

Metabolically Primed Multipotent Hematopoietic Progenitors Fuel Innate Immunity

Jason Cosgrove¹, Anne-Marie Lyne¹, Ildefonso Rodriguez², Vincent Cabeli¹, Cecile Conrad¹, Sabrina Tenreira-Bento¹, Emilie Tubeuf¹, Erica Russo⁴, Fanny Tabarin⁵, Yannis Belloucif³, Shayda Maleki-Toyserkani⁶, Sophie Reed⁶, Federica Monaco⁶, Ann Ager⁶, Camille Lobry³, Philippe Bousoo⁴, Pablo Jose Fernández-Marcos², Herve Isambert¹, Rafael J. Argüello⁷, Leïla Perié^{1,*}

1. Institut Curie, Université PSL, Sorbonne Université, CNRS UMR168, Laboratoire Physico Chimie Curie, Paris, France
2. Metabolic Syndrome Group – BIOPROMET, Madrid Institute for Advanced Studies – IMDEA, Madrid, Spain
3. Inserm U944, CNRS UMR7212, Institut de Recherche Saint Louis and Université Paris-Cité, 75010 Paris, France
4. Dynamics of Immune Responses Unit, Institut Pasteur, Université Paris Cité, INSERM U1223, 75015 Paris, France
5. BMBC platform, Institut Curie, Université PSL, Sorbonne Université, CNRS UMR168, Laboratoire Physico Chimie Curie, Paris, France
6. Systems Immunity University Research Institute, School of Medicine, Cardiff University, Cardiff, UK
7. Aix Marseille Univ, CNRS, INSERM, CIML, Centre d'Immunologie de Marseille-Luminy, Marseille, France

*Contact information: Leila.Perie@curie.fr

SUMMARY

Following infection, hematopoietic stem and progenitor cells (HSPCs) support immunity by increasing the rate of innate immune cell production but the metabolic cues that guide this process are unknown. To address this question, we developed MetaFate, a method to trace the metabolic expression state and developmental fate of single cells *in vivo*. Using MetaFate we identified a gene expression program of metabolic enzymes and transporters that confers differences in myeloid differentiation potential in a subset of HSPCs that express CD62L. Using single-cell metabolic profiling, we confirmed that CD62L^{high} myeloid-biased HSPCs have an increased dependency on oxidative phosphorylation and glucose metabolism. Importantly, metabolism actively regulates immune-cell production, with overexpression of the glucose-6-phosphate dehydrogenase enzyme of the pentose phosphate pathway skewing MPP output from B-lymphocytes towards the myeloid lineages, and expansion of CD62L^{high} HSPCs occurring to support emergency myelopoiesis. Collectively, our data reveal the metabolic cues that instruct innate immune cell development, highlighting a key role for the pentose phosphate pathway. More broadly, our results show that HSPC metabolism can be manipulated to alter the cellular composition of the immune system.

Introduction

Throughout the body, stem cells need to constantly adapt the amount and type of cells that they produce to maintain tissue homeostasis, compensate for cell loss and promote tissue repair. A key challenge in the stem cell field is to understand the molecular signals that selectively differentiate stem cells into specialized cell types *in vivo*. Much emphasis has been placed on the role of transcription and growth factors in guiding lineage specification, but the role of metabolism in this context has been overlooked. Under homeostasis, hematopoiesis is the dominant biosynthetic process in the human body, generating 10¹¹ cells each day^{1,2} and accounting for ~86% of daily cell turnover¹. Following infection, hematopoietic stem and progenitor cells (HSPCs) rapidly change the amount and the types of cells that they produce to support the immune system³. Despite the significant biosynthetic demands placed on HSPCs, the metabolic cues that regulate the magnitude and lineage-specificity of their cellular outputs are poorly understood.

Recently developed population-level metabolomics, genetics, and pharmacological approaches show that cellular metabolism is a key regulator of hematopoietic stem cell (HSC) function^{4,5}. At the top of the hematopoietic hierarchy, HSCs have been characterized by high rates of glycolysis⁶ and low rates of protein

49 synthesis⁷, while glutamine⁸, mitophagy and fatty acid oxidation^{9,10}, vitamin A¹¹, ascorbate¹² and aspartate¹³
50 have been shown to regulate HSC erythroid commitment, renewal, dormancy, abundance, and reconstitution
51 capacity respectively. As in many other stem cell systems, much work has focused on the metabolic regulation
52 of stemness and quiescence, with relatively little focus given to the metabolic cues that guide downstream
53 lineage specification^{5,14,15}. Consequently, the metabolic processes needed to differentiate HSCs into immune
54 cells are unknown. In addition, fate-mapping and cellular barcoding studies have shown that the downstream
55 multipotent progenitor (MPP) compartment acts as the major source of new blood cells in native
56 haematopoiesis¹⁶, and is where lineage branchpoints occur^{17,18}. Despite their functional importance, it is
57 unclear to what extent metabolism can shape the magnitude and lineage specificity of immune cell production
58 from MPPs.

59
60 The lack of metabolomics studies in HSPCs is due in part to the technical challenges associated with measuring
61 metabolic processes in rare cell types. Metabolites are typically short lived - on the order of minutes¹⁹ and
62 have a large structural diversity, limiting state of the art mass-spectrometry based assays to 10⁴ HSPCs^{12,20}. Such
63 limitations in sensitivity make it difficult to link relative metabolite measurements of bulk populations to the
64 functional heterogeneity of individual HSPCs, as characterised by lineage tracing and single cell transplantation
65 studies^{17,18,21,22}. Recent advances in high-dimensional mass cytometry²³, flow cytometric profiling of
66 translation²⁴, genetically encoded biosensors²⁵, and *in situ* dehydrogenase assays²⁶ are helping to address this
67 challenge. However, these techniques are typically destructive in nature, making it challenging to link metabolic
68 state to functional outcomes, particularly *in vivo*. This limitation is critical, with recent studies showing that -
69 omics profiling should be paired with functional measurements to resolve HSPC heterogeneity^{27,28}.

70
71 To address these challenges, we developed MetaFate, an *in situ* barcoding approach to combine the metabolic
72 gene expression state and differentiation fate of single cells *in vivo*. RNA expression has a complex association
73 with metabolite levels, due in part to non-linear enzyme kinetics and metabolites being processed by multiple
74 pathways²⁹. RNA measurements do however provide information about the expression patterns of metabolic
75 enzymes and transporters, and can be used to identify novel surface markers to purify functionally distinct cell-
76 subsets for downstream metabolomics profiling. Here, using MetaFate profiling of HSPCs, we identified a gene
77 expression program of metabolic enzymes/transporters that are associated with myeloid differentiation
78 potential and expression of the adhesion molecule CD62L. Fluorescence based metabolic profiling assays
79 corroborated MetaFate gene expression patterns, revealing a higher dependency on OXPHOS and glucose
80 metabolism to fuel increased rates of protein synthesis and ATP turnover in CD62L^{high} myeloid-biased MPPs. In
81 addition, we demonstrate that metabolism plays an active role in regulating immune cell production. MetaFate
82 identified the pentose phosphate pathway as a metabolic signature of myeloid development. Overexpression
83 of the glucose-6-phosphate dehydrogenase enzyme, rate limiting enzyme of the pentose phosphate pathway,
84 limited the production of B-cells from transplanted MPPs, skewing output towards the erythromyeloid lineages.
85 In the context of emergency myelopoiesis, the CD62L^{high} MPP compartment expands to meet increased
86 demands for innate immune cells.

87 Collectively, our data bridges understanding between the fields of stem cell biology, cellular metabolism and
88 immunology, revealing the metabolic cues that guide early innate immune cell development. Our results
89 highlight a key role for the pentose phosphate pathway in this process and show that by manipulating lineage
90 specific metabolic cues, it is possible to alter the specificity of regenerative processes *in vivo*.

91 **Results**

92 93 94 **MetaFate – a lineage tracing approach to obtain fate resolved RNA expression patterns of metabolic enzymes** 95 **and transporters**

96

97 In this study, we hypothesise that metabolism regulates the differentiation of HSPCs *in vivo*. To address this
98 question, we developed MetaFate, an approach that combines single cells transcriptomics with *in-situ*
99 barcoding to provide fate-resolved expression patterns of metabolic enzymes and transporters (**Figure 1a**). For
100 a single progenitor cell MetaFate provides 3 pieces of information: (i) gene expression data (ii) a lineage barcode
101 (iii) the frequency at which its lineage barcode is found across differentiated cell types. Collectively, this
102 information can be used to link metabolic enzyme/transporter RNA expression to differentiation behaviours in
103 single cells *in vivo* (**Figure 1a**).

104
105 To barcode cells in their native environment, we use the DRAG (**D**iversity through **R**AG) *in situ* barcoding
106 technology that allows inducible labelling of cellular lineages with heritable barcodes (**Figure 1a, Figure S1a**)
107 (*Urbanus and Cosgrove et al, under revision, manuscript provided within this submission*). Importantly, DRAG
108 barcoding is neutral with respect to hematopoietic differentiation, has a high barcode diversity, and can
109 quantify clonal output in low cell numbers (*Urbanus and Cosgrove et al*). These key attributes make DRAG
110 barcoding well-suited to studying the clonal dynamics of HSPCs *in vivo*. In brief, upon CRE induction by
111 tamoxifen, the cassette between two loxP sites is inverted, causing the expression of both the RAG1 and 2
112 enzymes and Terminal deoxynucleotidyl transferase (TdT). This leads to the generation of a heritable barcode
113 through the recombination of the synthetic V-, D- and J-segments, with barcode diversity being generated both
114 by RAG-mediated nucleotide deletion and TdT-mediated N-addition (**Figure S1a**). In addition, DRAG
115 recombination results in GFP expression, facilitating the purification of barcoded cells by fluorescence activated
116 cell sorting (FACS). To detect barcodes in large populations of differentiated cells, we use targeted amplification
117 of the invariant region common to all barcodes and deep sequencing of genomic DNA at the population level,
118 as in our original protocol (*Urbanus and Cosgrove et al, in revision*). To detect barcodes and gene expression
119 information in progenitor cells, we developed a custom targeted amplification approach using primers targeted
120 to the invariant region of DRAG barcodes to recover barcode transcripts from 10X genomics 3' scRNAseq
121 libraries (*supplementary information*).

