late HSPCs revealed similar differentiation preferences as
427 observed in larval life. For example, similar to the observation during larval stages, late-HSPC-
428 traced cells in the kidney marrow seemed biased towards the erythroid lineage as evidence by
429 more erythroid progenitors derived from the late-traced sample (**Fig. 6C&E**). In contrast, early
430 HSPCs contributed more to the lymphoid compartment in the kidney marrow, thymus, TLN and
431 to a lesser extent the intestine (**Fig. 6D**).

432 The developmental origin of barrier lymphoid immunity, especially ILCs, is poorly
433 understood, so we next examined the contribution of early *versus* late HSPCs to specific
434 lymphoid cell populations within the skin-TLN and intestine. In the skin-TLN, there was a 3.1-
435 fold difference between early and late traces especially the ILC1/NK-like and ILC3-like cells (**Fig.
436 6F**). The trend was similar in the intestine with 2.5-fold more ILC-like cells derived from early
437 HSPCs than late, particularly the ILC3 subtype (**Fig. 6F**). These findings indicate that although

438 there are distinct distributions of lymphoid cells across barrier tissues, ILC-like cells found in
439 them are mostly derived from HSPCs arising earlier in development. Overall, these results
440 suggests that the developmental bias to lymphopoiesis in early-arising HSPCs and
441 erythropoiesis in late-arising HSPCs persists not only to adult hematopoietic tissues like the
442 kidney and thymus but also key barrier immune organs like the skin and intestine.
443

444 **Developmental time imparts HSPCs with distinct transcriptional signatures and Runx1**
445 **sensitivities**

446 Given the observation that early-arising and late-arising HSPCs show persistent
447 differentiation preferences (**Fig. 6**), we posited that there could be different transcriptional
448 signatures in HSPCs from the early and late traces that hint at their final lineage biases. Cluster
449 4 within the kidney marrow samples expresses genes representative of HSCs and multipotent
450 progenitors (MPPs), thus we examined the differentially expressed genes between the early-
451 and late-derived cluster 4 HSC/MPPs to determine if the time of embryonic origination imparted
452 a long-lasting transcriptional signature to adult HSPCs. Early-HSPC trace derived HSC/MPPs
453 expressed higher levels of lymphoid and immune-related genes like *socs1a* and *hes6* and while
454 HSC/MPPs from the late-HSPC trace expressed higher levels of erythroid lineage genes like
455 *cahz* and *alas2* (**Fig. 7A**). KM HSC/MPPs derived from early-traced HSPCs expressed higher
456 levels of xenobiotic response genes while those from late HSPCs show enriched expression of
457 genes involved in erythrocyte and myeloid lineage differentiation (**Fig. 7B, Table S7**).

458 Next, we searched for upstream transcription factors (TFs) that might impart the
459 transcriptional differences that distinguish early and late arising HSPCs. We analyzed TF motifs
460 enriched across the genes differentially expressed between KM HSC/MPPs derived from early-
461 *versus* late-HSPC traces, which revealed several predicted upstream TFs with trace-specific
462 enrichment (**Fig. 7C, Table S7**). Among the high confidence TFs, ETV1, an ETS family
463 transcription factor linked to neurogenesis, was early-trace specific, while RUNX1, a Runt-family
464 transcription heavily involved in hematopoiesis, was late-trace specific. These data hint at
465 differential TF activity in the early and late HSPC traces. We also inspected expression of these
466 TFs in the KM HSC/MPP cluster as well as in the larval *myb+* HSC cluster to assess if they
467 displayed trace-specific gene expression. The zebrafish *etv1* gene is not expressed in our larvae

468 datasets, and *runx1* expression levels were similar in both traces at larval and adult stages (**Fig.**
469 **7D**).

470 These findings led to the hypothesis that early *versus* late arising HSPCs could be
471 differentially sensitive to Runx1 levels. RUNX1 is a transcription factor critical for HSPC
472 formation.[59,60] In mice, *Runx1* loss is embryonic lethal in large part due to the lack of
473 hematopoietic progenitors during fetal development.[59,61] In zebrafish, some *runx1* homozygous
474 loss-of-function mutants die due to hematopoietic failure but some can survive to adulthood
475 because of genetic compensatory mechanisms that seem to restore most of the hematopoietic
476 system.[62,63] To decipher if HSPCs forming during different developmental windows are
477 differentially sensitive to Runx1 levels, we did lineage tracing of early, mid, and late HSPCs in
478 *runx1*^{w84x} wild-type, heterozygous, and homozygous mutants and then used flow cytometric
479 quantification to examine the relative frequency of mCherry+ cells within all marrow cells and
480 within each blood lineage of 4 month old zebrafish (**Fig. 7E, Supplementary Fig. 7**). The early
481 and mid HSPCs contributed a similar percentage of mCherry+ cells to the total kidney marrow
482 across the genotypes (**Fig. 7F**). However, late HSPCs showed a reduction in the percentage of
483 mCherry cells in the *runx1*^{W84X} homozygous mutant. The data suggests that loss of Runx1
484 impacts the late HSPCs ability to contribute to adult hematopoiesis in the kidney marrow.
485 Furthermore, compared to wild-type controls, the late HSPCs in *runx1* homozygous mutants
486 contributed significantly less to the myeloid and erythroid lineages, while the lymphoid lineage
487 was not significantly affected (**Fig. 7F**). Combined, the data suggests that Runx1 loss may be
488 more severely impairing erythroid-myeloid biased HSCs that are labeled by the late trace in our
489 system compared to the lymphoid-biased cells labeled by the early trace.
490

491 **DISCUSSION**

492 In this study, we uncovered two major aspects of developmental hematopoiesis. We
493 illustrated that long-lasting lineage bias is etched into HSPCs at distinct developmental windows
494 with lymphoid lineage bias preceding other differentiation preferences, insights important for
495 understanding origins of clonal hematopoiesis and age-related hematologic health risks.
496 Furthermore, our characterization of innate lymphocytes in the larvae presents a new frontier for
497 zebrafish immunology with implications on understanding lymphocyte ontogeny and perinatal
498 lymphoid immune responses.

While hematopoietic transplantation is the gold standard for examining HSPC functionality, seminal studies measuring native hematopoiesis revealed greater complexity and functionality within the HSPC pool.[15,64] For example, studies examining native hematopoiesis in murine embryos illustrated that progenitors arising before definitive HSCs can differentiate to adaptive lymphocytes like T and B cells that persist to adulthood.[5,17] Furthermore using transposon-based lineage tracing, Patel *et al.* revealed that a subset of later arising HSCs are biased toward generating megakaryocytes.[5] In zebrafish, Tian *et al.* and He *et al.* showed that the progenitor generating the earliest wave of thymic seeding T-cells was HSC independent.[23,24] Our lineage-tracing findings build on these works by establishing that embryonic lymphoid-producing early HSPCs can generate a wide variety of adaptive and innate lymphocytes during larval life (**Fig. 1-4**). Furthermore, we demonstrated that HSPCs arising earlier during embryogenesis contribute to the life-long pool of lymphocytes residing in barrier tissues (**Fig. 6**). Prior work showed tissue resident macrophages arose from an earlier HSC-independent progenitor wave,[16,65] and our data now supports a broader contribution of early embryonic HSPCs to tissue resident lymphoid cells. Moreover, in a similar vein to Patel *et al.*, we observed that later arising HSPCs in zebrafish show a distinct lineage preference to those emerging earlier in embryonic development. The subset of HSPCs captured by our later lineage tracing paradigm showed a bias toward erythropoiesis in larvae and adults (**Fig. 1, 2, 6**).

Mechanistically, regulation of divergent HSPCs and their differentiation output remains to be elucidated. Recent work showed that HSPC heterogeneity is initiated in the hemogenic endothelium and is influenced by factors like apico-basal polarity of cells[66], and by Notch and Wnt signaling whose regulation in the embryo affects adulthood lineages.[67] For our work, we posited that transcriptional wiring imparted during development would leave hallmarks in adult HSPCs. Upstream regulator analysis in adult cells uncovered RUNX1 signatures in the late-derived HSPCs. Our assessment of hematopoiesis in zebrafish *runx1* mutants supports the transcriptional observation and illustrates that late-derived HSPCs are more dependent on Runx1 (**Fig. 7**). The effect was not due to level of *runx1* expression but rather differential functional requirement. Studies in mice also find temporal requirements for hematopoietic transcription factors C-MYB[68] and RUNX1[69]. Dependency on C-MYB is used to distinguish early primitive and later definitive waves of hematopoiesis,[16,65,68] and RUNX1 dependency can distinguish early from late yolk sac EMPs.[69] Given the embryonic lethality of homozygous murine

530 mutants, experiments connecting embryonic HSPCs and adult hematopoiesis remains limited.
531 As zebrafish can survive for weeks without blood due to their external development,[32] they offer
532 an opportunity to reveal novel insights on hematopoietic ontogeny, including our finding that
533 distinct adult-contributing HSPCs display differential Runx1 sensitivity. Future mechanistic
534 studies are needed to determine how Runx1 is imparting lineage bias on HSPCs during
535 embryogenesis.

536 The finding of an array of ILCs located across many hematopoietic and non-
537 hematopoietic depots throughout young larvae begs the question of why an organism would
538 produce ILCs so early in development. Perhaps the earlier wave is needed as a line of defense
539 against pathogens before the adaptive immune system matures in the fetus. Consistent with a
540 few limited studies on fetal NK cells following pathogen exposure,[53,54] our results show that in
541 addition to NK cells, ILC-like populations can respond to viral-like stimulation early in larval life.

542 To understand if or how ILC ontogeny can affect their function, the embryonic origins of
543 the multitude of tissue resident ILCs need better resolution.[24,70] In larval zebrafish, we find that
544 ILC3s are produced by the earlier-arising HSPCs (**Fig. 2**), which is similar to the developmental
545 origins of murine lymphoid tissue inducers, which are a subtype of ILC3s.[71,72] The other ILC
546 subtypes are not as well resolved in the larvae, but examination of the adult tissue resident ILCs
547 shows that a majority of ILCs and NK cells identified in the thymus, skin-TLN, and intestine are
548 derived from early-arising HSPCs (**Fig. 6**). The mechanism underlying how and when these ILCs
549 enter these tissues is unclear but work from Wang *et al.* showing that murine embryonic ILC
550 progenitors seed the liver, lung, and intestine during fetal development and later differentiate
551 into distinct ILC subsets suggest the pattern is established early in life.[73]

552 Overall, our work illustrates that HSPC lineage preference manifests in a temporal
553 manner during embryogenesis with lymphoid bias arising first. Moreover, we illuminated that
554 despite divergence in tissue distribution, tissue resident lymphocytes across many adult tissues
555 are derived from the lymphoid-biased early arising HSPC population. Understanding the
556 regulation of this temporal embryonic imprint has the potential to uncover a connection between
557 HSPC ontogeny to a variety of human autoimmune disorders and leukemias as well as age-
558 associated disorders linked to HSPC heterogeneity.

