Dissecting infant leukemia developmental origins with a hemogenic gastruloid model

Denise Ragusa[1,2], Chun-Wai Suen[3*], Gabriel Torregrosa-Cortés[4*], Fabio Pastorino[5*], Ayona Johns[1,2], Ylenia Cicirò[1,2], Liza Dijkhuis[1,12], Susanne van den Brink[6,7], Michele Cilli[8], Connor Byrne[1,13], Giulia-Andreea Ionescu[1,14], Joana Cerveira[9], Kamil R. Kranc[10,11], Victor Hernandez-Hernandez[1,2], Mirco Ponzoni[5], Anna Bigas[6,7], Jordi Garcia-Ojalvo[4], Alfonso Martinez Arias[4], Cristina Pina[1,2¶]*

[1]College of Health, Medicine and Life Sciences, [2]Centre for Genome Engineering and Maintenance, Brunel University London, UK; [3]Department of Genetics, University of Cambridge, UK; [4]Department of Medicine and Life Sciences, Universitat Pompeu Fabra, Spain; [5]Laboratory of Experimental Therapies in Oncology, IRCCS Istituto Giannina Gaslini, Genova, Italy; [6]Program in Cancer Research, Hospital de Mar Research Institute, Barcelona, Spain; [7]Josep Carreras Leukemia Research Institute, Barcelona, Spain; [8]Animal Facility, IRCCS Policlinico San Martino, Genova, Italy; [9]Department of Pathology, University of Cambridge, UK; [10]Centre for Haemato-Oncology, Barts Cancer Institute, Queen Mary University of London, UK; [11]Centre for In Vivo Modelling, Institute of Cancer Research, Sutton, UK.

[12]Current address: Sanquin Research Laboratory, Department of Hematopoiesis, Amsterdam, The Netherlands

[13]Current address: Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK

[14]Current address: Division of Infection and Immunity, Cardiff University School of Medicine, Cardiff, Henry Wellcome building, CF14 4XN

*These authors contributed equally to this work

¶Lead contact. Correspondence: cristina.pina@brunel.ac.uk. Brunel University London, Kingston Lane, Uxbridge UB8 3PH, United Kingdom.

ABSTRACT

Current in vitro models of developmental blood formation lack spatio-temporal accuracy and weakly replicate successive waves of hematopoiesis. Herein, we describe a mouse embryonic stem cell (SC)-derived 3D hemogenic gastruloid (hGx) that captures multi-wave blood formation, progenitor specification from hemogenic endothelium (HE), and generates hematopoietic SC precursors capable of short-term engraftment of immunodeficient mice upon maturation in an adrenal niche. We took advantage of the hGx model to interrogate the origins of infant acute myeloid leukemia (infAML). We focused on MNX1-driven leukemia, representing the commonest genetic abnormality unique to the infant group. Enforced MNX1 expression in hGx promotes the expansion and in vitro transformation of yolk sac-like erythroid-myeloid progenitors (EMP) at the HE-to-hematopoietic transition to faithfully recapitulate patient transcriptional signatures. By combining phenotypic, functional and transcriptional profiling, including at the single-cell level, we establish the hGx as a useful new model for the study of normal and leukemic embryonic hematopoiesis.

INTRODUCTION

Blood formation in the embryo is a multi-stage process that develops across multiple cellular niches, reflecting a complex hierarchy of tissue interactions. It involves sequential production of distinct cell types in so-called hematopoietic waves, separated in time and space (Lacaud and Kouskoff, 2017; Costa, Kouskoff and Lacaud, 2012; Medvinsky, Rybtsov and Taoudi, 2011; Dzierzak and Bigas, 2018). The first wave of hematopoiesis produces unipotent red blood cell and macrophage precursors in the yolk sac (YS) from an angioblast precursor which can also form endothelium; it occurs at mouse embryo day (E)7-E7.5 (Fujimoto *et al.,* 2001; McGrath and Palis, 2005; Moore and Metcalf, 1970). The subsequent wave, known as definitive, produces: first, oligopotent erythro-myeloid progenitors (EMPs) in the YS (E8-E8.5); and later myelo-lymphoid progenitors (MLPs – E9.5-E10), multipotent progenitors (MPPs – E10-E11.5), and hematopoietic stem cells (HSCs – E10.5-E11.5), in the aorta-gonad-mesonephros (AGM) region of the embryo proper (Ivanovs *et al.,* 2011; McGrath *et al.,* 2015; Palis *et al.,* 1999; Medvinsky and Dzierzak, 1996; Medvinsky, Rybtsov and Taoudi, 2011). Definitive blood waves are characteristically specified from a specialised hemogenic endothelium (HE) through a process of endothelial-to-hematopoietic transition (EHT) (Zovein *et al.,* 2008; Li, W. *et al.,* 2005). HSC specification from HE is the last event in hematopoietic development and occurs in characteristic cell clusters found on the ventral wall of the dorsal aorta (Ivanovs *et al.,* 2011; Boisset *et al.,* 2010). All blood cell types enter circulation from their original locations and migrate to the fetal liver (FL) (Ciriza *et al.,* 2013; Ghiaur *et al.,* 2008). Most embryonic blood progenitors are transient and sustain blood production exclusively during embryonic development, through differentiation. In contrast, HSCs make little or no contribution to embryonic and fetal blood: instead, they proliferate (i.e. self-renew) to expand their numbers. At least a subset of HSC migrates to the bone marrow (BM) niche at the end of gestation and balances self-renewal and differentiation to maintain blood production throughout the entire post-natal life (Bowie *et al.,* 2006; Kikuchi and Kondo, 2006; Ganuza *et al.,* 2022; Yokomizo *et al.,* 2022).