122
123 To test the metafate experimental and bioinformatics pipeline, RosaCreERT2^{+/-} DRAG^{+/-} mice were given
124 tamoxifen injections to induce barcode recombination at 8-11 weeks of age. 47-67 weeks post-induction,
125 barcoded HSPCs (Sca1⁺ cKit⁺ GFP⁺), Cd11b⁺ GFP⁺ Myeloid and Ter119⁺ CD44⁺ GFP⁺ nucleated erythroid cells were
126 isolated from the bone marrow of 5 mice using fluorescence activated cell sorting (FACS) (**Figure 1b; Figure S2**).
127 HSPCs were then processed for single cell RNA sequencing and targeted barcode amplification (materials and
128 methods). In nucleated erythroid and mature myeloid cells, barcodes were detected from cell populations at
129 the DNA level as described in Urbanus and Cosgrove et al. Following data pre-processing, integration and
130 quality control (**Figure S4a-c**) transcriptomes for 4,485 cells were retained post quality control filtering and a
131 median of 3,812 genes were detected per cell. Genes which mapped to enzymes and transporters of metabolic
132 pathways from the KEGG, GO and REACTOME database were classified as metabolically-associated for
133 downstream analyses (3095 genes). Following QC and filtering (materials and methods), we recovered RNA
134 barcode information for 668 hematopoietic stem and progenitor cells (14.9% recovery rate at the RNA level),
135 corresponding to 158 unique lineage barcodes (**table S1**). From the mature erythroid and myeloid bone marrow
136 compartment, we recovered 381 unique barcodes at the DNA level with high consistency between technical
137 replicates (**Figure S3**), with 97 barcodes overlapping between RNA and DNA detection. Comparison of lineage
138 barcodes detected from either DNA or RNA showed similar barcode lengths, as well as similar insertion and
139 deletion patterns, confirming that full barcode sequences could be accurately recovered from transcripts
140 (**Figure S1b-d**). To compute the probability that two independent cells were labelled with the same barcode,
141 we applied a mathematical model of DRAG barcode recombination (*Urbanus and Cosgrove et al, materials and*
142 *methods*) to infer the generation probability of each barcode. Most DNA and RNA detected barcodes had a low
143 generation probability (**Figure S1e**), with barcodes present in several mice having a higher probability than
144 barcodes present in one mice (**Figure S1f**). Therefore many barcodes with low probability to label several cells
145 both detected in RNA and DNA were available for lineage analysis. Collectively, these analyses demonstrate
146 that MetaFate permits flexibility in barcode recovery at the RNA or DNA level, a comprehensive bioinformatics
147 framework to filter spurious and probable barcodes, as well as providing a high degree of barcode diversity *in*

148 *in vivo*. In summary, MetaFate is a robust method to integrate *in situ* barcoding and single cell transcriptomics
149 measurements in rare cell types.

150
151 Once we had validated the MetaFate pipeline, we sought to study the metabolic regulation of lineage
152 commitment dynamics in HSPCs. By analysing the distribution of barcodes across the myeloid and erythroid
153 lineages, we observed that HSPCs are highly heterogeneous in the amount (**Figure 1c**) and type (**Figure 1d**) of
154 cells that they produced. To further characterize the functional heterogeneity of HSPCs, barcode labelled HSPCs
155 were classified as differentiation inactive (98 cells ; 61 unique barcodes) if we could not detect their barcode in
156 any mature cell compartments^{22,27,30,31}, or as erythroid-biased (143 cells ; 35 barcodes), myeloid-biased (143
157 cells ; 31 barcodes), or unbiased (284 cells ; 31 unique barcodes) depending on the relative abundance of the
158 barcode across the respective lineages. Specifically, barcodes that had more than 75% of its barcode reads in
159 the myeloid or erythroid lineage were classified as lineage-biased (**Figure 1d-e**), or otherwise classified as
160 unbiased. Similar results were obtained with thresholds close to 75% (**Figure S4e-g**). Extreme thresholds (90%
161 or below 55%) impacted the amount of barcodes and the magnitude of difference in gene expression,
162 precluding robust analysis (**Figure S4e-g**). Note that the barcode generation probability was low for most
163 barcodes, indicating that the labeling of multiple initial cells is not accounting for the classification of the
164 barcode in the different categories (**Figure S1g**). Within our UMAP representation of the data, the distribution
165 of differentiation inactive clones correlated with signatures of dormant HSCs³², while we observed a significant
166 overlap in the distribution of erythroid and myeloid-biased clones within MPP-associated regions of UMAP
167 space that was not resolved by unsupervised clustering on gene expression alone (**Figure S4d**) or mapped to
168 an existing known MPP subset (**Figure 1e-f**), suggesting that metafate revealed new biased MPP subsets to be
169 characterized. Differential expression analysis between myeloid-biased, erythroid-biased and differentiation
170 inactive clones identified a total of 464 differentially expressed genes associated with myeloid bias, 271 of
171 which were upregulated in myeloid-biased clones, defining the DRAG-Fate myeloid gene signature (Figure 1e-
172 f). Among these genes were existing markers of myeloid potential including *Mpo*, *Ctsg*, *Ms4a3* and *Cpa3*^{27,33,34},
173 confirming that metafate can identified myeloid biased cells. Interestingly, 57/271 genes within the DRAG-Fate
174 myeloid signature encoded enzymes and transporters from metabolic pathways of the KEGG, REACTOME and
175 GO reference databases (**Figure 1g-h**). This subset of metabolically-associated genes, hereafter called the
176 MetaFate myeloid signature (**Figure 1g-h**), comprises genes relating to OXPHOS (*Idh2*, *Idh3a*, *Cox7b*, *Ndufa4*,
177 *Uqcr10*), proteostasis and ribosome biogenesis (*Hdc*, *Kyat3*, *Sec61b*, *Slc35b1*, *Psmc4*), the pentose phosphate
178 pathway (*Tkt*, *Taldo1*, *Gpi1*, *Pgls*) as well as genes relating to the regulation of redox state (*Gsto1*, *Mgst2*,
179 *Txn2*, *Txndc11*, *Gpx1*) (**Figure 1g-h**). In summary, by combining analysis of the transcriptome and the lineage
180 barcode in the bone marrow, metafate identified a new subset of myeloid-biased MPP that upregulate specific
181 metabolic-associated genes. These results suggest that a metabolic program associated to a lineage bias is
182 active very early in differentiation.

183
184 Within HSPCs expression of the MetaFate and DRAGFate -myeloid gene signatures are highly correlated (**Figure**
185 **S5a**), suggesting that that enzyme/transporter expression state alone could be used to predict myeloid fate in
186 HSPCs. To compare the predictive power of metabolic-associated genes from the MetaFate myeloid signature
187 against other signatures, we computed the Spearman's correlation coefficient between gene signature
188 expression scores and the myeloid bias score of HSPCs (**Figure 1i**). The DRAG fate signature consistently
189 outperformed the MetaFate myeloid signature, suggesting that metabolism is not the only program
190 contributing to myeloid bias (**Figure 1h**). However the MetaFate myeloid signature predicted myeloid bias to a
191 greater extent that gene sets relating to transcription factor activity (**Figure 1i**), which are established regulators
192 of fate choice, highlighting the importance of metabolic regulation in cell fate decisions. Furthermore, in
193 comparison to known signatures of myeloid bias in HSPCs, the MetaFate signature had a higher correlation (ρ
194 = 0.24, p-value = 2.5×10^{-10}) with myeloid bias compared to the existing MPP3 signature³⁴ (289 genes) (ρ =
195 0.2, p-value = 1×10^{-7}) (**Figure 1i**). DRAG-barcode derived signatures also outperformed the MPP3 signature of
196 myeloid bias when we performed 4-fold cross-validation analysis, to assess the sensitivity of our result to
197 overfitting (**Figure S4h**), revealing the power of combined barcoding and transcriptome in the same cells to
198 identify lineage bias subsets. To assess the broader predictive power of the MetaFate-myeloid signature, we

199 quantified its expression across 3 independent published scRNAseq datasets of hematopoietic progenitors³⁵⁻
200 ³⁷. In these datasets, the MetaFate-myeloid signature was upregulated in myeloid progenitors relative to other
201 progenitor subsets (**Figure S5**). Taken together these analyses showed that the MetaFate myeloid signature,
202 comprising only genes associated with metabolism, was a robust predictor of myeloid differentiation potential
203 of MPPs in native hematopoiesis.

204
205 To assess whether the MetaFate myeloid expression program was maintained throughout development or was
206 transiently expressed in Lin⁻ Sca1⁺ cKit⁺ HSPCs, we assessed enzyme and transporter gene expression patterns
207 across different phases of myeloid development. We modelled early stages of myeloid differentiation (HSC ->
208 MPP -> cKit⁺ restricted potential progenitors) by applying the PAGA algorithm³⁸ to a published scRNAseq
209 dataset of cKit⁺ progenitors³⁶ (**Figure S6**). Using this developmental trajectory inference approach, we found
210 that the MetaFate metabolic program was not expressed in HSCs but was heterogeneously expressed within
211 the MPP compartment and increased as cells transition from the MPP to the cKit⁺ Sca1⁻ myeloid committed
212 progenitor compartments. This data is consistent with previous single cell studies showing that erythroid-
213 myeloid branching can occur before the common myeloid progenitor compartment within MPPs^{18,39}

214
215 In summary, MetaFate is the first approach to map metabolic gene expression states to developmental fate in
216 single cells in vivo. By combining expression and fate analysis, MetaFate showed that MPP's are metabolically
217 heterogeneous and that this heterogeneity confers differences in lineage potential. Metafate revealed a
218 metabolic-associated gene signature that starts to be expressed in MPPs following the exit of quiescence and
219 entry into myeloid development. This early expressed metabolic-associated program is a robust predictor of
220 myeloid potential in MPP and is reinforced upon lineage commitment and maturation. Together this suggests
221 that the metabolic regulation of fate decisions can occur in the earliest phases of hematopoietic development,
222 within multipotent progenitors. To build upon this result, we next assessed to what extent the metabolic-
223 associated gene expression patterns observed with Metafate are reflective of metabolic pathway activity in
224 MPPs.
225

226 **CD62L^{high} Multipotent Progenitors Are Characterised by a Reduced ATP/ADP Ratio and Higher Rates of** 227 **Protein Synthesis and Oxidative Phosphorylation**

228 Given that MetaFate identified a novel myeloid-biased MPP subset with a distinct expression program of
229 enzymes and transporters, we developed a purification strategy to isolate this subset such that we could further
230 assess their metabolic and functional properties. Differential expression analysis between barcoded HSPCs
231 (**Figure 1g**) highlighted *Sell*, the gene encoding the adhesion molecule CD62L, as a putative marker of cells
232 expressing the MetaFate-myeloid expression program. We also observed significant differences in *Sell* between
233 MetaFate^{low} and MetaFate^{high} cells (cells below and above the 75th percentile of MetaFate signature expression
234 respectively, $p < 0.001$) (**Figure 2a**). Flow cytometry profiling showed that CD62L is heterogeneously expressed
235 in HSPCs (**Figure 2b-c**), with high expression in a subset of MPP3 and MPP4 cells, and little to no expression in
236 LT-HSCs, ST-HSCs and MPP2s (**Figure 2b-c**), consistent with our MetaFate analyses. Based on these results, we
237 selected CD62L as a marker of MetaFate^{high} myeloid-biased MPPs (**Figure 2d**).

238
239 To measure the metabolic pathway activity of CD62L⁺ MPPs, we combined two complementary fluorescence
240 based assays (**Figure 2e,f**): (i) SCENITH (Single Cell Metabolism by Profiling Translation inhibition), a flow
241 cytometry-based method²⁴ based on profiling protein synthesis rates in response to metabolic inhibitors using
242 flow cytometry (ii) SPICE-Met which provides a measure of cellular ATP:ADP ratio using a genetically encoded
243 PercevalHR biosensor that can be measured using fluorescence microscopy or flow cytometry²⁵. PercevalHR is
244 composed of a mutated version of the ATP-binding bacterial protein GlnK1 and the circular permuted
245 monomeric Venus fluorescent protein. ATP but not ADP binding to the PercevalHR causes a ratiometric shift
246 in the probe fluorescence excitation spectrum providing a read out of ATP:ADP intracellular ratio. In these two
247 metabolic profiling assays, cells are purified from the bone marrow of wild-type B6j (SCENITH) or Vav-iCre

248 Perceval^{fl/fl} (SPICE-Met) mice and treated with either DMSO (control; Co), or small molecule inhibitors of
249 glycolysis (2-deoxy-D-glucose; 2-DG), OXPHOS (Oligomycin; O) and protein synthesis (harringtonine; H). By
250 comparing the fluorescent intensities of ATP, ADP and puromycin across different experimental conditions
251 (**Figure 2e,f**) we can then quantify the bioenergetic state of rare cell types such as HSPCs.