559

560 **Acknowledgements**

561 T.V.B. was supported by grants from the National Institutes of Health (NIH) R01DK121738,
562 R01DK131445, R01DK141169, and the Edward P. Evans Foundation. AN was supported by
563 NIH F31HL167600, NYSTEM award C34874GG (PI Frenette), and a Liang Zhu Memorial
564 Fellowship. BAU was supported by NIH F31HL152562 and T32GM007288 (PI Akabas), KJ was
565 supported by NIH F31HL180025, T32GM145438 (PI Query), and R01HL153920 (PI Zheng).
566 Instrumental feedback to the development of this project was provided by Dr. Kira Gritsman, Dr.
567 Ulrich Steidl, and Dr. Meelad Dawlady. We thank the Langenau lab for sharing *il2rga*^{Y91fs} mutant
568 lines generated as part of R24OD016761. We thank the core facilities instrumental for this work
569 that are supported by the Einstein National Cancer Institute's grant number P30CA013330,
570 including: Flow Cytometry Core Facility at Einstein and core members Yu (Joey) Zhang and
571 Ming Liu for assistance with FACS; the Einstein Genomics Core and core member David
572 Reynolds for scRNA library preparation; and the Analytical Imaging Facility. We thank Dr.
573 Roshan Sharma for meaningful feedback on scRNA-seq data and Dr. Eliseo Castillo and Dr.
574 Keir Balla for fruitful discussions on innate lymphoid cells. Thank you to Dr. Noura Ghazale for
575 meaningful discussions and technical assistance, and Dr. Kathryn Potts for meaningful review
576 and discussion of the manuscript. Graphics were created in BioRender with copyright
577 https://BioRender.com/rsdc34d and https://BioRender.com/riwx66c
578

579 **Author Contributions**

580 Conceptualization, TVB, AN, BAU; Methodology, TVB, AN, BAU, KJ, DZ; Investigation, TVB,
581 AN, BAU; Writing—Original Draft, TVB, AN; Writing—Review & Editing, TVB, AN, BAU, KJ, DZ;
582 Funding Acquisition, TVB, AN; Supervision, DZ, TVB.
583

584 **Declaration of Interests**

585 The authors declare no competing interests.
586

587 **METHODS**

588 **EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS**

589 Zebrafish were bred and maintained as described previously.[74] All fish were maintained
590 according to Institutional Animal Care and Use Committee (IACUC)-approved protocols in
591 accordance with the Albert Einstein College of Medicine research guidelines. Sex determination

592 in zebrafish is complex and cannot be accounted for at larval stages,[75] but for all adult stages
593 we used equal numbers of males and females.
594
595 **Transgenic and Mutant Lines**
596 For lineage tracing, we used *Tg(drl:creER^{T2})* and the *Tg(ubi:loxP-GFP-loxP-stop-mCherry)* also
597 known as *Tg(ubi:Switch).*[25,26] For all experiments, fish with two copies of
598 *Tg(drl:creER^{T2},ubi:Siwtch)* transgenes were crossed to WT AB fish (from ZIRC) so all resulting
599 experimental animals were hemizygous. We used the published *runx1^{W84X}* mutants and
600 *il2rga^{Y91fs}* mutants.[48,60] Genotyping of *runx1^{W84X}* mutants was done using a qPCR with a custom
601 designed Taqman SNP probe
602 GGTTCTGGTACCTTGAAGGCGATGGGCAGGGTCTTGTTGCAGCG[C/T]CAGTGTGTCGGC
603 AGGACGGAGCAGAGGAAGTTCGGGCTGTCGG from Life Technologies. The *il2rga^{Y91fs}*
604 genotyping was done as described in Tang *et al.*[48] For lineage tracing in *runx1^{W84X}* mutants, we
605 crossed *Tg(drl:creER^{T2}),runx1^{W84X}+/- to Tg(ubi:Switch),runx1^{W84X}+/-* zebrafish. To validate the
606 functional consequences of *rag1* crispants on T-cells, we used the *Tg(rag2:mCherry).*[76]
607
608 **Lineage Tracing Assay**
609 To induce fluorescent-based lineage tracing of HSPCs, embryos containing
610 *Tg(drl:creER^{T2},ubi:Siwtch)* were incubated in a 28°C incubator with 12 μM 4-hydroxytamoxifen
611 (4-OHT) (Sigma H7904) for 20 hours, starting at 30 hpf for early HSPCs, at 54 hpf for mid
612 HSPCs, and at 86 hpf for late HSPCs. Hydroxytamoxifen was washed out three times with
613 embryo water then larvae were kept in a clean dish until larval endpoint or raised in our zebrafish
614 aquatic system until adult time points.
615
616 **Single Cell RNA Sequencing**
617 *Cell isolation and Library Preparation*
618 <u>*Larvae:*</u> For the early-HSPCs and mid-HSPCs cohorts, 200-250 larvae were used at one-day-
619 post labelling, 6 dpf, and 10 dpf. To obtain enough fluorescent cells from the late-HSPCs derived
620 cohort at the three timepoints (4, 6, and 10 dpf), we processed 1,100, 200, and 170 zebrafish
621 larvae, respectively. For the *rag1* crispants/R848 treatment experiment, 35 8 dpf larvae were
622 processed per condition. For all samples, whole larvae were minced with a razor blade and

623 digested with a 1:65 dilution of a 5 mg/mL liberase (Sigma 5401127001) stock for 15 minutes at
624 37°C, followed by FBS inactivation as previously described.[22] The mCherry-expressing cells
625 were sorted into 500 μL IMDM + 10% FBS buffer using the BD FACSAria II instrument at room
626 temperature. The cells were immediately processed for cell capture by the Einstein genomics
627 facility, using the Chromium Next GEM Single Cell 3′ Kit v3.1 for samples in Figures 1-3 and the
628 v4 kit for *rag1* crispant/R848 samples in Figures 4 & 5. Cell numbers for each experiment are
629 listed in Tables S1-5.
630
631 *Adult Tissues:* For kidney, intestine, thymus, and Tessellated Lymphoid Network (TLN)
632 experiments, two adult fish were used per condition, one male and one female. All four organs
633 were collected from the same fish. Fish were euthanized with tricaine, and TLNs were dissected
634 first, then thymus, intestines, and kidneys last. TLN was dissected according to described
635 protocols.[58] In brief, a razor blade was used to scrape off the scales into 1X DPBS, and the
636 resulting suspension was filtered to remove the scales. Cells were centrifuged and resuspended
637 in DPBS for FACS. For preparing intestinal tissue, we developed a protocol based on two
638 publications.[77,78] Animals were not fed for 24 hours prior to dissection to improve the cleanliness
639 of the intestinal tissue. Intestines were dissected from the animals and then minced with scissors
640 in cold DPBS. The tissue was then digested with 1 mg/mL Collagenase IV (ThermoFisher
641 17104019) for 30 minutes at 37°C. The digested cells were filtered through a 40 μm cell strainer,
642 centrifuged, and resuspended in DPBS with 2 mM EDTA. Kidneys were dissected and then
643 mechanically triturated in FACS buffer to release the cells into suspension.[79] FACS buffer
644 contains 0.9XD-PBS, 5% FBS, and 1% Pen/Strep. To enrich for myeloid and lymphoid lineages
645 in the kidney, we dissected kidneys and then lysed the erythrocytes using a 1:10 dilution of ACK
646 buffer for 10 minutes. Cells were washed with 1XDPBS, filtered, and resuspended in FACS
647 buffer. Thymus was dissected under a fluorescent microscope and dissociated by trituration in
648 500 μL HBSS (No Calcium, No Magnesium) with 0.5%BSA + DAPI and filtered through a 40 μm
649 cell strainer. All FACS sorting was done on the BD FACS Aria II machine and mCherry positive
650 cells were sorted into 500 μL IMDM + 10% FBS. Libraries were prepared immediately by the
651 genomics core at Einstein using the Chromium Next GEM Single Cell 3′ Kit v4. The libraries
652 were sequenced on the Illumina NovaSeq 6000 machine. For early-HSPCs condition, we
653 sequenced 33K, 35K, 86K, 33K reads per cells for kidney, TLN, intestine, and thymus
654 respectively. For late-HSPCs condition we sequenced 42K, 49K, 130K, 51K reads per cell for

655 kidney, TLN, intestine, and thymus respectively. Cell numbers for each experiment are listed in
656 Tables S6.
657
658 *Quality Control and Analysis*
659 *Larvae datasets:* Fastq files were processed in Cell Ranger 7.0.1, except the *rag1* crispant/R848
660 experiment which was done in version 8.0.1. To generate the counts matrix in Cell Ranger, we
661 made a custom reference genome based on *Danio rerio* GRCz11 with the addition of mCherry,
662 GFP, and T cell specific genes: TCRaC, TCRbC1 and TCRbC2. All subsequent dimension
663 reduction and clustering was done in Seurat version 5.2.0.[29] No batch effect was apparent
664 between samples, so analysis was done without batch corrections (Figure S1-2). Cells with less
665 than 200 genes and more than 10% mitochondrial DNA were filtered out (Figure S1,5,6). For the
666 lineage tracing dataset in Figure 1, we used the 4,000 top variable features for initial dimensional
667 reduction, with 30 principal components (PC) and at cluster resolution of 1. After non-
668 hematopoietic cells were removed, the remaining hematopoietic cells were re-analyzed with top
669 2,000 variable features, 30 PCs and at cluster resolution of 1 (Figure S2A). Two more lower
670 quality clusters were identified and removed to obtain the final clustering using 2,000 variable
671 features, 20 PCs, and cluster resolution of 1. For the lymphoid only subset analysis in Figure 3,
672 we used 2,000 variable genes, 25 PCs, and 0.3 cluster resolution. For *rag1* crispant/R848
673 dataset in Figures 4-5, initial clustering was done using 4,000 variable genes, 30 PCs, and
674 cluster resolution of 1. After removing non-hematopoietic cells, final clustering was done using
675 2,000 variable genes, 20 PCs, and a cluster resolution of 1. For analysis of the lymphoid only
676 subset in Figure 4, we used 1,000 variable genes, 15 PCs, and 0.2 cluster resolution.
677
678 *Adult datasets:* All FASTQ files were processed in Cell Ranger 8.0.1 using the default counts
679 setting and the custom-built *Danio rerio* GRCz11 genome. All clustering analysis was done in
680 Seurat version 5.2.0.[29] We filtered low quality cells using the 10% mitochondrial genes threshold
681 and filtered out cells containing less than 200 genes. To filter out possible doublets, cells
682 containing a high number of genes were filtered out using a different cutoff for each sample. For
683 kidney samples, initial clustering revealed a batch effect among the neutrophil clusters, so we
684 did a batch correction method on the kidney sample. We used anchor based CCA integration
685 followed by clustering using 20 Principal Components (PCs), and 0.5 cluster resolution. Clusters
686 18 and 19 were removed from analysis because they were an unclear mix of cells. For thymus

687 samples, no batch effects were present, so clustering was done using 1,500 variable features,
688 20 PCs and cluster resolution of 0.5. For TLN, initial clustering was done using 4,000 variable
689 features, 20 PCs and a 0.5 cluster resolution. The dataset contained many erythrocytes, so we
690 further subset away the erythroid populations to examine only the lymphoid and myeloid immune
691 cells. The final clustering for TLN was done using 2,700 variable features, 20 PCs and 0.7 cluster
692 resolution. For intestine, initial clustering revealed batch effects and many ribosomal protein
693 genes across clusters. We removed all genes starting with *rps, rpl, rsl* and did anchor based
694 CCA integration to correct for batch effect. Final clustering was done using 20 PCs and 0.5
695 resolution, and 4 small clusters were removed due to unclear mix of cells.
696
697 *Cluster identification*
698 For all cluster identification, we used a combination of literature-based marker genes and an
699 assessment of top marker genes using the FindAllMarkers function in Seurat which compares
700 the target cluster to other clusters using the Wilcoxon Rank Sum test. Module scoring using the
701 built in AddModuleScore() function was also used to aid in cluster identification.
702
703 *Differential Expression and GO Enrichment Analysis*
704 For differential gene expression analysis, two clusters of interest were compared using the
705 FindMarkers function in Seurat using standard settings for Wilcoxon Rank Sum test. For 4-way
706 gene expression comparison, clusters were subset and compared using the FindMarkers
707 function with standard settings. For gene enrichment analysis, the list of genes obtained from
708 differential expression analysis was filtered to include only those with adjusted p-value of less
709 that 0.05, and an average fold-change greater than 1.5. Using the enrichGO function available
710 in ClusterProfiler package,[80] we identified overrepresented GO terms that belong to the
711 Biological Processes category. Displayed in Figure 2I, 7B, Supplemental Figure 2D, 5D,E are
712 selected GO categories from the output list with the full lists found in Supplemental Tables S2,
713 S3, S5, S7. The GSEA analysis for Hallmark gene set in Figures 1I and 5F were done by fold-
714 change ranking.
715
716 *Pseudotime Analysis*
717 The software Psupertime[36] was used to temporally order each cluster of cells in the
718 hematopoietic atlas dataset from the different lineage traces. We include in the supplement