Blood progenitors and HSCs can be targeted by genetic alterations that confer clonal advantage and/or drive initiation and development of leukemia. Leukemia is the most common malignancy and the most common cause of cancer-related death in the

72 paediatric age group (Steliarova-Foucher *et al.,* 2017; Dong *et al.,* 2020). Several
73 paediatric leukemias originate *in utero*, and a subset of these are driven by mutations
74 exclusive to the paediatric group (Masetti *et al.,* 2015; Cazzola *et al.,* 2021). Such
75 pediatric-exclusive mutations target blood cell types restricted to development or
76 depend on signals that are specific to embryonic hematopoietic niches, effectively
77 defining distinct biological entities to adult malignancies (Cazzola *et al.,* 2021; Bolouri
78 *et al.,* 2018; Chaudhury *et al.,* 2018). Dissection of these forms of developmental
79 leukemia is confounded by the inability to robustly identify and isolate their embryonic
80 cells of origin, and/or to fully recapitulate the disease in post-natal cells, which leads
81 to failure in generating relevant targeted therapies (Alexander and Mullighan, 2021).
82 One such example is t(7;12)(q36;p13) acute myeloid leukemia (AML) driven by
83 rearrangement and overexpression of the *MNX1* gene locus: *MNX1*-rearranged
84 (MNX1-*r*) AML is restricted to infants under the age of 2 and carries a high risk of
85 relapse and poor survival (Ragusa *et al.,* 2023). *MNX1* is a developmental gene
86 required for specification of motor neurons and pancreatic acinar cells (Thaler *et al.,*
87 1999; Harrison *et al.,* 1999). MNX1 is not known to participate in embryonic
88 haematopoiesis, however it has been discordantly described to be expressed in some
89 subsets of adult haematopoietic cells (Deguchi and Kehrl, 1991; Nagel *et al.,* 2005;
90 Wildenhain *et al.,* 2012; Taketani *et al.,* 2008). It is ectopically expressed in AML cells
91 carrying one of a variable set of t(7;12) translocations which place the normally silent
92 *MNX1* gene under the control of *ETV6* regulatory elements (Weichenhan *et al.,* 2023;
93 Bousquets-Muñoz *et al.,* 2024; Ragusa *et al.,* 2023; Ballabio *et al.,* 2009). *ETV6* is
94 required for hematopoietic cell specification from HE via VEGF signalling (Ciau-Uitz *et*
95 *al.,* 2010), potentially positioning *MNX1* within a critical embryonic hemogenic
96 regulatory network.

97 Faithful *in vitro* recapitulation of developmental blood cell types and their embryonic
98 niches has the potential to dissect the cellular and molecular mechanisms presiding
99 at initiation of developmental leukemia, reveal prognostic biomarkers, and indicate
100 therapeutic vulnerabilities which are otherwise challenging to test in the embryo.
101 Despite widespread use of embryonic stem (ES) cell and induced pluripotent stem
102 (iPS) cell-based models of blood specification, a system which recapitulates the spatial
103 and temporal complexity of embryonic blood formation within the respective time-

104 dependent niches is still lacking (Dijkhuis *et al.,* 2023). In the last few years, gastruloid
105 models have emerged as robust avatars of early development, self-organising
106 processes of symmetry breaking, elongation, multi-axis formation, somitogenesis and
107 early organogenesis, with remarkable similarity to embryo processes in space and
108 time (van den Brink, Susanne C and van Oudenaarden, 2021; Beccari *et al.,* 2018;
109 van den Brink, Susanne C *et al.,* 2020). Gastruloids are grown individually in a multi-
110 well format that facilitates tracking and screening of developmental processes.
111 Relevant to blood formation, a variant gastruloid culture system, which successful
112 organises a heart primordium (Rossi *et al.,* 2021), was also able to recapitulate YS
113 blood formation with erythroid and myeloid cellularity (Rossi *et al.,* 2021; Rossi *et al.,*
114 2022).