252
253 To assess whether these methods could be applied to study HSPCs, we benchmarked them by comparing the
254 metabolic profiles of HSCs (Lin⁻ cKit⁺ Cd48⁻ Slamf⁺) and MPPs, for which a number of metabolic differences have
255 already been reported^{7,40,41}. Consistent with previous reports^{7,40,41}, SCENITH and SPICE-Met profiling showed
256 that HSCs had a higher glycolytic capacity and lower protein synthesis rate than MPPs (**Figure S7b-d**), confirming
257 that our approach can be successfully applied to study other hematopoietic progenitor subsets. To assess
258 whether the MetaFate-myeloid signature was reflective of differences in metabolic pathway activity, we then
259 compared the metabolic profiles of CD62L^{neg} and CD62L^{high} MPPs (Lin⁻ cKit⁺ Sca1⁺) (**Figure 2g,k**). SCENITH
260 profiling showed that CD62L^{high} MPPs have a significantly higher rate of protein synthesis than CD62L^{neg} MPPs
261 ($p < 0.001$) and their translation rates are highly sensitive to oligomycin treatment ($p = 0.001$) (**Figure 2h-i**).
262 Similar results were obtained when we measured the uptake of mitochondrial membrane potential TMRE dye
263 in the different MPP subsets ($p < 0.001$) (**Figure 2j**). Using SPICE-Met, we observed that CD62L^{high} MPPs had a
264 lower ATP/ADP ratio ($p < 0.001$) than CD62L^{neg} MPPs (**Figure 2k,l**) and a higher OXPHOS-dependence ($p < 0.001$)
265 (**Figure 2m**), corroborating results from SCENITH. To understand if the relationship between translation rates
266 and OXPHOS was entirely glucose dependent, or to what extent the breakdown of fatty and amino acids via
267 the TCA cycle was also involved, we measured ATP:ADP ratios of MPPs following inhibition of glucose
268 metabolism using 2-DG. In this analysis, ATP:ADP ratios (**Figure 2n**), were more sensitive to glucose inhibition
269 in CD62L^{high} MPPs compared CD62L^{neg} MPPs ($p = 0.02$), suggesting that the higher rates of translation observed
270 in CD62L^{high} MPPs are typically fuelled by glucose, rather than through the oxidation of fatty and amino -acids.

271
272 Consistent with our MetaFate analyses, CD62L^{high} MPPs have distinct metabolic properties compared to other
273 multipotent progenitors. Specifically we observed higher rates of ATP turnover and protein synthesis, with
274 metabolic demands fuelled by increased glucose catabolism in the mitochondria, rather than by increasing the
275 rate of fatty and amino acid oxidation. Our results suggest that metabolic remodelling accompanies the earliest
276 phases of hematopoietic lineage specification, and so we next investigated whether metabolism plays an active
277 or a passive role in the decision making process.

279 **The Pentose Phosphate Pathway Actively Regulates Immune Cell Production**

280 MetaFate analyses and metabolic profiling highlighted a number of metabolic pathways that are associated
281 with early myeloid development. This led us to hypothesise that manipulating metabolic processes within MPPs
282 could influence the rate of immune cell production. To better discriminate between pathways that have an
283 active versus a passive role in myelopoiesis, we applied the MIIC causal network reconstruction algorithm⁴² to
284 enzyme and transporter expression data obtained from bulk RNA sequencing samples across the entire
285 hematopoietic system⁴³ (**Figure S8**). MIIC is an information theoretic method which learns graphical models
286 from observational data, including the effects of unobserved latent variables^{42,44}. MIIC network reconstruction
287 predicted that in myeloid cells, enzymes of the PPP (*Taldo1*, *G6pdx*) are strongly associated with redox state
288 (*Gsr*, *Mgst1*, *Mgst2*), glucose metabolism (*Hk2*, *Pkm*), NADPH-oxidase activity (*Ncf1*) and lipid metabolism
289 (*Scarb1*, *Abcd1*, *Acer3*), and collectively, these enzymes contribute to myeloid lineage specification. This
290 prediction was consistent with MetaFate results, which showed that enzymes associated with the PPP are
291 upregulated in myeloid-biased MPPs (**Figure 1h**). Together, this led to the hypothesis that manipulating the PPP
292 in MPPs could be a strategy to regulate the dynamics of immune cell production.

293 To test this hypothesis, we use a murine model where Glucose-6-Phosphate-Dehydrogenase (G6PD), the rate
294 limiting enzyme of the PPP, is overexpressed⁴⁵. In this G6PD overexpression system a large genomic fragment
295 (20.1Kb) of the entire human G6PD gene, including upstream and downstream regulatory sequences was

296 inserted into the genome of a transgenic mouse line (G6PD-Tg). In G6PD-Tg mice, G6PD expression is increased
297 2-fold at the RNA level compared to WT littermate controls, a phenotype associated with increased G6PD
298 enzyme activity and NADPH production rates⁴⁵. To assess the functional consequences of G6PD overexpression
299 on MPP differentiation, we transplanted G6PD-tg and WT MPPs and quantified their differentiation patterns *in*
300 *vivo* using a lentiviral cellular barcoding approach (**Figure 3a**). A single cell lineage tracing approach such as
301 lentiviral barcoding allows to follow the fate of heterogenous cells like MPPs. MPPs were purified from the
302 bone marrow of G6PD-Tg mice or WT littermate controls and infected with the LG2.2 lentiviral barcode library
303 as previously described⁴⁶. Cells were then transplanted into sub-lethally irradiated (6Gy) recipients and left to
304 engraft, divide and differentiate. At day 21 after transplantation, the timepoint where myeloid production from
305 transplanted MPPs peaks⁴⁷, barcoded (GFP⁺) erythroblasts (E; Ter119⁺ CD44⁺), myeloid cells (M;
306 Ter119⁻ CD19⁻ CD11b⁺), and B-cells (B; Ter119⁻ CD11b⁻ CD19⁺) were sorted from the bone marrow and their
307 barcode identity was assessed through PCR and deep sequencing from their bulk DNA (**Figure 3a and S9a**).

308 Analysis of chimerism post-transplantation using the proportion of GFP⁺ barcoded cells in the erythroid,
309 myeloid, and B-cell compartments showed a significant decrease in GFP⁺ cells in G6PD-Tg derived leukocytes
310 relative to WT controls, with a trending decrease in B-cell chimerism ($p = 0.057$) but no change in myeloid
311 chimerism ($p = 0.34$) (**Figure 3b-e**). This result show that over-expression of a myeloid associated enzyme, G6PD
312 lead to impaired B-lymphopoiesis. To better assess this phenotype, we analysed the distribution of lentiviral
313 barcodes amongst mature cell types. For G6PD-Tg and WT samples, we detected similar number of sequencing
314 reads (**Figure S9b**) as well as a high consistency between PCR duplicates (**Figure S9c-d**) and very little sharing
315 of barcodes between mice (**Figure S9e**). Following these QC steps, we analysed the diversity, clone size
316 distributions, and lineage bias of G6PD-Tg vs WT barcoded MPPs to understand the progenitor dynamics that
317 gave rise to reduced leukocyte chimerism. Consistent with chimerism measurements, we observed no
318 significant differences in the diversity, clone sizes, and bias of the erythroid- and myeloid- producing barcoded
319 MPPs between the G6PD-Tg and control group (**Figure S10c-h**). Within the B-cell lineage, the cumulative
320 barcode read distribution showed that only a small fraction ($11.9 \pm 2.5\%$ of WT and $21 \pm 9\%$ of G6PD-Tg) of
321 transplanted cells give rise to 95% of all donor-derived B cells (**Figure S10b**), contrary to the other two lineages
322 (**Figure S10c-h**). Focusing on these B-cell producing MPP barcodes, we observed a 2.7-fold reduction in the
323 number of B-cells derived from each barcoded MPP in the G6PD-Tg group compared to WT (**Figure 3f**), with a
324 median clone size of 1692 ± 689 cells for WT B-cell barcodes, but only 630 ± 434 cells for G6PD-Tg B-cell
325 barcodes (**Figure 3f, Figure S10i**). This reduced amount of lymphoid cells produced per MPP due to G6PD
326 overexpression was then partly compensated at the population level by an increase the total number of MPPs
327 producing B-cells in the G6PD-Tg group compared to WT (**Figure 3g**). Overall per individual MPP, G6PD over-
328 expression led to a net-skewing of cell production towards the erythro-myeloid lineages at the expense of the
329 lymphoid lineage (**Figure 3h-j**). In summary, by combining targeted genetics and cellular barcoding approaches,
330 we showed with single cell resolution that overexpression of the pentose phosphate pathway, a key pathway
331 from our meta fate myeloid derived gene signature, limits B-cell production *in vivo*. This result confirmed our
332 hypothesis that manipulating metabolic processes within MPPs can regulate the dynamics of immune cell
333 production *in vivo*.

334 CD62L^{high} MPPs Fuel Emergency Myelopoiesis during acute infection and bone marrow transplantation

335
336 Hematopoiesis is a highly dynamic system, and must adapt to meet changing requirements for blood and
337 immune cells. We hypothesised that the metabolically-primed myeloid-biased multipotent progenitors we
338 identified via MetaFate, could play a significant role in emergency myelopoiesis where the rate of myeloid cell
339 production increases significantly⁴⁸.

340
341 To assess the role of CD62L^{high} MPPs in infection we first used an LPS challenge model in which mice are given
342 35 μ g of LPS at 0 and 48 hours, and bone marrow samples are processed for flow cytometry analysis at 72
343 hours⁴⁹ (**Figure 4a**). In this established model of emergency myelopoiesis⁴⁹, we observe increased CD62L
344 expression in MPPs (**Figure 4b**) and an increased proportion of CD62L^{high} MPPs (Figure 4c). This change

345 correlated with large increases in both the GMP (cKit⁺ Sca1⁻ CD16/32⁺ CD34⁻) and myeloid (Cd11b⁺)
346 compartments of the bone marrow (**Figure 4c**). To assess whether the CD62L^{high} HSPC expansion can also occur
347 following infection with a live pathogen, we re-analysed a scRNAseq dataset of cKit⁺ progenitors from WT mice,
348 or mice infected with *Plasmodium Berghei* 7 days post infection⁵⁰ (**Figure 4d**). In this setting, our MetaFate-
349 myeloid expression program gene signature was significantly increased in HSPCs relative to control cells (**Figure**
350 **4f,g**). Additionally, we saw a concomitant increase in the number of cells which expressed *Sell* – the gene
351 encoding CD62L (**Figure 4e,h**). Taken together, our results show that the CD62L multipotent progenitor
352 compartment plays a key role in supporting immune responses by producing innate immune cells.