719 results from psupertime on neutrophils (Figure S3). Psupertime was run with default settings
720 with the exception of "min_expression" set to 0.00 instead of 0.01 to include genes that were
721 expressed in any cell rather than at least 1% to ensure that rare lineage genes were not missed
722 due to low abundance of some cell types in some traces.
723
724 *Upstream Transcription Factor and Motif Enrichment*
725 Upstream motif analysis was done using the RcisTarget. The list of zebrafish genes that are
726 differentially expressed between early- and late-HSPC trace derived adult kidney marrow
727 HSC/MPP was filtered by log fold-change greater than 0.6 and adjusted p-value less than 0.05.
728 Genes were converted to human orthologs using gProfiler. For motif rankings, we used the hg38
729 file with 10kb up and down of TSS. Displayed in Figure 7C are the predicted high confidence
730 transcription factors with NES greater than 3.0. The full list of predicted upstream regulators is
731 found in Table S7.
732
733 **Confocal Microscopy and Analysis**
734 The CSU-W1 Nikon spinning disc confocal was used for all high-resolution imaging of lineage
735 tracing in wild-type and mutant conditions and in drug-treated larvae. For live imaging, larvae
736 were anesthetized with 0.01% tricaine and mounted in 4% methylcellulose on glass bottom
737 dished. Images were acquired using 10x lens with the following parameters: 600 millisecond
738 exposure, and 561λ laser at 65% power and the z-stack was set to 6 μm size steps. The videos
739 were generated by imaging 6dpf larvae embedded in 1% low melt agarose and anesthetized
740 with tricaine. Larvae were mounted after 1 hour of R848 treatment, and then R848 was re-added
741 after mounting during image acquisition. Images were acquired for 4 hours total, with 4-minute
742 loops and a step size of 2.5 μm. Images were analyzed in Fiji (ImageJ) version 2.1.0.[81] To
743 calculate mCherry signal in the thymus and CHT, we used Corrected Total Cell Fluorescence
744 formular: (Integrated density of region of interest) - (Area of region of interest x Mean
745 fluorescence of 6 different background regions). All quantification of live imaging was done
746 blindly to genotype and/or condition, with the exception of images in Figure 1G and 2A.
747
748 ***In situ* Hybridization**
749 The *tcra* RNA probe was made from Addgene plasmid #11262 gifted by Lisa Steiner.[82] The *il7r*
750 RNA probe was made from a PCR template that was amplified from cDNA collected from 10 dpf

751 whole larvae. The following primers were used to generate the PCR template: (1)
752 TAATACGACTCACTATAGGGGGCCTTCCTTTCACTTTCAG and (2)
753 CATTAACCCTCACTAAAGGGAACATGGATTAATGTCTATAGCAGAGGA. Transcription of the
754 anti-sense *il7r* RNA was done using T3 RNA polymerase (Sigma 11031163001). To make
755 *eomesa* and *swift* probes, the DNA sequences were synthesized and put into pTWIST-kan high
756 copy vector by TWIST Bioscience. These probes are available from Addgene [*eomesa-*
757 *#*236391*; swift- #*236392; *il7r-* 236393]. For the *eomesa* probe, a 971-bp fragment containing
758 exon 7 through the 3'UTR region and flanked by T7 and T3 primers was synthesized. For the
759 *swift* probe, a 922-bp fragment was made containing the entire coding region of the gene. Gene-
760 specific primers that contained T3 RNA polymerase sequence (forward) and T7 RNA
761 polymerase sequence (reverse) at the 5'end were then used to generate a linear PCR-amplified
762 template for *in vitro* transcription using T7 RNA polymerase (Sigma 10881775001) and DIG RNA
763 Labeling Mix (Sigma 11277073910). The whole mount RNA *in situ* hybridization was done based
764 on established zebrafish protocols with modifications for older larvae.[83] For 6 & 8 dpf stages,
765 pigment was removed by bleaching with a solution of 0.8%KOH, 0.9% H_2O_2, and 0.1% Tween20
766 for 35 & 45 min, respectively. Proteinase K (Sigma 3115887001) working concentration was
767 increased to 15 mg/mL, and digestion time was extended to 35 & 40 min for 6 & 8 dpf,
768 respectively. After digestion, samples were fixed with 0.25% glutaraldehyde for 25 and 30
769 minutes for 6 & 8 dpf, respectively. Incubation in the RNA probes was done for two nights at
770 70°C to maximize detection. To visualize the RNA probe signal, we used anti-DIG (Sigma
771 11093274910) followed by colorimetric detection using NBT/BCIP (Fisher PR-S3771). Samples
772 were incubated in NBT/BCIP for 1-3 days for *eomesa* or *il7r*, and 8 hours for *swift,* all at room
773 temperature. For experiments with *il2rga*^{Y91fs} mutants, each embryo was imaged, quantified, and
774 then genotyped.
775
776 **CRISPR/Cas9 Mutagenesis**
777 The crispant approach was used to mutagenize the *rag1* locus, where three guide RNAs and
778 S.p. Cas9 Nuclease V3 protein (IDT 1081058) were assembled into a complex and injected into
779 one-cell stage embryos.[84] Each guide RNA mix was assembled by annealing a tracr RNA (IDT
780 1072532) with the locus specific crRNA for 5 min at 95°C, and cooled at room temperature for
781 10 min. Three guides were used to target the *rag1* locus, two of them were designed in

782 CHOPCHOP[85] (*rag1* guide #1 AATGATGCCCACATCCCAGG, guide#2:
783 AATGATGCCCACATCCCAGG), and one sequence from ZFIN (*rag1* guide#3:
784 GGACTGCTGCCATAAGGGGA, ZDB-CRISPR-201209-4) (Figure S4). As a control for
785 mutagenesis, an established *slc26a4* (golden) guide RNA was used:
786 GGACAGACGTGTTTCTCCAG.[86] Each single guide RNA was assembled into a complex with
787 Cas9 separately. For the control, *slc24a5* targeting complex alone was done, and for *rag1*
788 crispants, 3 *rag1* guides + 1 *slc24a5* guide were mixed in equal volumes to make a four-guide
789 complex. A T7 endonuclease assay (NEB M0302L) was used to confirm mutagenesis, where
790 cutting occurs only in the presence of heteroduplex DNA indicating mutagenesis (Figure S4).
791

792 **Resiquimod R848 Drug Treatment**

793 The 500 μg of R848 powder was obtained from InvivoGen (tlrl-r848-1) and reconstituted in 500
794 μL of water for a final concentration of 1 μg/μL or 2.85 mM. To treat larvae, the stock was diluted
795 in embryo water to a working concentration of 28.5 μM. Larvae were incubated with R848
796 starting at 6 dpf and analyzed at 8 dpf.
797

798 **Flow Cytometry**

799 Adult 4 month old zebrafish were euthanized on ice and kidneys were dissected and placed into
800 500 μL of FACS buffer containing 1μg/mL DAPI (Sigma D8417). The FACS buffer is composed
801 of 0.9XD-PBS, 5% FBS, and 1% Pen/Strep. To release cells into suspension, the dissected
802 tissue was gently triturated with a p1000 pipette about 20 times. The cell suspension was filtered
803 through a 40 μm cell strainer and collected into flow tubes by centrifugation at 2500 rpm for 40
804 seconds. For results shown in Figure 7, the red blood cells were not lysed, and samples were
805 assessed on LSRII (BD Bioscience) instrument. Voltages were set appropriately and kept
806 consistent between replicates. Analysis of data was done in FlowJo v10.8.1 software and gates
807 were set as shown in Supplemental Figure 7. A wild-type control sample was used to delineate
808 the fluorescent from non-fluorescent cells.
809

810 **QUANTIFICATION AND STATISTICAL ANALYSIS**

811 Analysis was done in GraphPad Prism 10.2.3. For all experiments with three or more groups, a
812 one-way ANOVA with Tukey multiple testing corrections was used. For two groups, an unpaired

813 t-test was used as indicated in the figure legends. A Chi-square test was used to compare
814 number of cells per cluster in the stacked bar plots. In all plots, one point is a single zebrafish
815 larva or adult. For single-cell RNA sequencing, one replicate was done per condition. Confocal
816 imaging experiments were done with at least three independent experiments, and *in situ*
817 hybridization experiments were done at least two independent times. All error bars show
818 standard deviation from the mean. In all figures: ns = not significant, * - $p < 0.05$, ** - $p \leq 0.01$, ***
819 - $p \leq 0.001$, **** - $p \leq 0.0001$.
820

821 **Material Availability**

822 Reagents used in this study will be made available upon request. Plasmids for making RNA
823 probes are deposited at Addgene IDs 236391, 236392, 236393.
824

825 **Data and Code Availability**

826 The sequencing data is available via the GEO accession number: GSE292726
827 The processed data can be viewed at:
828 https://sites.google.com/view/tvbowmanlab/developmental-hematopoiesis-atlas
829 Codes can be made available upon request.
830

831 **REFERENCES**
832

833 1 Muller-Sieburg, C. E., Cho, R. H., Karlsson, L., Huang, J. F. & Sieburg, H. B. Myeloid-
834 biased hematopoietic stem cells have extensive self-renewal capacity but generate
835 diminished lymphoid progeny with impaired IL-7 responsiveness. *Blood* **103**, 4111-4118
836 (2004). https://doi.org:10.1182/blood-2003-10-3448

837 2 Challen, G. A., Boles, N. C., Chambers, S. M. & Goodell, M. A. Distinct hematopoietic
838 stem cell subtypes are differentially regulated by TGF-beta1. *Cell Stem Cell* **6**, 265-278
839 (2010). https://doi.org:10.1016/j.stem.2010.02.002

840 3 Sieburg, H. B. *et al.* The hematopoietic stem compartment consists of a limited number
841 of discrete stem cell subsets. *Blood* **107**, 2311-2316 (2006).
842 https://doi.org:10.1182/blood-2005-07-2970

843 4 Ross, J. B. *et al.* Depleting myeloid-biased haematopoietic stem cells rejuvenates aged
844 immunity. *Nature* **628**, 162-170 (2024). https://doi.org:10.1038/s41586-024-07238-x

845 5 Patel, S. H. *et al.* Lifelong multilineage contribution by embryonic-born blood
846 progenitors. *Nature* **606**, 747-753 (2022). https://doi.org:10.1038/s41586-022-04804-z

847 6 Gore, A. V., Pillay, L. M., Venero Galanternik, M. & Weinstein, B. M. The zebrafish: A
848 fintastic model for hematopoietic development and disease. *Wiley Interdiscip Rev Dev*
849 *Biol* **7**, e312 (2018). https://doi.org:10.1002/wdev.312

850 7 Kobayashi, M. & Yoshimoto, M. Multiple waves of fetal-derived immune cells constitute
851 adult immune system. *Immunol Rev* **315**, 11-30 (2023).
852 https://doi.org:10.1111/imr.13192

853 8 Jagannathan-Bogdan, M. & Zon, L. I. Hematopoiesis. *Development* **140**, 2463-2467
854 (2013). https://doi.org:10.1242/dev.083147

855 9 Canu, G. & Ruhrberg, C. First blood: the endothelial origins of hematopoietic
856 progenitors. *Angiogenesis* **24**, 199-211 (2021). https://doi.org:10.1007/s10456-021-
857 09783-9