115 Herein, we describe a new protocol of gastruloid formation, i.e. hemogenic gastruloids
116 (hGx), that captures YS and AGM-like waves of definitive blood progenitor
117 specification from HE. The AGM-like wave includes cells capable of short-term
118 hematopoietic engraftment of the spleen and BM of immunodeficient mice, overall
119 suggesting nearly full EMP-to-MPP/HSC recapitulation of embryonic blood
120 development. We harnessed the embryonic haematopoietic context provided by hGx
121 for modelling MNX1-*r* infant leukemia by proxy of *MNX1* overexpression. MNX1-
122 overexpressing hGx showed expansion of cells at the HE-to-EMP transition, which
123 became susceptible to transformation *in vitro* and recapitulated MNX1-*r* AML patient
124 signatures, placing *MNX1*'s leukemogenic effects at a very early stage of
125 developmental blood formation.

126

127 **RESULTS**

128 **HGxs capture cellularity and topography of developmental blood formation**

129 With the aim of recapitulating embryonic haematopoiesis with spatial and temporal
130 accuracy, we modified existing gastruloid formation protocols to promote the
131 specification of hemogenic mesoderm derivatives and extend developmental time
132 beyond the E8.0-equivalent of early developmental gastruloids (Van den Brink,
133 Susanne C *et al.,* 2014) and the E9.0 timepoint of cardiogenic gastruloids (Rossi *et*
134 *al.,* 2022). We established a 216h hGx protocol (**Fig. 1A**) using a *Flk1-GFP* mouse ES

135 cell line (Jakobsson *et al.,* 2010), which allows the tracking of early hemato-endothelial
136 specification by expression of the *Flk1* (i.e. *Kdr*) *locus*. Activation of the *Flk1*
137 expression, visible by fluorescence of the GFP tag, necessitated induction of the TGF-
138 β signalling pathway by Activin A, in addition to the characteristic gastruloid patterning
139 via WNT activation by CHI99021 supplementation at 48h (**Fig. S1A**). We sought to
140 promote hemato-endothelial programmes through addition of VEGF and FGF2 at 72h
141 (Sroczynska *et al.,* 2009) (**Fig. 1A**), and observed extension of a polarised *Flk1-GFP*-
142 expressing cellular network (**Fig. 1B**), which included VE-cadherin$^+$ C-Kit$^+$ cells
143 suggestive of EHT (**Fig. 1C-D**). This process of HE specification and EHT peaked at
144 120-144h; it was followed by a surge in CD41$^+$ candidate hematopoietic cells at 144h
145 (**Fig. 1D**). In our experience, observation of early polarised activation of the *Flk1 locus*
146 at 96h (**Fig. 1B**) was predictive of CD41$^+$ detection 2 days later, suggesting that TGF-
147 β activation in early patterning is required for later hematopoietic development,
148 putatively by modifying mesoderm specification. We extended the protocol by
149 incorporation of a 24-hour pulse of sonic hedgehog (Shh) between 144-168h to mimic
150 the aortic patterning that precedes cluster formation (**Fig. 1A**) (Rybtsov *et al.,* 2014).
151 Continued exposure to VEGF after the Shh pulse resulted in specification of CD45$^+$
152 hematopoietic cells progenitors (**Fig. 1D-E**), which increased after 168h. The
153 proportion of CD45$^+$ cells was enhanced by the addition of SCF, FLT3L and TPO (**Fig.**
154 **S1B-C**), a combination of cytokines normally used for HSC maintenance; conversely,
155 it did not require continued addition of FGF2 after 168h (**Fig. S1D**), which was omitted.
156 CD45$^+$ cells appeared in a small number of 'cluster'-like structures adjacent to CD31$^+$
157 endothelium (**Fig. 1E**), which were reminiscent of HSC and progenitor cell budding
158 from the aortic HE, and suggested recapitulation of hemogenic tissue architecture.
159 Time-dependent specification of C-Kit$^+$, CD41$^+$ and CD45$^+$ cells in hGxs could be
160 obtained with other mouse ES cell lines, including commonly used E14Tg2a (E14)
161 cells (**Fig. S1E-G**), with successful specification of comparable frequencies of CD45$^+$
162 progenitors (**Fig. S1H**), attesting to the reproducibility of the protocol.