353
354 In the context of bone marrow transplantation, the immune system must be restored following conditioning
355 protocols to avoid life-threatening complications⁵¹. To assess whether CD62L^{high} MPPs preferentially
356 reconstitute the myeloid compartment following bone marrow transplantation, we purified CD62L^{high} and
357 CD62L^{neg} MPPs by FACS and transplanted them into irradiated recipient mice. However the use of the MEL-14
358 anti-CD62L antibody clone to sort cells and transplantation led to much poorer engraftment of CD62L^{high} MPPs
359 (**Figure S11a-b**), a result also supported by a report in the literature that MEL-14 inhibits CD62L function on
360 leukocytes⁵². To overcome this limitation, we lentivirally barcoded total MPPs and transplanted them into
361 irradiated recipient mice (**Figure 5a**). Barcodes present in the CD62L^{neg} and CD62L^{high} MPPs, as well as the
362 nucleated erythroid, B cells and myeloid cells were analysed 3 week post-transplantation – the timepoint when
363 myeloid production from MPPs peaks post-transplantation⁴⁷ (**Figure 5a-b**). We obtained 172 barcodes from all
364 the samples that passed QC and filtering with high consistency of sequencing read counts between PCR
365 technical duplicates and very little sharing of barcodes between mice (**Figure S12b-d**). When comparing the
366 distribution of barcodes between CD62L^{neg} and CD62L^{high} MPPs, we found 124 barcodes that are shared
367 between the two types of MPPs, suggesting that CD62L^{low} can give rise to CD62L^{high} MPPs and vice-versa or that
368 cells may transition directly between compartments without undergoing cell division (**Figure 5b**). We also
369 observed barcodes present in only one of the two MPP subsets (**Figure 5b**). Focusing on the differentiation
370 outcome of the barcodes that had more than 95% of its reads in either the CD62L^{low} (CD62L^{low enriched} ; 30
371 barcodes) or the CD62L^{pos} (CD62L^{high enriched} ; 18 barcodes) MPPs, we found that CD62L^{high enriched} barcodes
372 produce more myeloid cells and less B cells compared to CD62L^{low enriched} barcodes, while erythroid production
373 was similar (**Figure 5c**). When comparing the lineage bias score of CD62L^{neg enriched/high enriched} barcodes, we found
374 that CD62L^{high enriched} MPPs had a significantly higher myeloid bias than CD62L^{low} MPPs and significantly reduced
375 B-cell bias and similar erythroid bias (**Figure 5d**). These results were corroborated by unsupervised clustering
376 of the data where CD62L^{high enriched} barcodes cluster most closely with the myeloid lineage than the CD62L^{low}
377 ^{enriched} indicating that CD62L^{high} HSPCs produce more myeloid cells (**Figure 5e**). Importantly, the total number
378 of unique barcodes detected was similar for both the CD62L^{high} and the CD62L^{neg} subsets (**Figure S13a**). This
379 confirmed that the reduced number of CD62L^{neg} barcodes in myeloid cells could not be explained by sampling
380 or sensitivity issues, or differences due to the relative engraftment rates of the subsets. Furthermore, similar
381 patterns in myeloid bias for CD62L^{pos} MPPs was observed when we transplanted lentivirally barcoded CD150⁺
382 HSCs, and purified CD62L MPP subsets at 12 months post transplantation (**Figure S14**). Together, our cellular
383 barcoding experiments show that CD62L^{high} MPPs play a key role in repopulating and maintaining myeloid cell
384 numbers in transplantation hematopoiesis. In summary, our infection and transplantation experiments show
385 that metabolically-primed multipotent progenitors play a key role in fuelling emergency myelopoiesis.

386 387 Discussion

388
389 A key goal for the field of stem cell biology is to identify the molecular signals that induce stem cells to
390 selectively differentiate into a cell type of interest *in vivo*. Previous work has highlighted the critical role of
391 transcription and growth factors in regulating lineage commitment, but the role of metabolism is less clear in
392 this context. In this study we developed MetaFate to trace the metabolic state and developmental fate of single
393 HSPCs *in vivo*. Using this innovative approach, we characterise the metabolic cues that instruct early myeloid
394 development, showing that the pentose phosphate pathway plays an active role in this process and can be
395 manipulated to alter the rate of immune cell regeneration.

396

397 To add a functional dimension to metabolic studies of rare cell types *in vivo* we developed MetaFate, a lineage
398 tracing approach to perform state-fate mapping focused on metabolically-associated gene modules. Within a
399 single mouse, there are an estimated 1.4×10^5 MPPs, representing just 0.03% of total bone marrow cellularity².
400 Consequently, cell yields falling far below the sensitivity limits of many metabolomics methods, making it
401 technically challenging to study the metabolic properties of these rare subsets. Furthermore, existing methods
402 are destructive – making it difficult to link the metabolic and functional properties of heterogeneous cell
403 populations. Because of these technical challenges, much of our understanding of stem cell metabolism comes
404 from population based approaches, limiting our understanding of how metabolism regulates lineage
405 commitment specificity. MetaFate can also interface with other metabolomics methods – identifying
406 functionally resolved sets of enzymes and transporters that can be verified at the protein level using spatially
407 resolved metabolomics methods such as high dimensional mass cytometry²³ or by *in situ* dehydrogenase
408 assays²⁶. We anticipate that combining emerging metabolomics technologies such as SCENITH and SPICE-Met
409 with lineage tracing tools like MetaFate will yield significant insights into how cellular metabolism regulates the
410 function of rare cell types in both health, ageing and disease. Given the diverse array of lineage tracing and
411 metabolomics technologies that are emerging, our strategy can be readily adapted to other stem cell and
412 developmental systems, including human tissues using human-compatible retrospective lineage tracing
413 methods.

414

415 Using MetaFate, we have identified an expression program of enzymes and transporters that confers
416 differences in myeloid lineage potential within a subset of MPPs. Leveraging the ability of the DRAG barcoding
417 system to detect barcodes at both the RNA or the DNA level we were able to measure barcode abundances in
418 both HSPCs and the much larger mature myeloid and nucleated erythroid progenitor compartments. This
419 experimental design enabled us to trace lineage commitment over much longer developmental trajectories
420 compared to studies that measure barcodes only in progenitors^{27,37}. Using only genes encoding metabolic
421 enzymes and transporters, our signature had a higher correlation with myeloid bias than the existing MPP3
422 signature^{34,53}, prompting us to develop a novel purification strategy using the surface marker CD62L. Through
423 *in situ* and lentiviral barcoding experiments, we show that CD62L enriches for myeloid bias in MPPs. This is
424 consistent with reports showing that CD62L enriches for MPPs, rather than HSCs within the LSK compartment⁵⁴
425 and that CD62L enriches for myeloid potential in CMPs⁵⁵. In humans, the CD62L gene *SELL* has been associated
426 with abnormal myeloid cell counts^{56,57}, suggesting that CD62L may have implications in the regulation of
427 hematopoiesis in humans as well as mice. Metabolically, CD62L^{high} MPPs have higher rates of ATP turnover and
428 protein synthesis compared to CD62L^{neg} MPPs. CD62L^{high} MPPs also have a higher dependence on glucose
429 metabolism via OXPHOS to meet their energetic requirements compared to other MPP subsets which had a
430 higher reliance on glycolysis and on fatty/amino acid oxidation.

431

432 Importantly, manipulating metabolic processes in progenitors alters the rate of immune cell production, with
433 overexpression of G6PD altering the dynamics of B-cell producing MPPs, resulting in a net skewing towards the
434 erythromyeloid lineages. This result is consistent with reports that pharmacological inhibition of the pentose
435 phosphate pathway blocks erythropoiesis *in vitro*⁸ and that the pathway regulates the function of dendritic
436 cells⁵⁸ and macrophages^{59,60}. This work bridges understanding between the fields of metabolism, stem cell
437 biology and immunology, highlighting the pentose phosphate pathway as a regulator of immune cell
438 production. While much focused has been placed on the roles of glycolysis and oxidative phosphorylation as
439 key modulators of stem cell metabolism^{14,15}, further work is required to assess whether the pentose phosphate
440 pathway regulate can regulate stem cell function in other systems.

441

442 To understand why metabolic priming of multipotent progenitors may be functionally important we assessed
443 the role of the CD62L^{high} MPP compartment in 2 different emergency myelopoiesis models: infection and
444 transplantation. In both models, myelopoiesis was fuelled by CD62L^{high} MPPs, consistent with the idea that
445 metabolic priming of multipotent progenitors toward the myeloid lineage facilitates the production of innate
446 immune cells in response to injury. Our work therefore shows for the first time that MPPs are metabolically

447 heterogeneous and that a subset of metabolically primed MPPs contribute to innate immunity in emergency
448 settings.

449 In our study, we use transcriptomic changes that occur during fate decisions to infer metabolic differences that
450 we validate using metabolic pathway activity measures. Differences in metabolite levels or others changes not
451 captured by transcriptomic analysis may also regulate fate decisions. In particular, changes in metabolites
452 acting as substrates for chromatin modifiers may precede transcriptomic changes and influence fate. Recent
453 advances in single cell techniques to study the epigenome will help to address this limitation. In this context,
454 metabolic differences could arise by cell extrinsic mechanisms such as differential location within the niche, or
455 cell intrinsic mechanisms whereby the asymmetric distribution of metabolites/cell organelles following cell
456 division could influence lineage potential. Further exploration of these topics is required to better understand
457 the role of metabolism in shaping fate decisions.

458 Understanding the nutrients and metabolites that regulate hematopoiesis can inform the development of novel
459 bone marrow organoid technologies to maintain and differentiate haematopoietic precursors *ex vivo*. Our data
460 can also help to inform the development of nutrient/metabolite biomarker panels for stem cell function. Such
461 tools can inform dietary interventions to promote HSPC function, particularly in prospective recipients of bone
462 marrow transplants, a high-risk procedure that can lead to malnutrition and significant nutrient deficiencies^{61,62}.
463 Lastly, our results suggest that therapeutic interventions to alter the metabolism of HSPCs may not target all
464 cells uniformly, given their underlying metabolic heterogeneity. Our approach may therefore be useful in
465 studying whether this phenomenon occurs in other systems, such as cancer stem cells.

466 Collectively, we have identified the metabolic cues that guide the earliest stages of innate immune cell
467 development, highlighting a key role for the pentose phosphate pathway. More broadly, our results suggest
468 that manipulating lineage-specific metabolic cues can alter the cellular composition of the immune system *in*
469 *vivo*.

470 **Materials and Methods**

471 A detailed description of all materials and method is provided in the supplementary information

472

473 **Acknowledgements**

474 We thank the Institute Curie flow cytometry, next-generation sequencing, animal, and UMR168 BMBC facility.

475 We thank all the members of the Perié team for helpful discussion.

476

477 **Author Contributions**

478 **J.C.** conceptualization, performed experiments, data curation, data analysis, methodology, writing, funding
479 acquisition. **A.M.L.** data analysis, methodology, writing - review and editing. **I.R.** performed experiments **V.C.**
480 data analysis, writing, review and editing. **C.C.** methodological development, performed experiments **S.T.B.**
481 methodological development. **E.T.** methodological development. **E.R.** performed experiments. **F.T.** performed
482 experiments. **Y.B.** performed experiments. **S.M.T.** performed experiments. **S.R.** performed experiments. **F.M.**
483 performed experiments. **A.A.** provided expertise and reagents, writing, review and editing **C.L.** provided
484 expertise and reagents, performed experiments, writing - review and editing. **P.B.** provided expertise and
485 reagents, Writing - review and editing. **P.J.F.M.** provided expertise and reagents, writing - review and editing.
486 **H.I.** Formal analysis, Writing - review and editing. **R.J.A.** provided expertise and reagents, writing - review and
487 editing. **L.P.** conceptualization, data analysis, funding acquisition, methodology, supervision, writing.