858 10 McGrath, K. E., Frame, J. M. & Palis, J. Early hematopoiesis and macrophage
859 development. *Semin Immunol* **27**, 379-387 (2015).
860 https://doi.org:10.1016/j.smim.2016.03.013

861 11 Le Guyader, D. *et al.* Origins and unconventional behavior of neutrophils in developing
862 zebrafish. *Blood* **111**, 132-141 (2008). https://doi.org:10.1182/blood-2007-06-095398

863 12 Dege, C. *et al.* Potently Cytotoxic Natural Killer Cells Initially Emerge from Erythro-
864 Myeloid Progenitors during Mammalian Development. *Dev Cell* **53**, 229-239 e227
865 (2020). https://doi.org:10.1016/j.devcel.2020.02.016

866 13 Yoshimoto, M. *et al.* Embryonic day 9 yolk sac and intra-embryonic hemogenic
867 endothelium independently generate a B-1 and marginal zone progenitor lacking B-2
868 potential. *Proc Natl Acad Sci U S A* **108**, 1468-1473 (2011).
869 https://doi.org:10.1073/pnas.1015841108

870 14 Iturri, L. *et al.* Megakaryocyte production is sustained by direct differentiation from
871 erythromyeloid progenitors in the yolk sac until midgestation. *Immunity* **54**, 1433-1446
872 e1435 (2021). https://doi.org:10.1016/j.immuni.2021.04.026

873 15 Ginhoux, F. *et al.* Fate mapping analysis reveals that adult microglia derive from
874 primitive macrophages. *Science* **330**, 841-845 (2010).
875 https://doi.org:10.1126/science.1194637

876 16 Schulz, C. *et al.* A lineage of myeloid cells independent of Myb and hematopoietic stem
877 cells. *Science* **336**, 86-90 (2012). https://doi.org:10.1126/science.1219179

878 17 Kobayashi, M. *et al.* HSC-independent definitive hematopoiesis persists into adult life.
879 *Cell Rep* **42**, 112239 (2023). https://doi.org:10.1016/j.celrep.2023.112239

880 18 Neo, W. H., Lie, A. L. M., Fadlullah, M. Z. H. & Lacaud, G. Contributions of Embryonic
881 HSC-Independent Hematopoiesis to Organogenesis and the Adult Hematopoietic
882 System. *Front Cell Dev Biol* **9**, 631699 (2021). https://doi.org:10.3389/fcell.2021.631699

883 19 Cherrier, D. E., Serafini, N. & Di Santo, J. P. Innate Lymphoid Cell Development: A T
884 Cell Perspective. *Immunity* **48**, 1091-1103 (2018).
885 https://doi.org:10.1016/j.immuni.2018.05.010

886 20 Spits, H. *et al.* Innate lymphoid cells--a proposal for uniform nomenclature. *Nat Rev*
887 *Immunol* **13**, 145-149 (2013). https://doi.org:10.1038/nri3365

888 21 Yuan, T. *et al.* Innate lymphoid cells and infectious diseases. *Innate Immun* **30**, 120-135
889 (2024). https://doi.org:10.1177/17534259241287311

890 22 Ulloa, B. A. *et al.* Definitive hematopoietic stem cells minimally contribute to embryonic
891 hematopoiesis. *Cell Rep* **36**, 109703 (2021).
892 https://doi.org:10.1016/j.celrep.2021.109703

893 23 Tian, Y. *et al.* The first wave of T lymphopoiesis in zebrafish arises from aorta
894 endothelium independent of hematopoietic stem cells. *J Exp Med* **214**, 3347-3360
895 (2017). https://doi.org:10.1084/jem.20170488

896 24 He, S. *et al.* In vivo single-cell lineage tracing in zebrafish using high-resolution infrared
897 laser-mediated gene induction microscopy. *Elife* **9** (2020).
898 https://doi.org:10.7554/eLife.52024

899 25 Mosimann, C. *et al.* Ubiquitous transgene expression and Cre-based recombination
900 driven by the ubiquitin promoter in zebrafish. *Development* **138**, 169-177 (2011).
901 https://doi.org:10.1242/dev.059345

902 26 Mosimann, C. *et al.* Chamber identity programs drive early functional partitioning of the
903 heart. *Nat Commun* **6**, 8146 (2015). https://doi.org:10.1038/ncomms9146

904 27 Henninger, J. *et al.* Clonal fate mapping quantifies the number of haematopoietic stem
905 cells that arise during development. *Nat Cell Biol* **19**, 17-27 (2017).
906 https://doi.org:10.1038/ncb3444

907 28 Rubin, S. A. *et al.* Single-cell analyses reveal early thymic progenitors and pre-B cells in
908 zebrafish. *J Exp Med* **219** (2022). https://doi.org:10.1084/jem.20220038

29 Hao, Y. *et al.* Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat Biotechnol* **42**, 293-304 (2024). https://doi.org:10.1038/s41587-023-01767-y

30 Ouyang, J. F., Kamaraj, U. S., Cao, E. Y. & Rackham, O. J. L. ShinyCell: simple and sharable visualization of single-cell gene expression data. *Bioinformatics* **37**, 3374-3376 (2021). https://doi.org:10.1093/bioinformatics/btab209

31 Murayama, E. *et al.* Tracing hematopoietic precursor migration to successive hematopoietic organs during zebrafish development. *Immunity* **25**, 963-975 (2006). https://doi.org:10.1016/j.immuni.2006.10.015

32 Soza-Ried, C., Hess, I., Netuschil, N., Schorpp, M. & Boehm, T. Essential role of c-myb in definitive hematopoiesis is evolutionarily conserved. *Proc Natl Acad Sci U S A* **107**, 17304-17308 (2010). https://doi.org:10.1073/pnas.1004640107

33 Gao, X., Xu, C., Asada, N. & Frenette, P. S. The hematopoietic stem cell niche: from embryo to adult. *Development* **145** (2018). https://doi.org:10.1242/dev.139691

34 Ginhoux, F. & Guilliams, M. Tissue-Resident Macrophage Ontogeny and Homeostasis. *Immunity* **44**, 439-449 (2016). https://doi.org:10.1016/j.immuni.2016.02.024

35 Kirchberger, S. *et al.* Comparative transcriptomics coupled to developmental grading via transgenic zebrafish reporter strains identifies conserved features in neutrophil maturation. *Nature Communications* **15**, 1792 (2024). https://doi.org:10.1038/s41467-024-45802-1

36 Macnair, W., Gupta, R. & Claassen, M. psupertime: supervised pseudotime analysis for time-series single-cell RNA-seq data. *Bioinformatics* **38**, i290-i298 (2022). https://doi.org:10.1093/bioinformatics/btac227

37 Langenau, D. M. *et al.* In vivo tracking of T cell development, ablation, and engraftment in transgenic zebrafish. *Proc Natl Acad Sci U S A* **101**, 7369-7374 (2004). https://doi.org:10.1073/pnas.0402248101

38 Langenau, D. M. & Zon, L. I. The zebrafish: a new model of T-cell and thymic development. *Nat Rev Immunol* **5**, 307-317 (2005). https://doi.org:10.1038/nri1590

39 Hernandez, P. P. *et al.* Single-cell transcriptional analysis reveals ILC-like cells in zebrafish. *Sci Immunol* **3** (2018). https://doi.org:10.1126/sciimmunol.aau5265

40 Dorshkind, K. & Crooks, G. Layered immune system development in mice and humans. *Immunol Rev* **315**, 5-10 (2023). https://doi.org:10.1111/imr.13198

941 41 Tang, Y. *et al.* Emergence of NK-cell progenitors and functionally competent NK-cell
942 lineage subsets in the early mouse embryo. *Blood* **120**, 63-75 (2012).
943 https://doi.org:10.1182/blood-2011-02-337980

944 42 Schneider, C. *et al.* Tissue-Resident Group 2 Innate Lymphoid Cells Differentiate by
945 Layered Ontogeny and In Situ Perinatal Priming. *Immunity* **50**, 1425-1438 e1425
946 (2019). https://doi.org:10.1016/j.immuni.2019.04.019

947 43 Petrie-Hanson, L., Hohn, C. & Hanson, L. Characterization of rag1 mutant zebrafish
948 leukocytes. *BMC Immunol* **10**, 8 (2009). https://doi.org:10.1186/1471-2172-10-8

949 44 Yoder, J. A. *et al.* Developmental and tissue-specific expression of NITRs.
950 *Immunogenetics* **62**, 117-122 (2010). https://doi.org:10.1007/s00251-009-0416-5

951 45 Naito, M. & Kumanogoh, A. Group 2 innate lymphoid cells and their surrounding
952 environment. *Inflamm Regen* **43**, 21 (2023). https://doi.org:10.1186/s41232-023-00272-
953 8

954 46 Liu, C. *et al.* Delineating spatiotemporal and hierarchical development of human fetal
955 innate lymphoid cells. *Cell Res* **31**, 1106-1122 (2021). https://doi.org:10.1038/s41422-
956 021-00529-2

957 47 Colonna, M. Innate Lymphoid Cells: Diversity, Plasticity, and Unique Functions in
958 Immunity. *Immunity* **48**, 1104-1117 (2018). https://doi.org:10.1016/j.immuni.2018.05.013

959 48 Tang, Q. *et al.* Dissecting hematopoietic and renal cell heterogeneity in adult zebrafish
960 at single-cell resolution using RNA sequencing. *J Exp Med* **214**, 2875-2887 (2017).
961 https://doi.org:10.1084/jem.20170976

962 49 Sertori, R. *et al.* Conserved IL-2Rgammac Signaling Mediates Lymphopoiesis in
963 Zebrafish. *J Immunol* **196**, 135-143 (2016). https://doi.org:10.4049/jimmunol.1403060

964 50 Xia, J. *et al.* A single-cell resolution developmental atlas of hematopoietic stem and
965 progenitor cell expansion in zebrafish. *Proc Natl Acad Sci U S A* **118** (2021).
966 https://doi.org:10.1073/pnas.2015748118

967 51 Lamason, R. L. *et al.* SLC24A5, a putative cation exchanger, affects pigmentation in
968 zebrafish and humans. *Science* **310**, 1782-1786 (2005).
969 https://doi.org:10.1126/science.1116238

970 52 Veneziani, I., Alicata, C., Moretta, L. & Maggi, E. Human toll-like receptor 8 (TLR8) in
971 NK cells: Implication for cancer immunotherapy. *Immunol Lett* **261**, 13-16 (2023).
972 https://doi.org:10.1016/j.imlet.2023.07.003

973 53 Gee, S. *et al.* The legacy of maternal SARS-CoV-2 infection on the immunology of the
974 neonate. *Nat Immunol* **22**, 1490-1502 (2021). https://doi.org:10.1038/s41590-021-
975 01049-2

976 54 Li, J. *et al.* Impaired NK cell antiviral cytokine response against influenza virus in small-
977 for-gestational-age neonates. *Cell Mol Immunol* **10**, 437-443 (2013).
978 https://doi.org:10.1038/cmi.2013.31

979 55 Veneziani, I. *et al.* Toll-like receptor 8 agonists improve NK-cell function primarily
980 targeting CD56(bright)CD16(-) subset. *J Immunother Cancer* **10** (2022).
981 https://doi.org:10.1136/jitc-2021-003385

982 56 Lopez-Cuevas, P. *et al.* Reprogramming macrophages with R848-loaded artificial
983 protocells to modulate skin and skeletal wound healing. *J Cell Sci* **137** (2024).
984 https://doi.org:10.1242/jcs.262202

985 57 Zhou, J., Nouri-Shirazi, M., Tang, H., Shen, Y. & Zeng, M. R848/TLR7-Mediated
986 Stronger CD8+ T Immunity Is Dependent on DC-NK Cell Interactions. *Int Arch Allergy*
987 *Immunol* **183**, 860-875 (2022). https://doi.org:10.1159/000522364