163 Overall, our extended hGx model could capture successive specification of HE, CD41$^+$
164 and CD45$^+$ cells, with temporal and spatial congruence relative to the embryo, thus
165 showing initial promise as a tool for the exploration of normal and perturbed
166 developmental hematopoiesis.

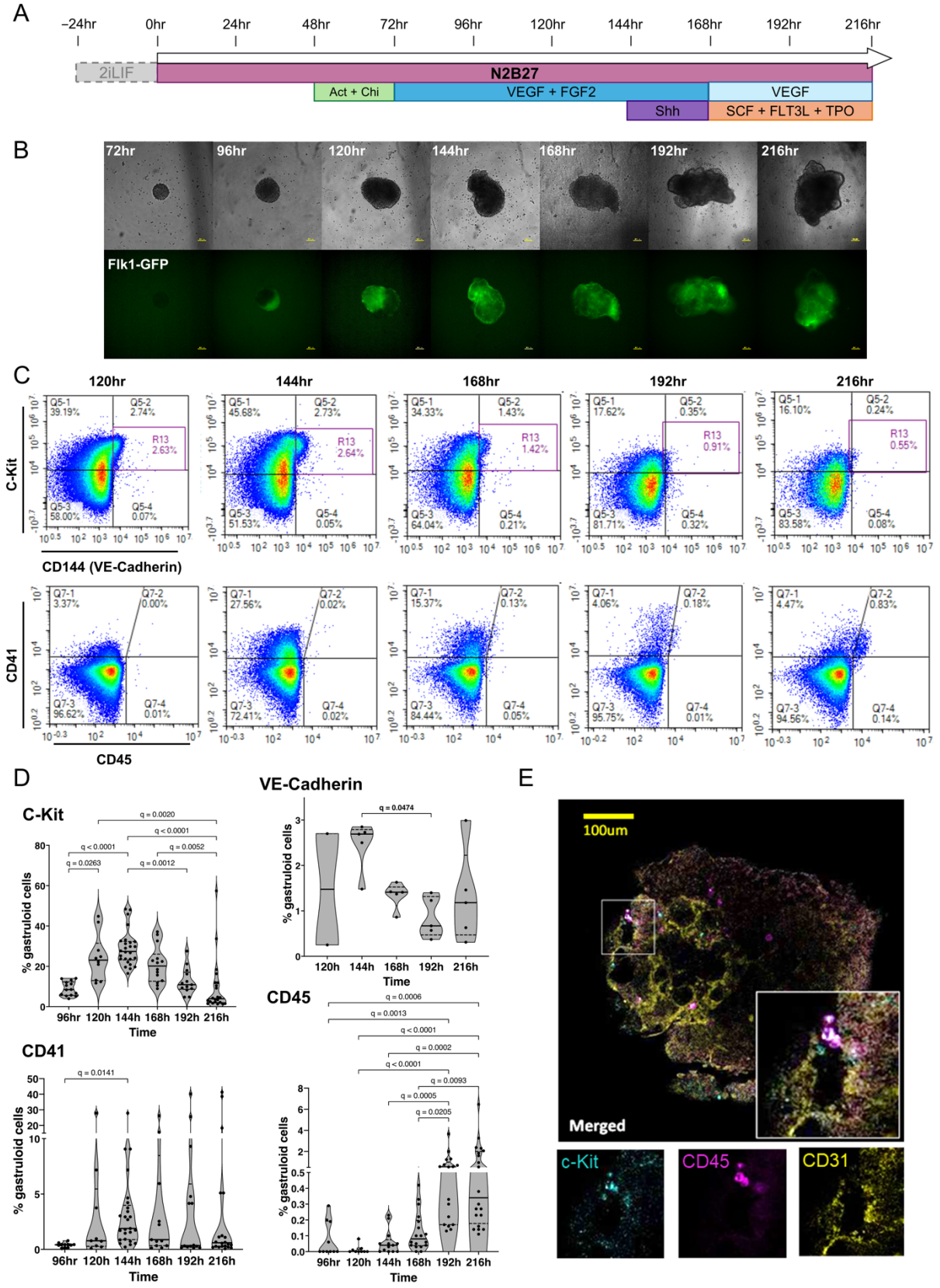
A
−24hr
0hr
24hr
48hr
72hr
96hr
120hr
144hr
168hr
192hr
216hr
2iLIF
N2B27
Act + Chi
VEGF + FGF2
VEGF
Shh
SCF + FLT3L + TPO
B
72hr
96hr
120hr
144hr
168hr
192hr
216hr
Flk1-GFP
C
120hr
144hr
168hr
192hr
216hr
C-Kit
CD144 (VE-Cadherin)
CD41
CD45
D
C-Kit
VE-Cadherin
CD41
CD45
% gastruloid cells
Time
96hr 120hr 144hr 168hr 192h 216hr
E
100um
Merged
c-Kit
CD45
CD31