488

489 **Funding**

490 This work was supported by grants from the *Labex Cell(n)Scale* (ANR-11-LABX-0038 , ANR-10-IDEX-0001-02 PSL
491) (to L.P.). This work is part of a project that has received funding from the European Research Council (ERC)
492 under the European Union’s Horizon 2020 research and innovation programme 758170-Microbar (to L.P.). J.C.
493 was supported by a Foundation ARC fellowship and by the Agence Nationale de Recherche (DROPTREP: ANR-
494 16-CE18-0020-03).

495

496 **Ethics**

497 All the experimental procedures were approved by the local ethics committee CEEA-IC (Comité d’Ethique en
498 expérimentation animale de l’Institut Curie) under approval numbers DAP 2016 006, DAP 2021-010 and DAP
499 2021-013

500

501 **Data Availability**

502 All datasets generated or reanalysed during this study are available at:
503 <https://github.com/TeamPerie/Cosgrove-et-al-2022>

504

505 **Code Availability**

506 All source code generated during this study is available at: <https://github.com/TeamPerie/Cosgrove-et-al-2022>

507

508 **References**

- 509 1. Sender, R. & Milo, R. The distribution of cellular turnover in the human body. *Nat. Med.*
510 **27**, 45–48 (2021).
- 511 2. Cosgrove, J., Hustin, L. S. P., de Boer, R. J. & Perié, L. Hematopoiesis in numbers. *Trends*
512 *Immunol.* S1471-4906(21)00211–8 (2021) doi:10.1016/j.it.2021.10.006.
- 513 3. Boettcher, S. & Manz, M. G. Regulation of Inflammation- and Infection-Driven
514 Hematopoiesis. *Trends Immunol.* **38**, 345–357 (2017).
- 515 4. Oburoglu, L., Romano, M., Taylor, N. & Kinet, S. Metabolic regulation of hematopoietic
516 stem cell commitment and erythroid differentiation. *Curr. Opin. Hematol.* **23**, 198–205
517 (2016).
- 518 5. Morganti, C., Cabezas-Wallscheid, N. & Ito, K. Metabolic Regulation of Hematopoietic
519 Stem Cells. *HemaSphere* **6**, e740 (2022).
- 520 6. Takubo, K. *et al.* Regulation of glycolysis by Pdk functions as a metabolic checkpoint for
521 cell cycle quiescence in hematopoietic stem cells. *Cell Stem Cell* **12**, 49–61 (2013).

- 522 7. Signer, R. A. J., Magee, J. A., Salic, A. & Morrison, S. J. Haematopoietic stem cells require
523 a highly regulated protein synthesis rate. *Nature* **509**, 49–54 (2014).
- 524 8. Oburoglu, L. *et al.* Glucose and glutamine metabolism regulate human hematopoietic stem
525 cell lineage specification. *Cell Stem Cell* **15**, 169–184 (2014).
- 526 9. Ito, K. *et al.* A PML–PPAR- δ pathway for fatty acid oxidation regulates hematopoietic
527 stem cell maintenance. *Nat. Med.* **18**, 1350–1358 (2012).
- 528 10. Ito, K. *et al.* Self-renewal of a purified Tie2⁺ hematopoietic stem cell population relies on
529 mitochondrial clearance. *Science* **354**, 1156–1160 (2016).
- 530 11. Cabezas-Wallscheid, N. *et al.* Vitamin A-Retinoic Acid Signaling Regulates Hematopoietic
531 Stem Cell Dormancy. *Cell* **169**, 807-823.e19 (2017).
- 532 12. Agathocleous, M. *et al.* Ascorbate regulates haematopoietic stem cell function and
533 leukaemogenesis. *Nature* **549**, 476–481 (2017).
- 534 13. Qi, L. *et al.* Aspartate availability limits hematopoietic stem cell function during
535 hematopoietic regeneration. *Cell Stem Cell* **28**, 1982-1999.e8 (2021).
- 536 14. Ly, C. H., Lynch, G. S. & Ryall, J. G. A Metabolic Roadmap for Somatic Stem Cell Fate.
537 *Cell Metab.* **31**, 1052–1067 (2020).
- 538 15. Shapira, S. N. & Christofk, H. R. Metabolic Regulation of Tissue Stem Cells. *Trends Cell*
539 *Biol.* **30**, 566–576 (2020).
- 540 16. Busch, K. *et al.* Fundamental properties of unperturbed haematopoiesis from stem cells in
541 vivo. *Nature* **518**, 542–546 (2015).
- 542 17. Naik, S. H. *et al.* Diverse and heritable lineage imprinting of early haematopoietic
543 progenitors. *Nature* **496**, 229–232 (2013).
- 544 18. Perié, L., Duffy, K. R., Kok, L., de Boer, R. J. & Schumacher, T. N. The Branching Point
545 in Erythro-Myeloid Differentiation. *Cell* **163**, 1655–1662 (2015).

- 546 19. Shamir, M., Bar-On, Y., Phillips, R. & Milo, R. SnapShot: Timescales in Cell Biology. *Cell*
547 **164**, 1302-1302.e1 (2016).
- 548 20. DeVilbiss, A. W. *et al.* Metabolomic profiling of rare cell populations isolated by flow
549 cytometry from tissues. *eLife* **10**, e61980 (2021).
- 550 21. Sun, J. *et al.* Clonal dynamics of native haematopoiesis. *Nature* **514**, 322–327 (2014).
- 551 22. Pei, W. *et al.* Polylox barcoding reveals haematopoietic stem cell fates realized in vivo.
552 *Nature* **548**, 456–460 (2017).
- 553 23. Hartmann, F. J. *et al.* Single-cell metabolic profiling of human cytotoxic T cells. *Nat.*
554 *Biotechnol.* (2020) doi:10.1038/s41587-020-0651-8.
- 555 24. Argüello, R. J. *et al.* SCENITH: A Flow Cytometry-Based Method to Functionally Profile
556 Energy Metabolism with Single-Cell Resolution. *Cell Metab.* **32**, 1063-1075.e7 (2020).
- 557 25. Russo, E. *et al.* SPICE-Met: profiling and imaging energy metabolism at the single-cell
558 level using a fluorescent reporter mouse. *EMBO J.* e111528 (2022)
559 doi:10.15252/embj.2022111528.
- 560 26. Miller, A. *et al.* Exploring Metabolic Configurations of Single Cells within Complex Tissue
561 Microenvironments. *Cell Metab.* **26**, 788-800.e6 (2017).
- 562 27. Rodriguez-Fraticelli, A. E. *et al.* Single-cell lineage tracing unveils a role for TCF15 in
563 haematopoiesis. *Nature* **583**, 585–589 (2020).
- 564 28. Pei, W. *et al.* Resolving Fates and Single-Cell Transcriptomes of Hematopoietic Stem Cell
565 Clones by PolyloxExpress Barcoding. *Cell Stem Cell* **27**, 383-395.e8 (2020).
- 566 29. Cavicchioli, M. V., Santorsola, M., Balboni, N., Mercatelli, D. & Giorgi, F. M. Prediction
567 of Metabolic Profiles from Transcriptomics Data in Human Cancer Cell Lines. *Int. J. Mol.*
568 *Sci.* **23**, 3867 (2022).
- 569 30. Sawai, C. M. *et al.* Hematopoietic Stem Cells Are the Major Source of Multilineage
570 Hematopoiesis in Adult Animals. *Immunity* **45**, 597–609 (2016).

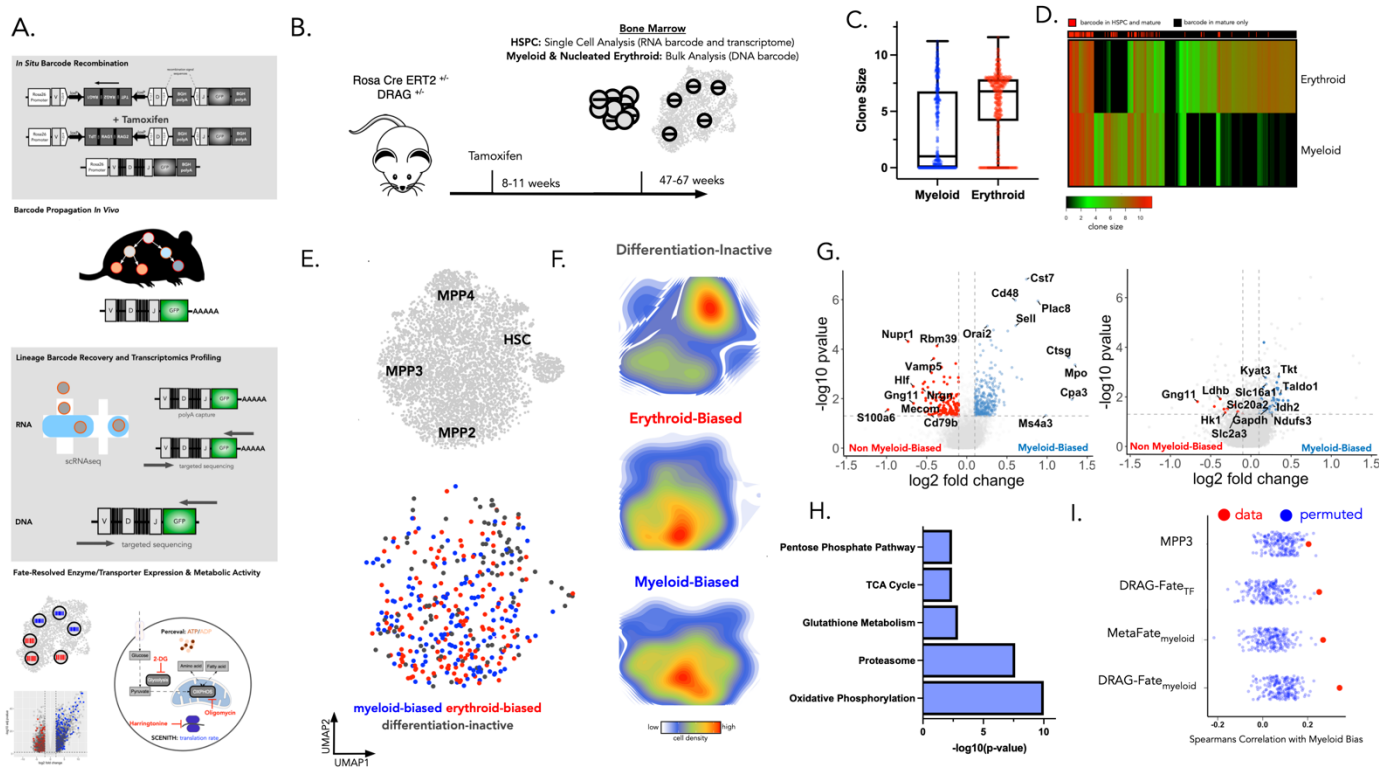
- 571 31. Bowling, S. *et al.* An Engineered CRISPR-Cas9 Mouse Line for Simultaneous Readout of
572 Lineage Histories and Gene Expression Profiles in Single Cells. *Cell* **181**, 1410-1422.e27
573 (2020).
- 574 32. Wilson, N. K. *et al.* Combined Single-Cell Functional and Gene Expression Analysis
575 Resolves Heterogeneity within Stem Cell Populations. *Cell Stem Cell* **16**, 712–724 (2015).
- 576 33. Giladi, A. *et al.* Single-cell characterization of haematopoietic progenitors and their
577 trajectories in homeostasis and perturbed haematopoiesis. *Nat. Cell Biol.* **20**, 836–846
578 (2018).
- 579 34. Pietras, E. M. *et al.* Functionally Distinct Subsets of Lineage-Biased Multipotent
580 Progenitors Control Blood Production in Normal and Regenerative Conditions. *Cell Stem*
581 *Cell* **17**, 35–46 (2015).
- 582 35. Tusi, B. K. *et al.* Population snapshots predict early haematopoietic and erythroid
583 hierarchies. *Nature* **555**, 54–60 (2018).
- 584 36. Dahlin, J. S. *et al.* A single-cell hematopoietic landscape resolves 8 lineage trajectories and
585 defects in Kit mutant mice. *Blood* **131**, e1–e11 (2018).
- 586 37. Weinreb, C., Rodriguez-Fraticelli, A., Camargo, F. D. & Klein, A. M. Lineage tracing on
587 transcriptional landscapes links state to fate during differentiation. *Science* **367**, (2020).
- 588 38. Wolf, F. A. *et al.* PAGA: graph abstraction reconciles clustering with trajectory inference
589 through a topology preserving map of single cells. *Genome Biol.* **20**, 59 (2019).
- 590 39. Paul, F. *et al.* Transcriptional Heterogeneity and Lineage Commitment in Myeloid
591 Progenitors. *Cell* **163**, 1663–1677 (2015).
- 592 40. Chen, C. *et al.* TSC–mTOR maintains quiescence and function of hematopoietic stem cells
593 by repressing mitochondrial biogenesis and reactive oxygen species. *J. Exp. Med.* **205**,
594 2397–2408 (2008).