988 58 Robertson, T. F. *et al.* A tessellated lymphoid network provides whole-body T cell
989 surveillance in zebrafish. *Proc Natl Acad Sci U S A* **120**, e2301137120 (2023).
990 https://doi.org:10.1073/pnas.2301137120

991 59 Chen, M. J., Yokomizo, T., Zeigler, B. M., Dzierzak, E. & Speck, N. A. Runx1 is required
992 for the endothelial to haematopoietic cell transition but not thereafter. *Nature* **457**, 887-
993 891 (2009). https://doi.org:10.1038/nature07619

994 60 Jin, H. *et al.* Definitive hematopoietic stem/progenitor cells manifest distinct
995 differentiation output in the zebrafish VDA and PBI. *Development* **136**, 647-654 (2009).
996 https://doi.org:10.1242/dev.029637

997 61 Chen, M. J. *et al.* Erythroid/myeloid progenitors and hematopoietic stem cells originate
998 from distinct populations of endothelial cells. *Cell Stem Cell* **9**, 541-552 (2011).
999 https://doi.org:10.1016/j.stem.2011.10.003

1000 62 Sood, R. *et al.* Development of multilineage adult hematopoiesis in the zebrafish with a
1001 runx1 truncation mutation. *Blood* **115**, 2806-2809 (2010). https://doi.org:10.1182/blood-
1002 2009-08-236729

1003 63 Bresciani, E. *et al.* Redundant mechanisms driven independently by RUNX1 and
1004 GATA2 for hematopoietic development. *Blood Adv* **5**, 4949-4962 (2021).
1005 https://doi.org:10.1182/bloodadvances.2020003969

1006 64 Sun, J. *et al.* Clonal dynamics of native haematopoiesis. *Nature* **514**, 322-327 (2014).
1007 https://doi.org:10.1038/nature13824

1008 65 Gomez Perdiguero, E. *et al.* Tissue-resident macrophages originate from yolk-sac-
1009 derived erythro-myeloid progenitors. *Nature* **518**, 547-551 (2015).
1010 https://doi.org:10.1038/nature13989

1011 66 Torcq, L., Vivier, C., Schmutz, S., Loe-Mie, Y. & Schmidt, A. A. Single-cell and in situ
1012 spatial analyses reveal the diversity of newly born hematopoietic stem cells and of their
1013 niches. *Development* **152** (2025). https://doi.org:10.1242/dev.204454

1014 67 Ghersi, J. J. *et al.* Haematopoietic stem and progenitor cell heterogeneity is inherited
1015 from the embryonic endothelium. *Nat Cell Biol* **25**, 1135-1145 (2023).
1016 https://doi.org:10.1038/s41556-023-01187-9

1017 68 Tober, J., McGrath, K. E. & Palis, J. Primitive erythropoiesis and megakaryopoiesis in
1018 the yolk sac are independent of c-myb. *Blood* **111**, 2636-2639 (2008).
1019 https://doi.org:10.1182/blood-2007-11-124685

1020 69 Tober, J., Yzaguirre, A. D., Piwarzyk, E. & Speck, N. A. Distinct temporal requirements
1021 for Runx1 in hematopoietic progenitors and stem cells. *Development* **140**, 3765-3776
1022 (2013). https://doi.org:10.1242/dev.094961

1023 70 Dignum, T. *et al.* Multipotent progenitors and hematopoietic stem cells arise
1024 independently from hemogenic endothelium in the mouse embryo. *Cell Rep* **36**, 109675
1025 (2021). https://doi.org:10.1016/j.celrep.2021.109675

1026 71 Simic, M. *et al.* Distinct Waves from the Hemogenic Endothelium Give Rise to Layered
1027 Lymphoid Tissue Inducer Cell Ontogeny. *Cell Rep* **32**, 108004 (2020).
1028 https://doi.org:10.1016/j.celrep.2020.108004

1029 72 Elsaid, R. *et al.* A wave of bipotent T/ILC-restricted progenitors shapes the embryonic
1030 thymus microenvironment in a time-dependent manner. *Blood* **137**, 1024-1036 (2021).
1031 https://doi.org:10.1182/blood.2020006779

1032 73 Wang, X. *et al.* Innate lymphoid cells originate from fetal liver-derived tissue-resident
1033 progenitors. *Sci Immunol* **10**, eadu7962 (2025).
1034 https://doi.org:10.1126/sciimmunol.adu7962

1035 74 Lawrence, C. in *Methods in Cell Biology* Vol. 104 (eds H. William Detrich, Monte
1036 Westerfield, & Leonard I. Zon) 429-451 (Academic Press, 2011).

1037 75 Aharon, D. & Marlow, F. L. Sexual determination in zebrafish. *Cell Mol Life Sci* **79**, 8
1038 (2021). https://doi.org:10.1007/s00018-021-04066-4

1039 76 Harrold, I. *et al.* Efficient transgenesis mediated by pigmentation rescue in zebrafish.
1040 *Biotechniques* **60**, 13-20 (2016). https://doi.org:10.2144/000114368

1041 77 Jones, L. O. *et al.* Single-cell resolution of the adult zebrafish intestine under
1042 conventional conditions and in response to an acute Vibrio cholerae infection. *Cell Rep*
1043 **42**, 113407 (2023). https://doi.org:10.1016/j.celrep.2023.113407

1044 78 Jiang, M. *et al.* Characterization of the Zebrafish Cell Landscape at Single-Cell
1045 Resolution. *Front Cell Dev Biol* **9**, 743421 (2021).
1046 https://doi.org:10.3389/fcell.2021.743421

1047 79 Mahony, C. B. & Monteiro, R. Protocol for the analysis of hematopoietic lineages in the
1048 whole kidney marrow of adult zebrafish. *STAR Protoc* **5**, 102810 (2024).
1049 https://doi.org:10.1016/j.xpro.2023.102810

1050 80 Yu, G., Wang, L. G., Han, Y. & He, Q. Y. clusterProfiler: an R package for comparing
1051 biological themes among gene clusters. *OMICS* **16**, 284-287 (2012).
1052 https://doi.org:10.1089/omi.2011.0118

1053 81 Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat*
1054 *Methods* **9**, 676-682 (2012). https://doi.org:10.1038/nmeth.2019

1055 82 Danilova, N. *et al.* T cells and the thymus in developing zebrafish. *Dev Comp Immunol*
1056 **28**, 755-767 (2004). https://doi.org:10.1016/j.dci.2003.12.003

1057 83 Thisse, C. & Thisse, B. High-resolution in situ hybridization to whole-mount zebrafish
1058 embryos. *Nat Protoc* **3**, 59-69 (2008). https://doi.org:10.1038/nprot.2007.514

1059 84 Burger, A. *et al.* Maximizing mutagenesis with solubilized CRISPR-Cas9
1060 ribonucleoprotein complexes. *Development* **143**, 2025-2037 (2016).
1061 https://doi.org:10.1242/dev.134809

1062 85 Labun, K. *et al.* CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome
1063 editing. *Nucleic Acids Res* **47**, W171-W174 (2019). https://doi.org:10.1093/nar/gkz365

1064 86 Avagyan, S. *et al.* Resistance to inflammation underlies enhanced fitness in clonal
1065 hematopoiesis. *Science* **374**, 768-772 (2021). https://doi.org:10.1126/science.aba9304
1066
1067

FIGURES AND FIGURE LEGENDS

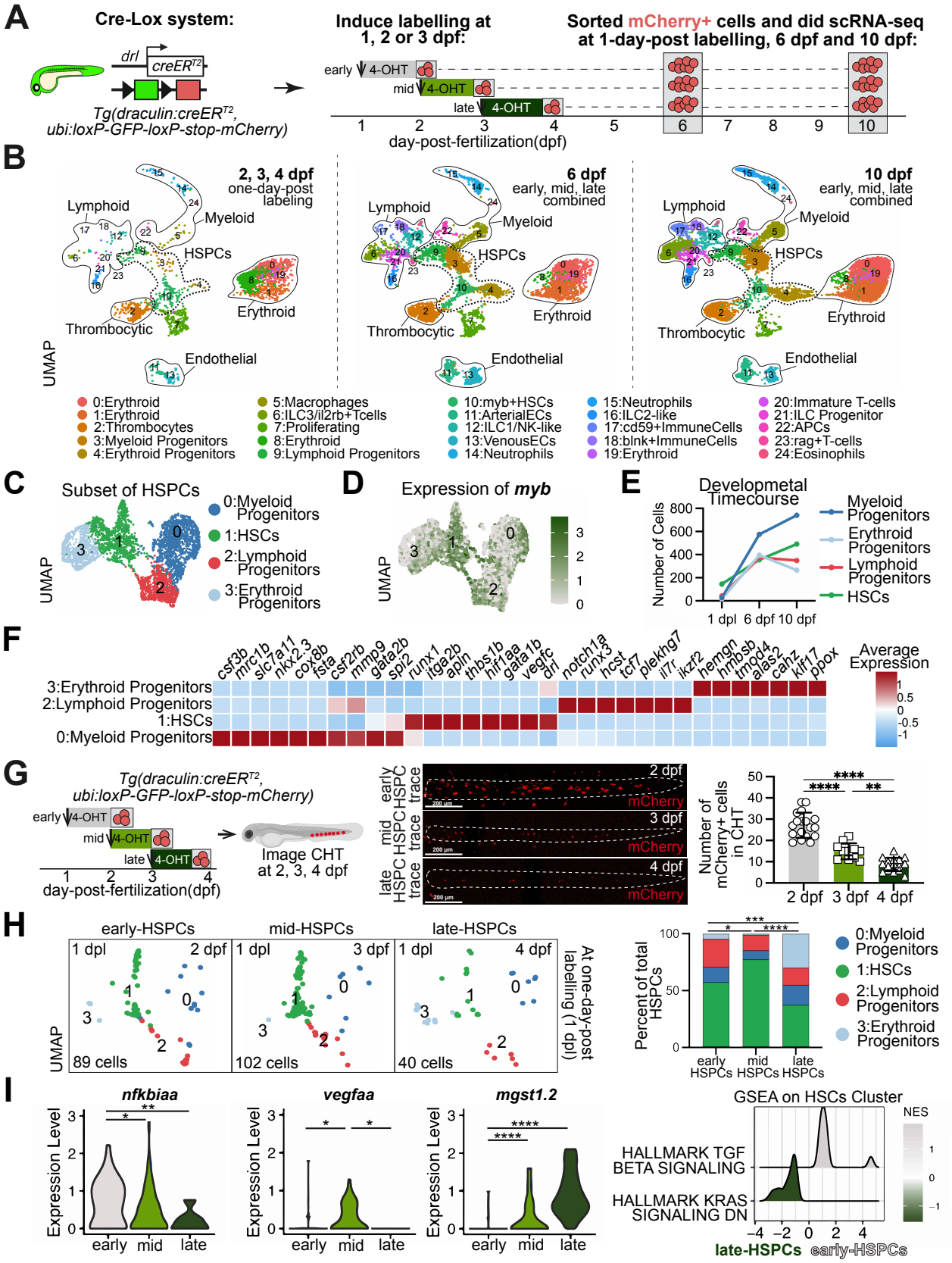


Figure 1. Lineage tracing of early, mid, and late HSPCs in zebrafish development.