Figure 1

169

170 **HGx capture time-dependent emergence of YS-like and AGM-like progenitors**
171 **from hemogenic endothelium and contain elements of hemogenic niches**

172 To better characterize the extent and progression of developmental hematopoiesis in
173 hGx, we performed single-cell RNA-sequencing (scRNA-seq) time-course analysis of
174 gastruloid cell specification. We sorted cells from 2 independent hGx cultures at 120,
175 144, 168, 192 and 216h, and profiled a total of 846 cells using the Smart-Seq2 protocol
176 (Picelli *et al.,* 2014) (**Fig. S2A**). We analysed unfractionated live single hGx cells at
177 120-192h and enriched the dataset with sorted CD41$^+$ cells at their peak time of 144h,
178 and sorted CD45$^+$ cells at 192 and 216h. At timepoints of significant enrichment in
179 CD41$^+$ (144h) and CD45$^+$ cells (192h) (**Fig. 1D**), we also sorted C-Kit$^+$ cells. Library
180 preparation and sequencing generated an average of 120000 reads/cell, which were
181 mapped to an average of 4000 genes/cell, with minimal signs of cell stress / death as
182 measured by mitochondrial DNA fraction (**Fig. S2B**). Read and gene counts were
183 similar between biological replicates, cell types and at different time points (**Fig. S2C**).
184 The exception was 120h unfractionated hGx cells, which despite similar sequencing
185 depth (average read counts), were mapped to twice the number of genes (**Fig. S2C**),
186 possibly reflecting multi-gene program priming at the onset of hemogenic specification.
187 We selected highly varying genes (HVG) before principal component analysis (PCA)
188 dimensionality reduction and retained the most relevant dimensions. In the PCA
189 reduced space, we constructed a KNN graph and used uniform manifold
190 approximation and projection (UMAP) to visualize the data on 2 dimensions, looking
191 for cell communities using Leiden clustering (**Fig. 2A**). We identified 12 cell clusters of
192 which 2 (clusters 0 and 5) almost exactly mapped to C-Kit$^+$ cells (also Sca-1$^+$), 2
193 (clusters 1 and 8) contained CD45$^+$ sorted cells, and cluster 4 uniquely captured the
194 CD41$^+$ cells observed at 144h (**Fig. 2A**). Although some unfractionated cells
195 overlapped with the hemogenic cell clusters (**Fig. 2A** - center), most occupied different
196 transcriptional spaces, reflecting the frequency of hematopoietic cells and suggesting
197 the presence of potential hemogenic niches.

198 To explore tissue and lineage affiliation of the different clusters, we performed
199 differential gene expression against all other cells using Wilcoxon ranking test, and
200 established cluster classifier gene lists (**Supplemental File S1**). Top classifier genes

201 for clusters 0 and 5 had endothelial affiliation (**Supplemental File S1**), and the cellular
202 composition of the clusters broadly reflected their time signature (**Fig. 2A**): cluster 5
203 comprised 144h cells and cluster 0 included most C-Kit$^+$ cells at 192h. Cluster 5
204 occupied the vicinity of 144h-CD41$^+$ cells (**Fig. 2A**), which carried an erythroid-biased
205 signature (**Supplemental File S1**). On the other hand, 192/216h-CD45$^+$ cells in cluster
206 8 expressed myeloid and lymphoid genes (**Supplemental File S1**). The remaining
207 clusters largely corresponded to time-restricted unfractionated cells (**Fig. 2A** - right),
208 with clusters of cells present at later time-points matching mesenchymal stromal cell
209 and autonomic neuron signatures (**Fig. S2D**), potentially configuring incipient
210 formation of niches relevant to production of HSC (Kapeni *et al.,* 2022; Fitch *et al.,*
211 2012).