- 595 41. Yu, W.-M. *et al.* Metabolic regulation by the mitochondrial phosphatase PTPMT1 is
596 required for hematopoietic stem cell differentiation. *Cell Stem Cell* **12**, 62–74 (2013).
- 597 42. Verny, L., Sella, N., Affeldt, S., Singh, P. P. & Isambert, H. Learning causal networks with
598 latent variables from multivariate information in genomic data. *PLOS Comput. Biol.* **13**,
599 e1005662 (<https://doi.org/10.1371/journal.pcbi.1005662>) (2017).
- 600 43. Choi, J. *et al.* Haemopedia RNA-seq: a database of gene expression during haematopoiesis
601 in mice and humans. *Nucleic Acids Res.* **47**, D780–D785 (2019).
- 602 44. Cabeli, V. *et al.* Learning clinical networks from medical records based on information
603 estimates in mixed-type data. *PLOS Comput. Biol.* **16**, e1007866 (2020).
- 604 45. Nóbrega-Pereira, S. *et al.* G6PD protects from oxidative damage and improves healthspan
605 in mice. *Nat. Commun.* **7**, 10894 (2016).
- 606 46. Eisele, A. S. *et al.* Erythropoietin directly remodels the clonal composition of murine
607 hematopoietic multipotent progenitor cells. *eLife* **11**, e66922 (2022).
- 608 47. Boyer, S. W. *et al.* Clonal and Quantitative In Vivo Assessment of Hematopoietic Stem
609 Cell Differentiation Reveals Strong Erythroid Potential of Multipotent Cells. *Stem Cell*
610 *Rep.* **12**, 801–815 (2019).
- 611 48. Schultze, J. L., Mass, E. & Schlitzer, A. Emerging Principles in Myelopoiesis at
612 Homeostasis and during Infection and Inflammation. *Immunity* **50**, 288–301 (2019).
- 613 49. Boettcher, S. *et al.* Cutting edge: LPS-induced emergency myelopoiesis depends on TLR4-
614 expressing nonhematopoietic cells. *J. Immunol. Baltim. Md 1950* **188**, 5824–5828 (2012).
- 615 50. Haltalli, M. L. R. *et al.* Manipulating niche composition limits damage to haematopoietic
616 stem cells during Plasmodium infection. *Nat. Cell Biol.* **22**, 1399–1410 (2020).
- 617 51. Mitroulis, I., Kalafati, L., Hajishengallis, G. & Chavakis, T. Myelopoiesis in the Context of
618 Innate Immunity. *J. Innate Immun.* **10**, 365–372 (2018).

- 619 52. Pizcueta, P. & Luscinskas, F. W. Monoclonal antibody blockade of L-selectin inhibits
620 mononuclear leukocyte recruitment to inflammatory sites in vivo. *Am. J. Pathol.* **145**, 461–
621 469 (1994).
- 622 53. Sommerkamp, P. *et al.* Mouse multipotent progenitor 5 cells are located at the interphase
623 between hematopoietic stem and progenitor cells. *Blood* **137**, 3218–3224 (2021).
- 624 54. Cho, S. & Spangrude, G. J. Enrichment of functionally distinct mouse hematopoietic
625 progenitor cell populations using CD62L. *J. Immunol. Baltim. Md 1950* **187**, 5203–5210
626 (2011).
- 627 55. Ito, Y., Nakahara, F., Kagoya, Y. & Kurokawa, M. CD62L expression level determines the
628 cell fate of myeloid progenitors. *Stem Cell Rep.* **16**, 2871–2886 (2021).
- 629 56. Orrù, V. *et al.* Complex genetic signatures in immune cells underlie autoimmunity and
630 inform therapy. *Nat. Genet.* **52**, 1036–1045 (2020).
- 631 57. Vuckovic, D. *et al.* The Polygenic and Monogenic Basis of Blood Traits and Diseases. *Cell*
632 **182**, 1214-1231.e11 (2020).
- 633 58. Everts, B. *et al.* TLR-driven early glycolytic reprogramming via the kinases TBK1- $IKK\epsilon$
634 supports the anabolic demands of dendritic cell activation. *Nat. Immunol.* **15**, 323–332
635 (2014).
- 636 59. Tannahill, G. M. *et al.* Succinate is an inflammatory signal that induces IL-1 β through HIF-
637 1 α . *Nature* **496**, 238–242 (2013).
- 638 60. Haschemi, A. *et al.* The sedoheptulose kinase CARKL directs macrophage polarization
639 through control of glucose metabolism. *Cell Metab.* **15**, 813–826 (2012).
- 640 61. *The EBMT Handbook: Hematopoietic Stem Cell Transplantation and Cellular Therapies.*
641 (Springer Nature, 2019). doi:10.1007/978-3-030-02278-5.
- 642 62. Akbulut, G. & Yesildemir, O. Overview of nutritional approach in hematopoietic stem cell
643 transplantation: COVID-19 update. *World J. Stem Cells* **13**, 1530–1548 (2021).