1071 (**A**) Inducible Cre-Lox labelling system was activated with 12 μM hydroxy-tamoxifen (4-OHT) at
1072 either 1, 2, or 3 days-post-fertilization (dpf), termed “early-HSPCs”, “mid-HSPCs” and “late-
1073 HSPCs,” respectively. scRNA-seq of mCherry-expressing progeny of each cohort/trace was
1074 done at three developmental timepoints: one-day-post labelling (1 dpl), 6 dpf, and 10 dpf. (**B**)
1075 UMAP showing hematopoietic clusters derived from early, mid, and late HSPCs across all tested
1076 time points. (**C**) UMAP showing the HSPC subsets in all traces. (**D**) Feature plots of *myb*
1077 expression across HSPC clusters. (**E**) Line plot showing the quantification of larval HSPCs
1078 detected by scRNA-seq in all traces. (**F**) Heatmap showing select genes in the HSPC subsets
1079 from a 4-way differential expression analysis. (**G**) Confocal imaging of mCherry signal in the
1080 CHT at one-day-post labeling for each tamoxifen induced labeling timepoint. Graph showing
1081 reduced mCherry signal at 3 & 4 dpf (mid & late trace) timepoints compared to the 2 dpf (early
1082 trace) signal. Each point represents one fish. Significance computed with one-way ANOVA with
1083 Tukey multiple testing corrections. (**H**) UMAP of 1 dpl HSPCs and their relative percentages in
1084 the bar plot. (**I**) Violin plots showing the trace-selective expression of genes in cluster 10
1085 myb+HSCs. Significance computed with one-way ANOVA with Tukey multiple testing
1086 corrections. Plot on the left is GSEA analysis of the same cluster. *$p<0.05$, **$p \leq 0.01$, ****
1087 $p \leq 0.0001$
1088
1089

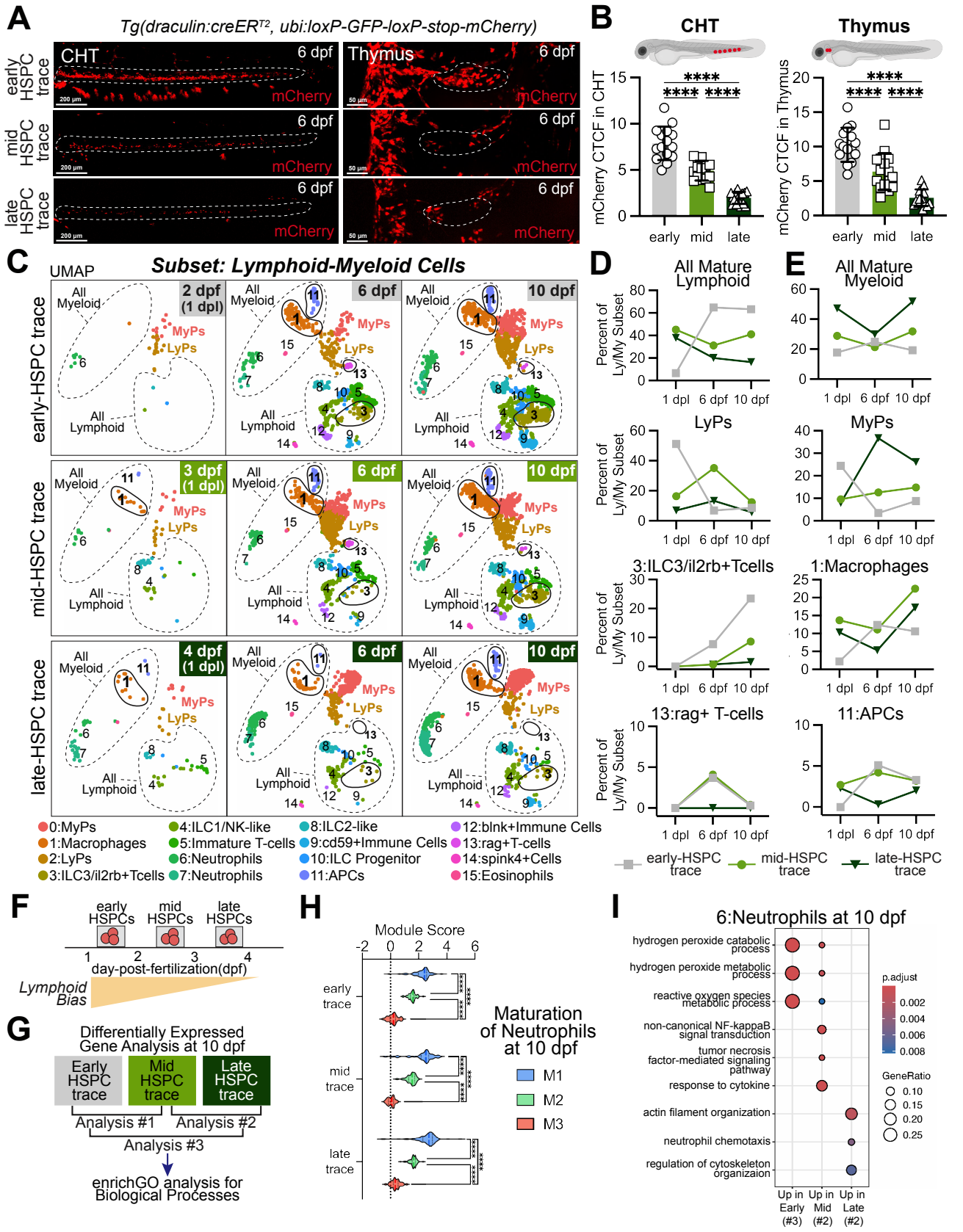
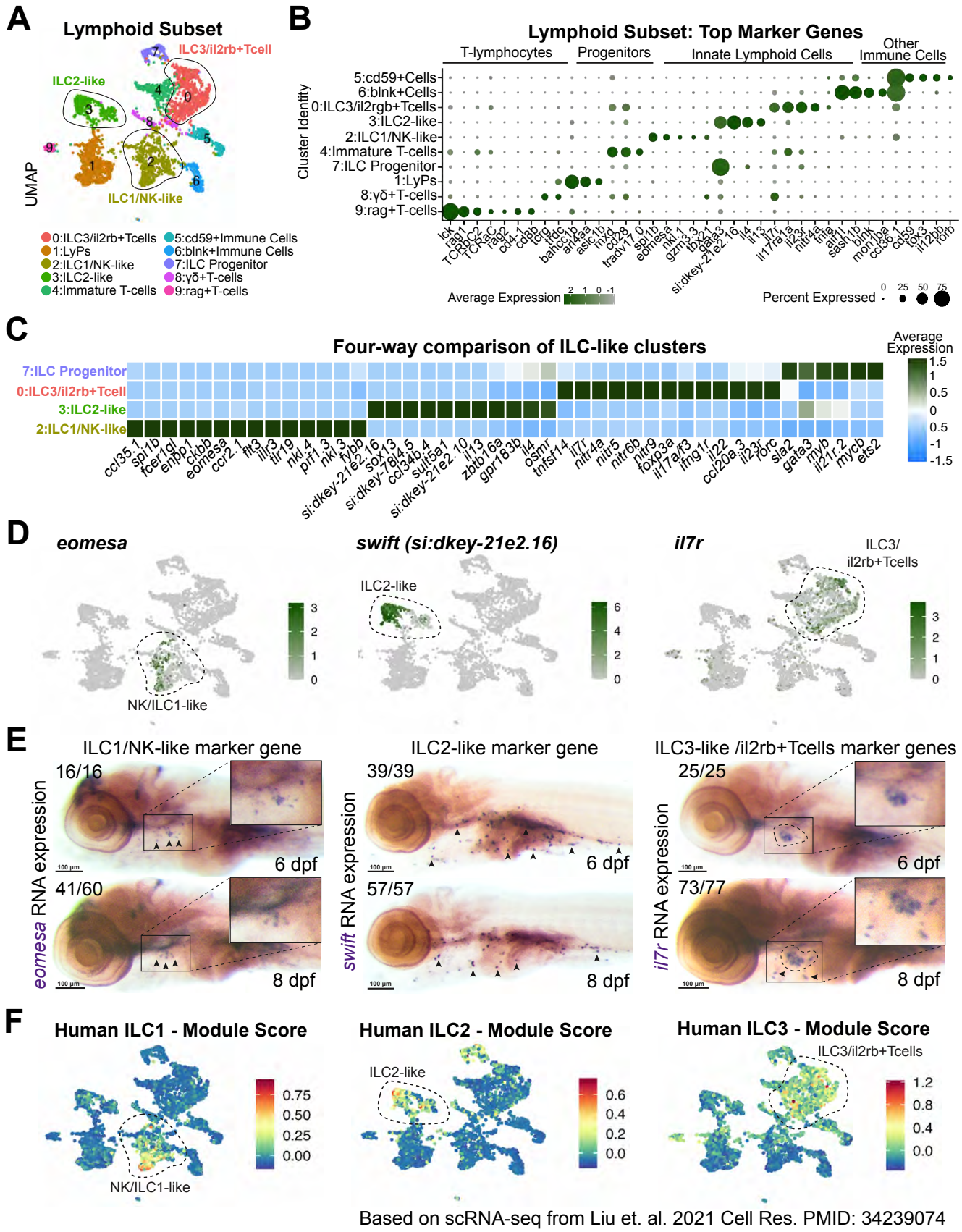


Figure 2. Differentiation of early, mid, and late HSPCs to myeloid and lymphoid lineages

1092 **during development.** (**A**) Confocal imaging of mCherry signal derived from early, mid, and late
1093 HSPCs at 6dpf with the thymus and caudal hematopoietic tissue (CHT) circled with dotted white
1094 lines. (**B**) Bar plots showing hither CHT and thymus signal from early HSPCs. Each point
1095 represents one fish. Significance computed with one-way ANOVA with Tukey multiple testing
1096 corrections. (**C**) UMAP showing the lymphoid-myeloid subset of cells, separated by timepoint
1097 and trace with dotted lines around the myeloid and lymphoid lineage cells and solid lines around
1098 clusters highlighted in D,E. (**D,E**) Plots show percentage of cells in a given cluster by trace across
1099 developmental timepoints. Lymphoid cells shown in (D) and Myeloid cells shown in (E). (**F**)
1100 Schema of the proposed model of lymphoid lineage bias established by developmental timing.
1101 (**G**) Experimental schema of differential gene expression between 10dpf cells derived from early,
1102 mid, and late HSPCs. (**H**) Module scoring of maturation stage (M1-M3) from least to most mature
1103 neutrophils across the early, mid, and late traces. Significance computed with one-way ANOVA
1104 with Tukey multiple testing corrections. (**I**) Biological Process EnrichGO pathway enrichment
1105 analysis showing pathways enriched in 10 dpf Neutrophils. **** $p \leq 0.0001$

1106



Based on scRNA-seq from Liu et. al. 2021 Cell Res. PMID: 34239074

1107

1108 **Figure 3. HSPCs differentiate to a variety of lymphoid lineage cells including ILC-like cells.**
1109 (**A**) UMAP of lymphoid lineage subset with the three types of ILCs circled. (**B**) Dot plot showing
1110 select top marker genes of each cluster. (**C**) Heatmap of a four-way differential gene expression
1111 analysis comparing the putative ILC progenitor and ILC-like clusters. (**D**) Feature plot showing
1112 expression of three genes, *eomesa, si:dkey21e2.16, il7r*, selected for *in situ* hybridization
1113 analysis of the three ILC populations. (**E**) Images of RNA *in situ* hybridization for ILC marker
1114 genes at 6 (top) and 8 (bottom) dpf. Triangles point to quantified puncta and dotted lines circle
1115 the thymus. Numbers of images represent the number of larvae with similar phenotype over the
1116 total larvae tested. (**F**) Module scoring using human fetal ILC marker genes obtained from a
1117 scRNA-seq dataset.[46]