212 To further characterise the identity of hGx cells, we interrogated lineage priming in
213 individual hemato-endothelial cells (**Fig. 2B-C**), to probe the presence of putative
214 multilineage YS (EMP) and AGM (MLP/MPP) progenitors, and HE cells. The
215 heatmaps (**Fig. 2B-C**) show expression of lineage markers and regulators, with
216 unsupervised clustering of cells (**Fig. S2E-F**) and genes. The gene list used was
217 identical in both analyses, with only expressed genes shown. Concordance between
218 sorting antigens and expression of the respective genes (*Kit*, *Itga2b* [CD41] and *Ptprc*
219 [CD45]) was partial potentially reflecting low levels of transcript expression in the hGx
220 system. Summary of endothelial, co-expressing erythro-myeloid (EM) and co-
221 expressing myelo-lymphoid (ML) programs (**Fig. 2D**) clearly associates the first with
222 C-Kit$^+$ and the latter with CD45$^+$ cells; CD41$^+$ cells capture a heterogeneous
223 population, which includes a proportion of cells co-expressing erythroid and myeloid
224 genes that constitute candidate EMPs.

225 Clustering of cells sorted as C-Kit$^+$ did not separate the 144h and 192h time-points
226 (**Figs. 2B** and **S2E**), indicating that the relative separation on the global UMAP was
227 not based on endothelial identity genes. We observed widespread expression of
228 arterial genes (e.g. *Flt1*, *Efnb2*) with relative reciprocity with venous endothelial genes
229 (e.g. *Flt4*) supporting specification of both types of endothelia (**Fig. 2B**). There was
230 limited co-expression of hematopoietic genes, but genes associated with endothelial-
231 to-hematopoietic transition (EHT), e.g. *Gata2*, *Notch1*, and *Procr*, were significantly

232 more frequent within the C-Kit$^+$ population (**Fig. 2E**). Specific associations of *Cd44*
233 and *Procr* with the 192h time-point (**Fig. 2F**) putatively represent aorta-like EHT.

234 Clustering of CD41$^+$ and CD45$^+$ cells (**Fig. 2C** and **S2F**) more clearly separates unique
235 hematopoietic signatures. A subset of cells, predominantly CD41$^+$, expresses high
236 levels of globin genes. While embryonic β-globin chains are highly expressed, there
237 is a gain in fetal *Hbb-y* and adult *Hbb-bt* and *Hbb-bs* (C57Bl/6 strain-specific β major
238 and β minor, respectively) in CD45$^+$ cells at 216h, indicative of AGM-like progenitor
239 identity (**Fig. 2C**). 192 and 216h cells with high expression of *Ptprc* have an MLP profile
240 (**Fig. 2C-D**), co-expressing myeloid and lymphoid genes including immunoglobulin
241 chains. There is progression of the lymphoid signature between 192 and 216h (**Fig.**
242 **2D** and **S2G**). Other CD45$^+$ cells have low-frequency expression of HSC-associated
243 transcription factors, including *Myb*, *Hlf* and *Hoxa9* (**Fig. 2G** and **S2I-J**), suggesting
244 incipient specification of AGM-like MPP and pre-HSC, albeit with incomplete
245 programming.

246 Altogether, single-cell RNA-seq data suggest that hGx specifies YS-like and AGM-like
247 progenitors in a time-dependent manner, with concomitant emergence of
248 hematopoietic-supporting niche cells. The data suggested putative incipient
249 specification of rare candidate MPP or pre-HSC, which we investigated further in
250 comparison with mouse AGM datasets and at a functional level.