644
645
646
647
648
649
650

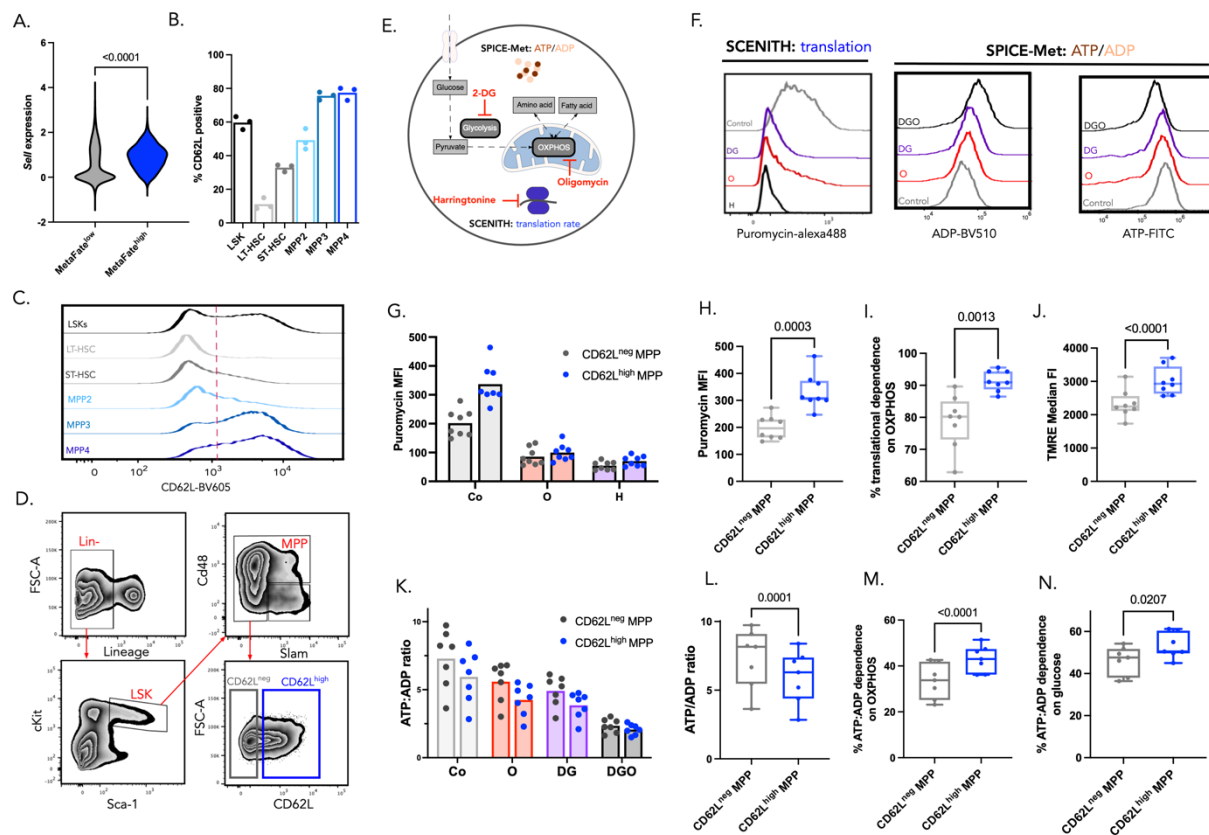
Figures



651
652

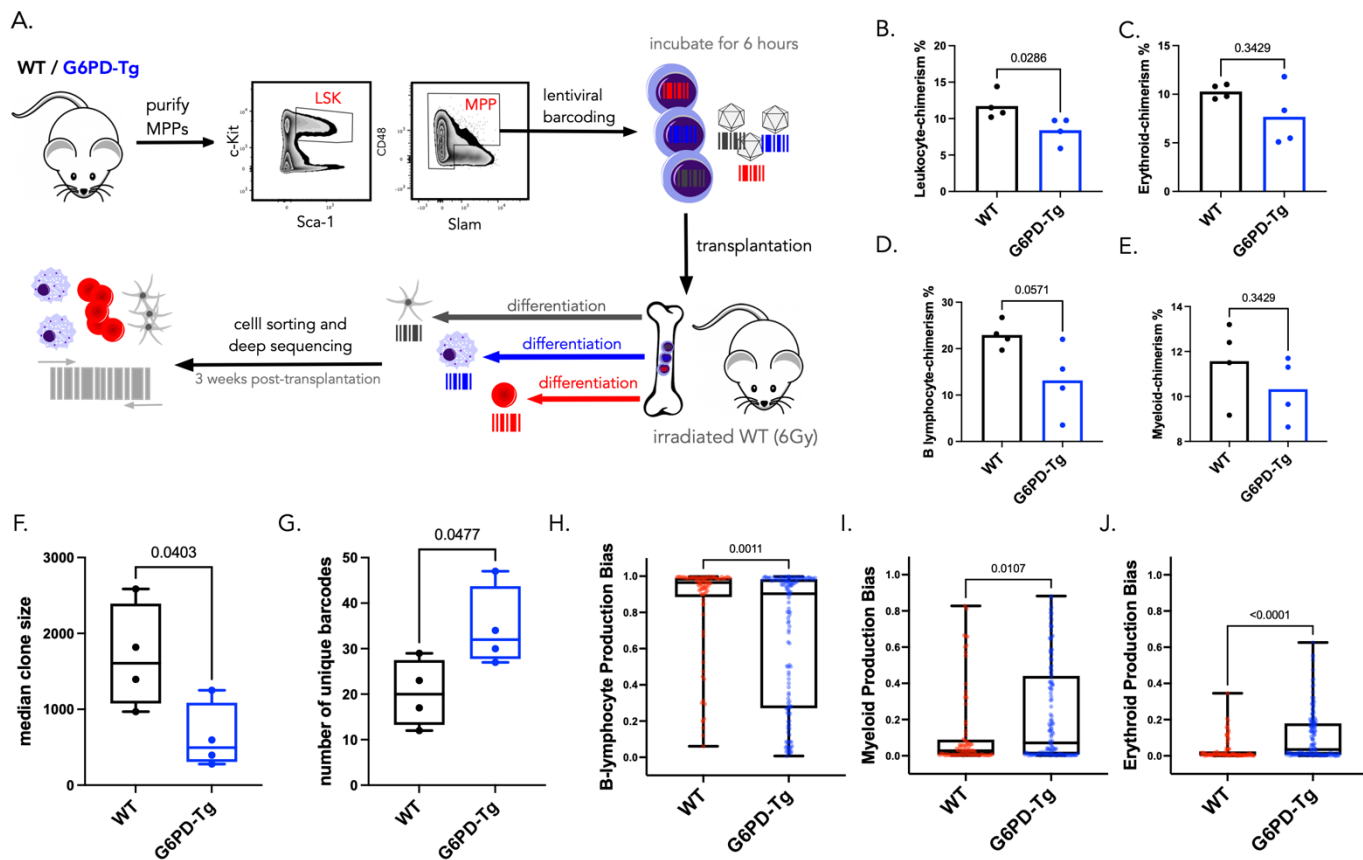
653 **Figure 1. Myeloid-biased HSPCs have a distinct expression program of enzymes and transporters.** (A) Overview of MetaFate and
654 experimental set up: it consists in the induction of a lineage barcode in cells, the propagation of this barcode in vivo when cells divide
655 and differentiate and then the recovery of the transcriptomes of HSPC and their barcode from RNA as well as the recovery of the barcode
656 of their mature progeny by DNA. Tamoxifen injection in mice induces the recombination of the lineage barcode in situ. After division
657 and differentiation of the barcoded cells, all offspring inheriting the barcode and a GFP tag. At the end time of the experiment, GFP-
658 expressing HSPCs (Sca1⁺ cKit⁺), nucleated erythroid progenitors (CD44⁺ Ter119⁺) and myeloid (Cd11b⁺) cells are isolated from the bone
659 marrow using FACS. Then bulk of mature cells were processed through nested PCR for barcode detection at the DNA level and sequenced.
660 In parallel, HSPCs were processed through 10X scRNAseq to recover their transcriptome. Specific targeted PCR amplification were
661 performed on the cDNA obtained by scRNAseq to recover the barcodes from the HSPC. Then, the metaFate bioinformatic pipeline
662 consolidates expression and lineage barcode data and identifies enzymes and transporters that can be targeted for functional studies,
663 as well as surface markers to purify cell subsets for downstream metabolomics profiling. (B) Experimental timeline for induction and
664 collection of HSPC and mature cells for metaFate profiling (C) Clone sizes, number of cells per barcode, in the erythroid and myeloid
665 lineage, the y-axis is transformed using the hyperbolic arcsin function. Each point represents a single barcode. (D) Heatmap

666 representation of DNA barcode expression in myeloid and erythroid cells. Normalized and hyperbolic arcsin transformed cell counts
 667 (clone size) data were clustered by hierarchical clustering using Euclidian distance. Color indicates hyperbolic arcsin transformed cell
 668 counts (clone size) The top column indicates barcodes that are found both in HSPCs and mature cells (red), or barcodes found only in
 669 mature cells (black). (E) UMAP representation of the MetaFate dataset of LSK cells overlaying the positioning of known HSC and MPP
 670 subsets (top left) as well as the localisation of myeloid-biased (blue), erythroid-biased (red), and differentiation inactive (grey)
 671 progenitors based on lineage barcode (bottom left). This figure represents 4,485 Sca1⁺ cKit⁺ GFP⁺ cells (668 RNA-barcoded cells ; 158
 672 unique barcodes). (F) Density map highlighting the localization of lineage biased barcoded cells on our UMAP representation of the data.
 673 (G) Volcano plots showing differentially expressed genes between myeloid-biased barcoded cells and other (erythroid-biased and
 674 differentiation inactive) barcoded subsets. Left hand side highlights the top differentially expressed genes. All genes upregulated in
 675 myeloid-biased barcoded cells compared to erythroid and differentiation inactive-barcoded cells form a gene-signature called
 676 DRAGFate-Myeloid (271 genes) and are shown in the plot on the left hand side. The plot on the right hand side highlights differentially
 677 expressed enzymes and transporters. Downregulated and upregulated genes encoding enzymes and transporters (57 genes) are
 678 highlighted in the right hand plot in red and blue respectively. The subset of genes from the DRAGFate-myeloid signature relating to
 679 cellular metabolism form the MetaFate-myeloid signature. The signature score corresponds to the average expression values of these
 680 gene sets for each cell and is projected onto the UMAP visualisation of the data. (H) Metabolic pathways from the KEGG database that
 681 are enriched amongst genes upregulated in myeloid biased barcoded progenitors compared to erythroid and differentiation inactive
 682 barcoded cell subsets. (I) Spearman's Correlation between different transcriptomic signatures and the myeloid bias of lineage barcodes.
 683 Red points represent the correlation between signature scores and myeloid bias score, while blue points represent the correlations
 684 observed for randomised gene-sets of an equivalent size. The MPP3 signature is taken from Sommerkamp et al (2021), each point
 685 represents a different mouse. In this figure all data was taken from 5 mice from 3 independent experiments.
 686



687
 688 **Figure 2. CD62L^{high} Multipotent Progenitors Are Characterised by a Reduced ATP/ADP Ratio and Higher Rates of Protein Synthesis,**
 689 **Oxidative Phosphorylation and Glucose Dependency:** (A) Comparison of *Self* expression between MetaFate^{low/high} expressing
 690 populations. MetaFate-low cells are defined as HSPCs in the bottom 25th percentile of MetaFate signature expression. MetaFate-high
 691 cells are defined as HSPCs in the top 75th percentile of MetaFate signature expression (B-C) L-selectin expression across different HSPC
 692 subsets (n = 3 mice). Gating strategy for HSPC are in figure S7a. (D) Flow cytometry gating strategy to purify L-selectin expression MPPs
 693 (E) Overview of our strategy to profile the metabolic state of different HSPC subsets. Cells are incubated in the presence of DMSO
 694 (control) or inhibitors of glycolysis (2-DG), OXPHOS (Oligomycin) or translation (Harringtonine). The scheme shows where the inhibitors
 695 are acting. Following incubation of cells with these inhibitors, we measure either translation rate (puromycin labelling), and or the
 696 percevalHR biosensor as a measure of ATP: ADP ratio. (F) Example staining profiles for SCENITH and Perceval metabolic profiling
 697 experiments on Lin- Sca1⁺ cKit⁺ bone marrow progenitors. (G) Median fluorescence intensity values for puromycin across CD62L MPP

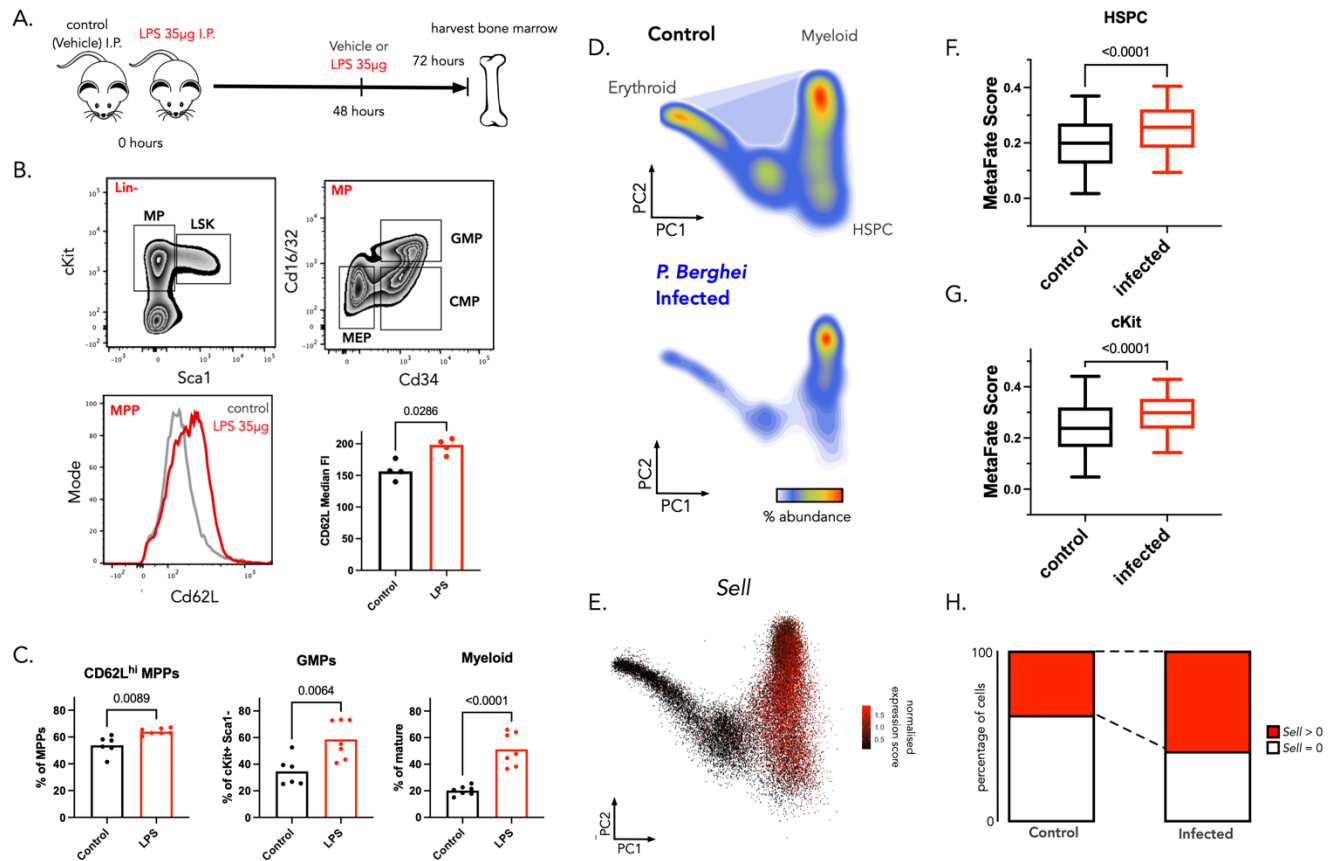
698 subsets and experimental conditions. Each point represents 1 mouse, data pooled from 2 independent experiments (N = 8 mice). (H)
 699 Median fluorescent intensity of puromycin labelling in CD62L^{neg} and CD62L^{hi} MPPs. (N = 8 mice, data pooled from two independent
 700 experiments). (I) Comparison of mitochondrial dependence measures calculated based on changes in puromycin labelling across
 701 control, oligomycin and harringtonine treated conditions. Formula for data transformation is provided in the materials and methods (J)
 702 Median fluorescent intensity measures for tetramethylrhodamine, ethyl ester (TMRE) labelling. N = 8 mice, data pooled from 2
 703 experiments. Each point represents 1 mouse. Statistical comparisons were made using a paired T-test. (K) ATP:ADP measurements
 704 obtained by dividing the median fluorescent intensity of the ATP channel by the median fluorescent intensity of ADP channel. (L)
 705 Comparison of ATP:ADP ratio in CD62L^{neg} and CD62L^{hi} MPP control samples (M-N) Comparison of mitochondrial and glucose dependence
 706 measures inferred from changes in ATP:ADP ratio across control, 2-DG and oligomycin treated conditions. Formula for data
 707 transformation is provided in the materials and methods (N = 8 mice, data pooled from two independent experiments). CD62L^{neg}
 708 samples are highlighted in grey and CD62L^{hi} MPPs are highlighted in blue. Normality of the data was assessed using a Shapiro-Wilk test and
 709 statistical differences for h-j and L-N was assessed using a paired T-test. Barplots represent the mean value across all mice. Boxplots
 710 represent the median and interquartile range with whiskers extending to the minimum and maximum values.
 711