1118

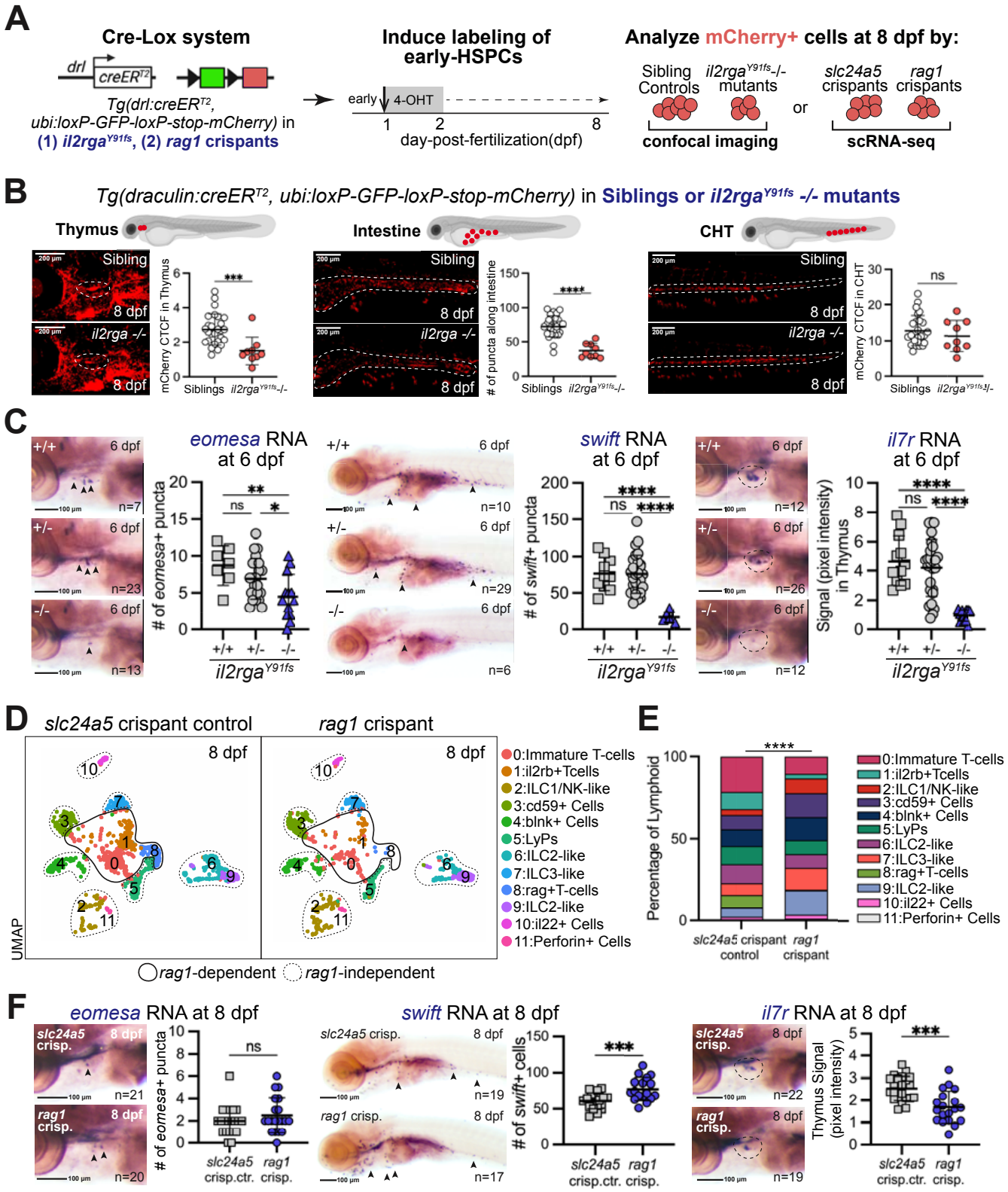


Figure 4. Zebrafish ILCs are Rag1-independent and Il2rγa-dependent. (**A**) Lineage tracing of early-HSPCs in *il2rga*^{Y91fs} mutants or in *rag1* crispants at 8dpf. The *il2rga*^{Y91fs} mutants were analyzed by confocal imaging. Crispants were subjected to scRNA-seq. The *slc24a5* (*golden*)

1123 locus was mutagenized as a control for the crispant experiment. (**B**) Confocal imaging of
1124 mCherry cells derived from early HSPCs in siblings and *il2rga*^{Y91fs} homozygous mutants at 8 dpf
1125 (left). Graph showing quantification tissue-specific mCherry signal (right). Each point represents
1126 one fish. Significance computed with unpaired t-test. (**C**) Images of RNA *in situ* hybridization of
1127 *eomesa*, *swift*, and *il7r* (left). Graphs showing significant reduction of ILC markers in *il2rga*^{Y91fs}
1128 homozygous mutants. Triangles point to quantified puncta and dotted lines circle the thymus.
1129 Each point represents one fish. Significance computed with one-way ANOVA with Tukey multiple
1130 testing corrections. (**D**) UMAP showing the lymphoid lineage subset of clusters in controls and
1131 *rag1* crispants with solid lines circling the *rag1*-dependent clusters and dotted lines circling *rag1*-
1132 independent clusters. (**E**) Stacked bar plot showing percentage of cells in each cluster relative
1133 to total lymphoid cells. Significance computed with chi-square test. See supplemental for
1134 absolute numbers. (**F**) RNA *in situ* hybridization for ILC marker genes at 8 dpf in controls and
1135 *rag1* crispants (left). Graphs of quantification of *eomesa, swift,* and *il7r* expression in the *rag1*
1136 crispants (right). Each point represents one fish. Significance computed with unpaired t-test.
1137 **$p \leq 0.01$, ***$p \leq 0.001$, **** $p \leq 0.0001$

1138

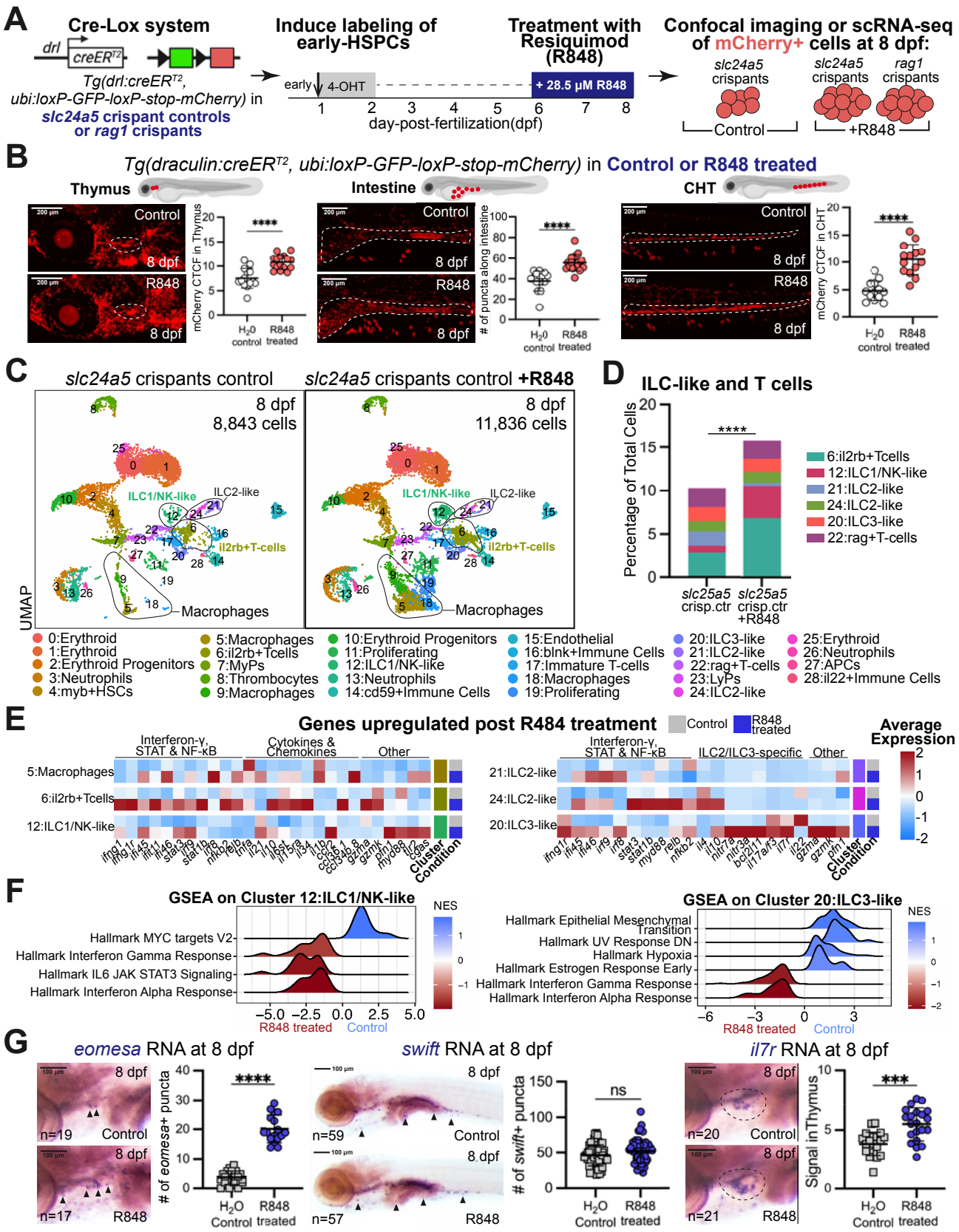


Figure 5. Viral mimicry elicits a response from the ILC-like cells in zebrafish larvae. (**A**)

1141 Early-HSPC trace-derived mCherry+ cells from controls and *rag1* crispants from either vehicle
1142 control or R848 treated groups were sequenced at 8 dpf. (**B**) Confocal imaging of mCherry+
1143 cells derived from early HSPCs from 8 dpf controls and R848-treated larvae (left). Graphs
1144 depicting an increase in mCherry+ signal in the thymus, intestine and CHT of R848-treated
1145 larvae (right). Each point represents one fish. Significance computed with unpaired t-test. (**C**)
1146 UMAP showing all hematopoietic clusters detected at 8 dpf per condition with lines circling
1147 Macrophages, ILC1/NK-like, ILC2-like, and T-cells clusters that show the greatest response to
1148 the R848 agonist. (**D**) Bar plots showing the relative percentages of ILC/T-cell clusters in control
1149 and R848 treated groups out of all the cells shown in (C), Significance computed with chi-square
1150 test. See supplemental figure for absolute numbers. (**E**) Heatmaps showing average expression
1151 of select genes across the designated clusters in control and R848-treated groups. (**F**) GSEA
1152 analysis showing upregulation of interferon response genes in ILC clusters 12 and 20 in R848
1153 treated larvae. (**G**) RNA *in situ* hybridization of *eomesa*, *swift,* and *il7r* with *eomesa* and *il7r*
1154 significantly increased in response to R848. Triangles point to quantified puncta and dotted lines
1155 circle the thymus. Each point represents one fish. Significance computed with unpaired t-test.
1156 ***$p \leq 0.001$, **** $p \leq 0.0001$