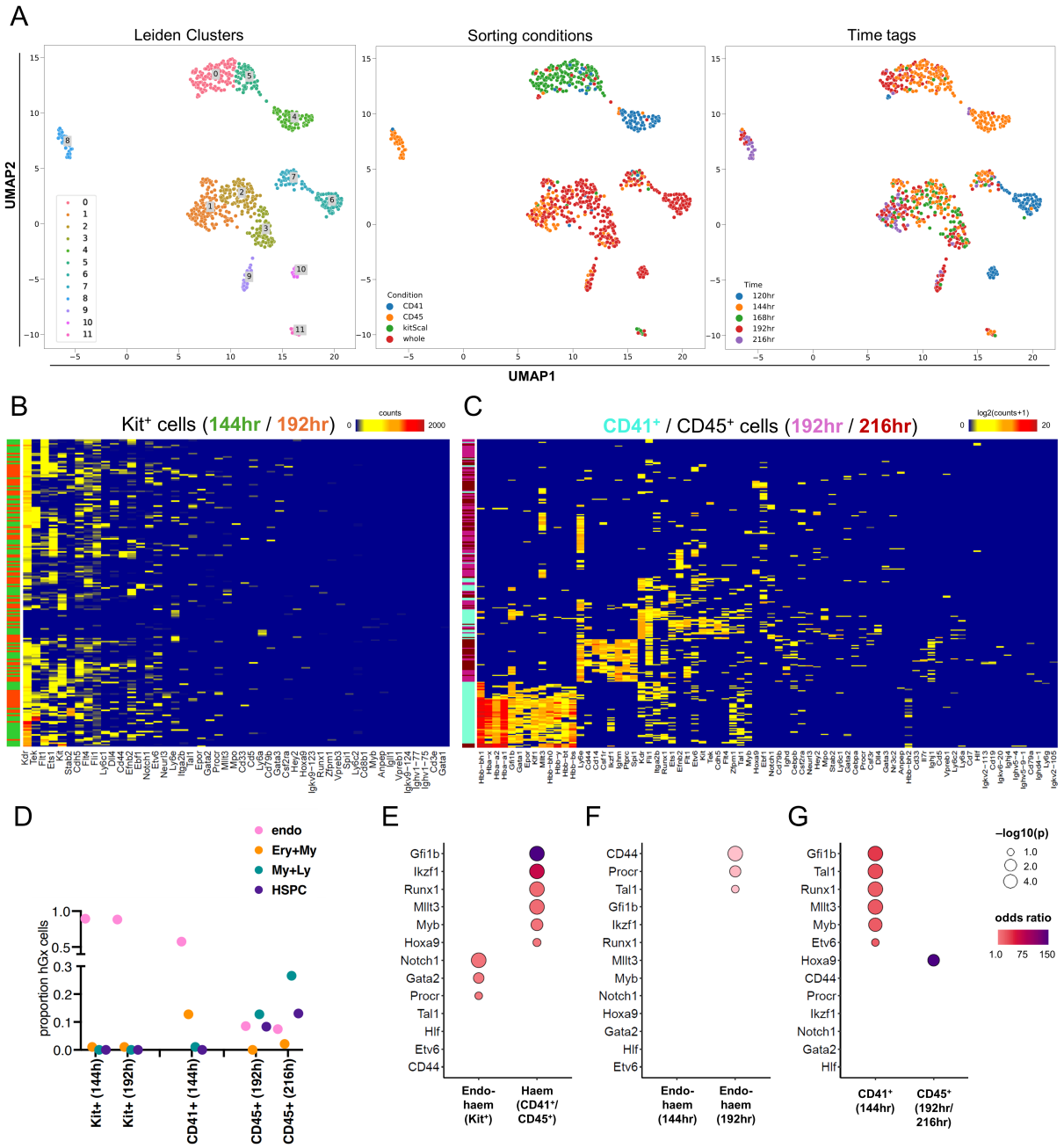
A
Leiden Clusters
Sorting conditions
Time tags
UMAP2
UMAP1
Condition
CD41
CD45
kitScal
whole
Time
120hr
144hr
168hr
192hr
216hr
B
Kit⁺ cells (144hr / 192hr)
counts
C
CD41⁺ / CD45⁺ cells (192hr / 216hr)
log2(counts+1)
D
endo
Ery+My
My+Ly
HSPC
proportion hGx cells
Kit+ (144h)
Kit+ (192h)
CD41+ (144h)
CD45+ (192h)
CD45+ (216h)
E
Endo-haem (Kit⁺)
Haem (CD41⁺/ CD45⁺)
F
Endo-haem (144hr)
Endo-haem (192hr)
G
CD41⁺ (144hr)
CD45⁺ (192hr/ 216hr)
−log10(p)
odds ratio

Figure 2

253
254 **Late-stage hGx cells transcriptionally resemble mouse AGM progenitors**

255 We made use of a dataset recently published by one of us (Thambyrajah *et al.,* 2024)
256 that interrogates the mouse AGM at the point of HSC emergence from the dorsal aorta
257 HE. In this study, a combination of surface markers and *Gfi1* locus reporting of EHT,
258 tested functionally through transplantation, identified HE and HSC-enriched clusters
259 (**Fig. S3A** - clusters 1 and 2, respectively, with cluster 4 a putatively heterogenous
260 cluster at the transition between the 2 states). We projected cell transcription profiles
261 from hGx to AGM using a version of scmap algorithm (Kiselev, Yiu and Hemberg,
262 2018) with 1-cell and 3-cell neighbourhoods for lower and higher levels of confidence
263 on similarity projections (respectively gray and black circles on **Fig. S3A** - left). The
264 numbers of cluster-specific hGx cells projected onto AGM clusters is quantified in **Fig.**
265 **S3B**. Top differentially expressed genes (DEG) between AGM HE and HSC-projected
266 hGx cells are represented in Fig. **S3C** and **Supplemental File S2**.