712

713 **Figure 3: Upregulation of Glucose-6-Phosphate-Dehydrogenase in MPPs Inhibits B-Lymphopoiesis Post-Transplantation** (A) Overview
 714 of the lentiviral barcoding experiment. MPPs (say phenotype) were purified from the bone marrow of WT or G6PD-Tg mice by FACS and
 715 were infected with the LG2.2 lentiviral barcoding library. 6 hours later transduced MPPs were injected I.V. into 6Gy irradiated WT
 716 recipients. 3 weeks post-transplantation bone marrow was harvested and cells were sorted and their bulk DNA was processed for
 717 barcode detection through nested PCR and sequencing. (n = 4 mice per condition). (B-E) % chimerism quantified by measuring the
 718 proportion of GFP cells relative to total live cell numbers in the respective lineage compartments by flow cytometry for the WT (black)
 719 and G6PD-Tg (blue) transplanted MPPs. Each point represents a single mouse with N=4 mice per experimental condition. Pairwise
 720 comparisons were made using a Mann-Whitney test. (F) Median clone sizes for the top B-cell producing barcodes (the top *n* barcodes
 721 defined as contributing to 95% of all read counts for the B-cell lineage) from WT (black) and G6PD-Tg (blue) transplanted MPPs. Each
 722 point represents a single mouse. Pairwise comparisons are made using a Student's T-test. (G) Barcode diversity as defined by the number
 723 of unique barcode for B-cell producing barcodes from WT (black) and G6PD-Tg (blue) MPPs, each point represents a single mouse.
 724 Pairwise comparisons are made using a Student's T-test (H-J) The lineage bias of each B-cell producing barcode is shown per lineage
 725 and per experimental condition. The bias is calculated by computing the frequency of barcode *i* in lineage *j*, and then comparing the
 726 relative frequencies of each barcode across all lineages. Each point represents a single barcode (81 WT barcodes (red), 138 G6PDtg
 727 barcodes (blue)) with data pooled from 4 mice per condition. Pairwise comparisons were made using a Mann-Whitney test. Boxplots
 728 represent the median and interquartile range with whiskers extending to the minimum and maximum values.

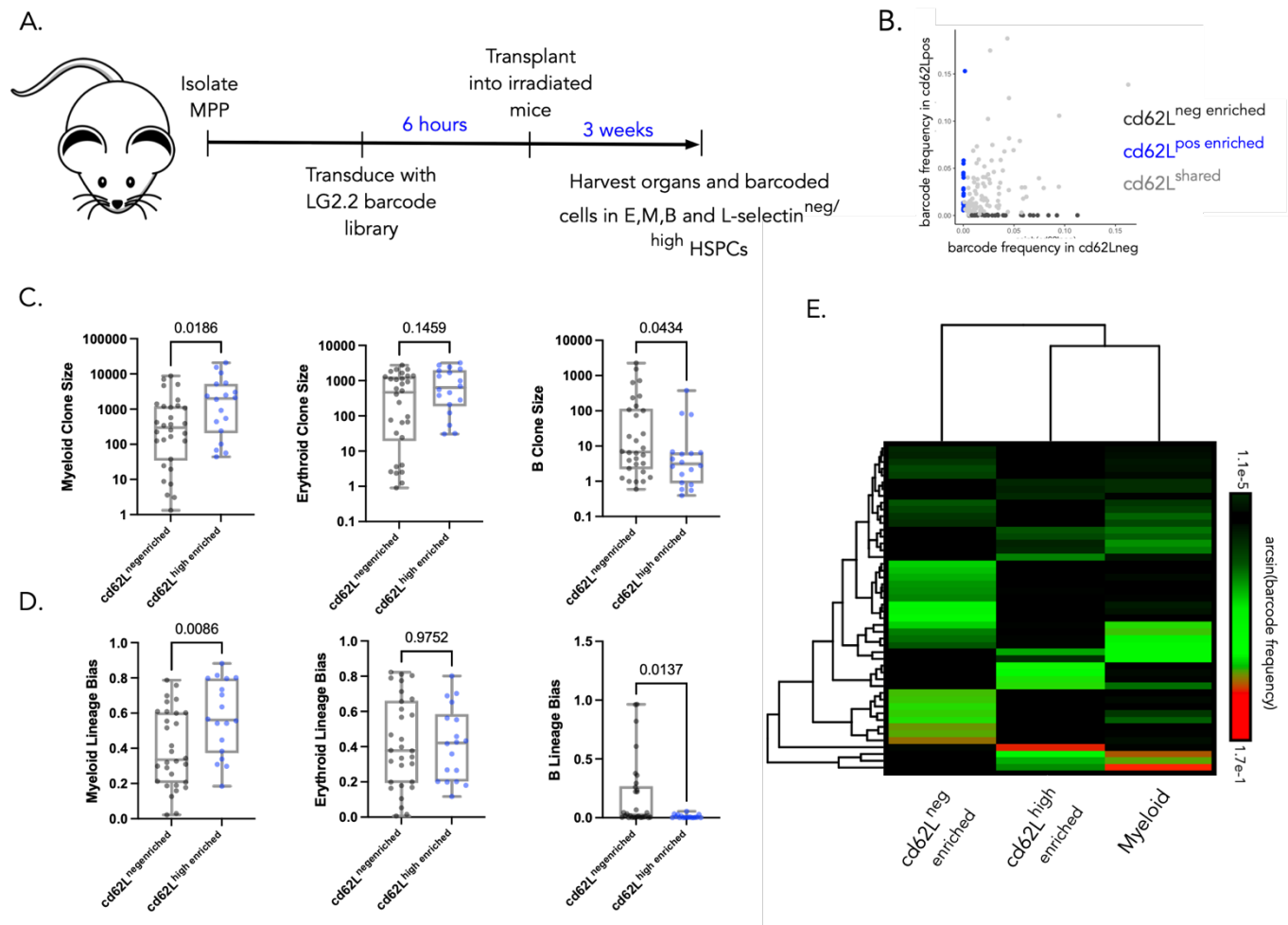
729
730
731



732

733 **Figure 4: The CD62L^{hi} MPP compartment expands to fuel emergency myelopoiesis during acute infection** (A) Overview of the LPS
734 challenge model. 18 week old B6j Mice were injected with LPS (35ug/mouse) I.P. at 0 hours and 48 hours. At 72 hours bone marrow
735 cells were harvested and analysed by flow cytometry. (B) Gating strategy for analysis of the progenitors in the bone marrow from
736 a representative mouse and median fluorescence intensity of CD62L expression in multipotent progenitors in control and LPS treated
737 mice. Each point is a mouse, N= 4 mice. Statistical significance was assessed using a Mann-Whitney test (C) Quantification of the
738 percentage of the different cell subsets in control (black) and LPS treated (red) mice. Normality of the data was assessed using a Shapiro-
739 Wilk test and significance was assessed using a T-test. N = 13 mice and data was pooled from 2 independent experiments. Barplots
740 represent the mean expression value and each point represents a different mouse. (D) scRNAseq data reanalysed from Haltalli et al
741 (2020), where mice were treated with vehicle control or *P. berghei*. 7 days post infection cKit+ progenitors were purified from the bone
742 marrow of each group and processed for scRNAseq profiling. Data are represented using a density projection of the cell abundances on
743 to a PCA embedding of the data. Cell type annotations of the data were taken from the original publication. Control samples have 14193
744 cells and infected samples have 13905 cells. (E) *Sell* (gene encoding CD62L) normalised gene expression projected onto the PCA
745 embedding of the data. (F) Boxplot showing MetaFate signature expression score in HSPCs (defined as primitive HSPCs in the original
746 article) from control (black) and infected (red) mice. Pairwise comparisons were made using a Student's T-test. Boxplot showing mean
747 and sd over cells? (G) Same as F but for all cKit+ hematopoietic progenitors from control and infected mice. (H) The proportion of cells
748 in control and infected mice that have non-zero expression of *Sell*. Boxplots represent the interquartile range and median values and
749 whiskers represent the 5th and 95th percentile of the data.

750



751

752 **Figure 5: The CD62L^{hi} MPP compartment reconstitutes the myeloid compartment following bone marrow transplantation** (A) MPPs
 753 were purified from male donor B6 WT mice by FACS and infected with the LG2.2 lentiviral barcoding library for 6 hours. Transduced cells
 754 were transplanted into 4 sublethally irradiated (6Gy) control recipient mice (male littermate controls) and 3 weeks later CD62L^{neg/hi}
 755 MPPs, CD19⁺ B cells, CD44⁺ Ter119⁺ erythrocytes, and CD11b⁺ myeloid cells, as well as CD62L^{high} or ^{neg} MPPs from the bone marrow
 756 were sorted and processed for targeted sequencing of lentiviral lineage barcodes. (B) read abundance of barcode in the CD62L^{neg} enriched
 757 and the CD62L^{high} enriched HSPC (LSK) fraction, each dot is a barcode and the axis are transformed using the hyperbolic arcsin function.
 758 Barcodes that had more than 95% of its reads in either the CD62L^{low} or the CD62L^{pos} MPPs were classified as CD62L^{shared} (light grey; 124
 759 barcodes), CD62L^{neg} enriched (dark grey; 30 barcodes) and CD62L^{high} enriched (blue; 18 barcodes) MPPs. (C) Number of cells of a given
 760 lineage (the myeloid, erythroid and B-cell lineages) produced per barcode, is clone size, for the barcode categories CD62L^{neg} enriched (grey)
 761 and CD62L^{high} enriched (blue) MPP subsets. Each point represents a distinct barcode. Statistical comparisons were made using a Mann-
 762 Whitney test. (D) lineage bias value for the myeloid, erythroid and B-cell lineages for barcodes in CD62L^{neg} enriched (grey) and CD62L^{high}
 763 enriched (blue) MPP subsets. Lineage bias represents the relative frequency of each barcode across the 3 mature cell lineage. Each point
 764 represents a distinct barcode. Statistical comparisons were made using a Mann-Whitney test. (E) Unsupervised clustering (using the
 765 Euclidean distance) and heatmap visualisation showing hyperbolic arcsin transformed barcode abundances for CD62L^{neg} enriched and
 766 CD62L^{high} enriched barcodes. Every row is a barcodes and every column is a cell type. All Boxplots represent the median and interquartile
 767 range with whiskers extending to the minimum and maximum values. N= 4 mice.

768

769

770

771