1157

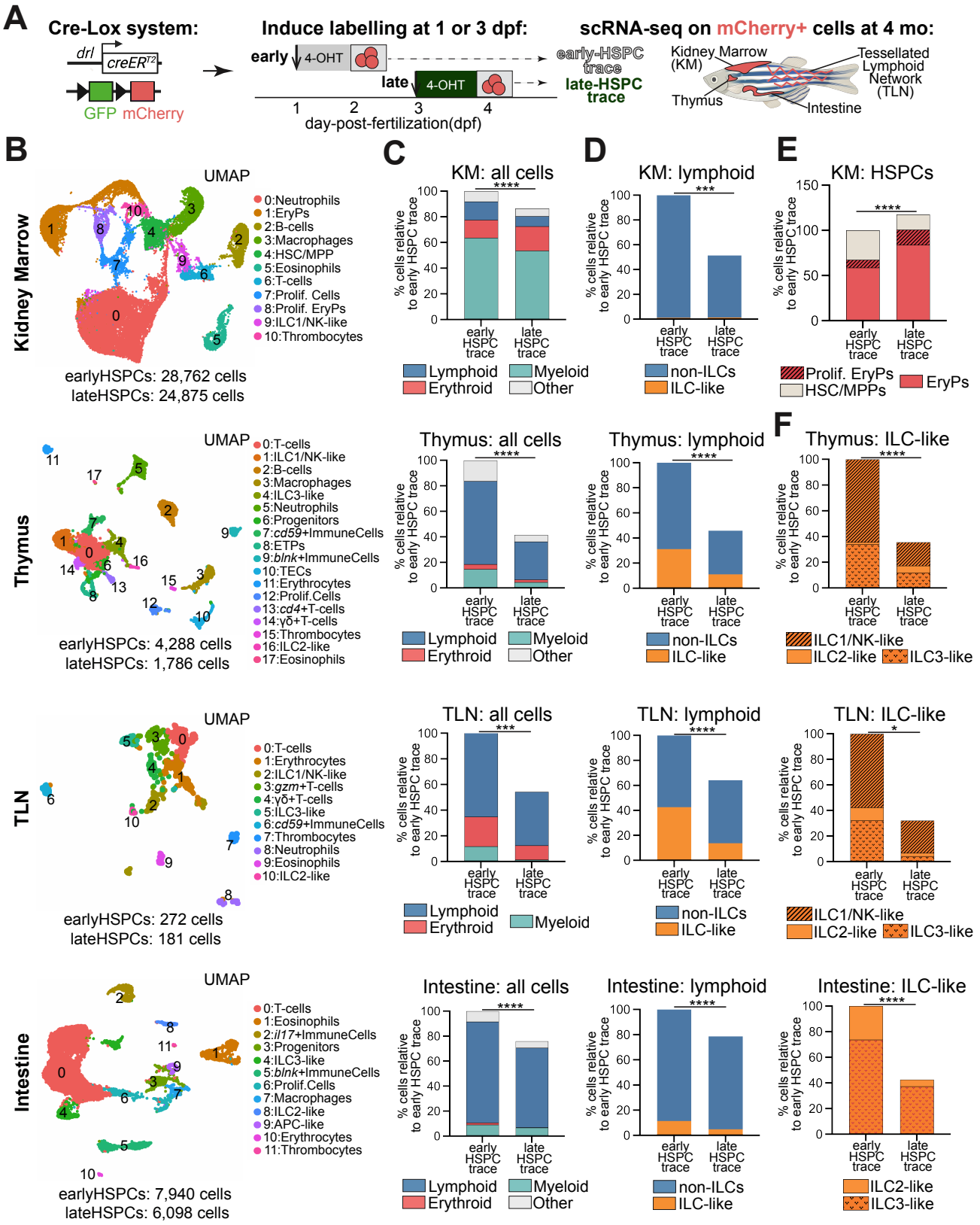
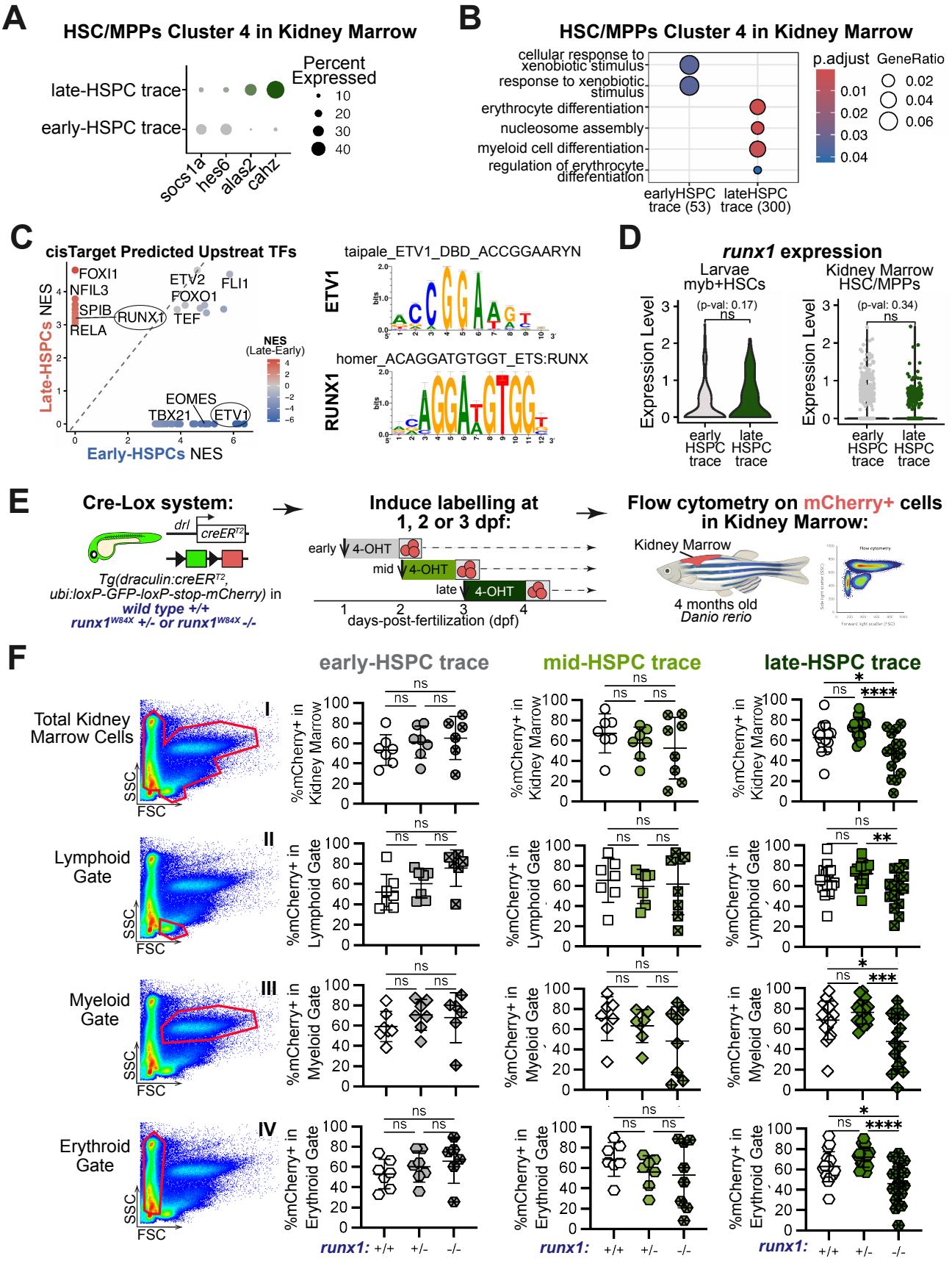


Figure 6. Early-arising HSPCs are the main contributors to adult tissue-resident

1160 **lymphocytes.** (**A**) Schema of early or late HSPC lineage tracing induced during embryonic
1161 development using the Cre-Lox inducible system followed by scRNA-seq of mCherry+ cells from
1162 kidney marrow, thymus, intestine and TLN isolated from 4 month adult zebrafish. (**B**) UMAPs
1163 showing cluster annotations per adult tissue with both traces combined. (**C**) Graphs showing
1164 percentage of hematopoietic cells across tissues relative to the early-HSPC trace. Significance
1165 computed with chi-square test. (**D**) Graphs showing percentage of lymphoid lineage cells across
1166 tissues relative to the early-HSPC trace, binned by ILC-like versus all other non-ILC
1167 lymphocytes. Significance computed with chi-square test. (**E**) Graph depicting percentage of
1168 HSCs and EryPs in kidney marrow relative to the early-HSPC trace. Significance computed with
1169 chi-square test. (**F**) Graphs depicting percentage of ILC-like cells across tissues relative to the
1170 early-HSPC trace. Significance computed with chi-square test. *$p<0.05$, ***$p \leq 0.001$,
1171 ****$p \leq 0.0001$
1172
1173


A
HSC/MPPs Cluster 4 in Kidney Marrow
Percent Expressed
10
20
30
40
late-HSPC trace
early-HSPC trace
socs1a
hes6
alas2
cahz
B
HSC/MPPs Cluster 4 in Kidney Marrow
cellular response to xenobiotic stimulus
response to xenobiotic stimulus
erythrocyte differentiation
nucleosome assembly
myeloid cell differentiation
regulation of erythrocyte differentiation
earlyHSPC trace (53)
lateHSPC trace (300)
p.adjust
0.01
0.02
0.03
0.04
GeneRatio
0.02
0.04
0.06
C
cisTarget Predicted Upstreat TFs
FOXI1
NFIL3
SPIB
RELA
RUNX1
ETV2
FLI1
FOXO1
TEF
EOMES
TBX21
ETV1
Late-HSPCs NES
Early-HSPCs NES
NES (Late-Early)
taipale_ETV1_DBD_ACCGGAARYN
ETV1
homer_ACAGGATGTGGT_ETS:RUNX
RUNX1
D
runx1 expression
Larvae myb+HSCs
Kidney Marrow HSC/MPPs
(p-val: 0.17)
(p-val: 0.34)
ns
Expression Level
early HSPC trace
late HSPC trace
E
Cre-Lox system:
drl
creER^T2
Tg(draculin:creER^T2, ubi:loxP-GFP-loxP-stop-mCherry) in
wild type +/+
runx1^W84X +/- or runx1^W84X -/-
Induce labelling at 1, 2 or 3 dpf:
early 4-OHT
mid 4-OHT
late 4-OHT
days-post-fertilization (dpf)
Flow cytometry on mCherry+ cells in Kidney Marrow:
Kidney Marrow
4 months old
Danio rerio
F
early-HSPC trace
mid-HSPC trace
late-HSPC trace
Total Kidney Marrow Cells
Lymphoid Gate
Myeloid Gate
Erythroid Gate
SSC
FSC
%mCherry+ in Kidney Marrow
%mCherry+ in Lymphoid Gate
%mCherry+ in Myeloid Gate
%mCherry+ in Erythroid Gate
ns
runx1: +/+ +/- -/-

1174

1175 **Figure 7. Early and late HSPCs show unique sensitivities to the transcription factor**
1176 **Runx1.** (**A**) Dot plot showing the expression of select lymphoid and erythroid lineage genes in
1177 early and late HSPC trace-derived kidney marrow cluster 4 HSC/MPPs. (**B**) EnrichGO Biological
1178 Processes pathway enrichment analysis comparing kidney marrow HSC/MPPs between the two
1179 traces. (**C**) Scatter plot showing the predicted upstream transcription factors regulating early and
1180 late HSPCs based on upstream motif analysis of genes differentially expressed between early
1181 and late trace-derived kidney marrow cluster 4 HSC/MPP. The RUNX1 (late) and ETV1 (early)
1182 motifs are highlighted as examples of putative trace-specific upstream regulators. (**D**) Violin plots
1183 showing the expression of *runx1* in larval and adult HSC clusters in early and late HSPC traces.
1184 Significance computed with a Wilcoxon test. (**E**) Lineage tracing in wild-type, heterozygous and
1185 homozygous *runx1*^{W84X} mutants following embryonic early, mid, and late HSPC labeling. Kidney
1186 marrow cells from 4-month-old zebrafish were analyzed for percentage of early, mid, or late
1187 trace-derived mCherry+ cells by flow cytometry. (**F**) Flow plots showing the lineage(s) examined:
1188 I- total kidney marrow; II- Lymphoid gate; III- Myeloid gate; IV- Erythroid gate. The full flow
1189 strategy is shown in Figure S7. Graphs quantifying the percentage of mCherry+ cells from the
1190 three traces across *runx1* genotypes. Each point represents one fish. Significance computed
1191 with one-way ANOVA with Tukey multiple testing corrections. *$p<0.05$, **$p \leq 0.01$, ***$p \leq 0.001$,
1192 **** $p \leq 0.0001$
1193