267 With respect to HE specification, we observed that a subset of hGx C-Kit$^+$ cells present
268 in endothelial clusters 5 and 0 (**Fig. 2A**) projected onto the AGM HE-identity cluster 1
269 (**Fig. S3A**), with more frequent similarity amongst hGx 192h cluster 0 (41%) than 144h
270 cluster 5 (21%). Overall, this supports the notion that specification of endothelial and
271 HE cells in hGx follows with time-dependent developmental progression into putative
272 AGM-like HE. Lineage identity genes supported the nature of HE-projected hGx cells
273 (**Fig. S3D**), which expressed *Gata2* but not *Runx1*, *Myb*, or *Gfi1b*, indicating that EHT
274 occurs downstream of these cells. We then asked if any of the hGx cells in hemopoietic
275 clusters bore resemblance to cells from the HSC-enriched AGM cluster 2. We
276 observed that most (85%) hGx CD45$^+$ cells in cluster 8 (**Fig. 2A**) projected onto the
277 HSC-enriched cluster 2, which also showed similarity with 76% of CD41$^+$-enriched
278 hGx cluster 4 cells. The relative positioning of hGx cells cluster 4 and cluster 8
279 progressively further away from the AGM HE cluster was in line with the time-
280 dependent specification in the hGx. However, the lineage marker programs (Fig. **S3D**)
281 expressed by the hGx cells, although definitive in nature (e.g. there is expression of
282 immunoglobulin chains and adult β-globins), lacked detection of HSC genes like *Hlf*,
283 *Gfi1*, or *Hoxa9*. On the other hand, *Mllt3* was present (Calvanese *et al.,* 2019), even
284 in the absence of an erythroid signature (Pina *et al.,* 2008), suggesting a developing

285 pre-HSC program compatible with the co-expression of multi-lineage markers (**Fig.**
286 **S3D**).
287
288 **Late-stage hGx cells can engraft hematopoietic tissues upon maturation**
289 We inspected the cellularity of hGx at the 216h time-point more closely through a
290 combination of multi-parameter flow cytometry (**Fig. 3A-B**), colony-forming progenitor
291 assays (CFC) (**Fig. 3C-D**) and *in vivo* engraftment of haematopoietic tissues (**Fig. 3E-**
292 **G**). In order to facilitate the analysis of engrafted animals, we engineered constitutive
293 expression of BFP from the *Rosa26* locus in the *Flk1-GFP* mES cell line (**Fig. S4A**),
294 which we also used in some of the *in vitro* analyses. Phenotypic characterisation of
295 216h-hGx (4 replicates, 1 with *Rosa26-BFP::Flk1-GFP* line) (**Fig. 3A**, **S4B**) showed
296 the presence of pre-HSC-like $CD45^{+}CD41^{lo}$c-Kit^{+}VE-$Cadherin^{+}$ cells (1.8% of all
297 $CD45^{+}CD41^{lo}$ cells) (**Fig. 3B**), corresponding to an estimated 8 cells per hGx. Analysis
298 of the hGx progenitor content using multipotential CFC assays (**Fig. 3C-D**) revealed
299 the presence of >50 CFC/hGx, of which an average of 7 have mixed-lineage
300 granulocytic-monocytic-erythroid (GEM) potential, numerically matching the estimated
301 number of pre-HSC. Other progenitors were predominantly granulo-monocytic (GM)
302 and erythroid (E) progenitors (**Fig. 3C-D**). Differentiated cells of G, M and E lineages
303 (**Fig. S4C-D**) were also detectable in 216h hGx, putatively reflecting the progeny of
304 EMP-like cells specified earlier at the 144h timepoint.
305 As an ultimate test of the presence of pre-HSC, we tested whether late-stage 216h
306 hGx contained cells capable of hematopoietic engraftment of immunodeficient mice.
307 In early experiments, we transplanted dissociated hGx cells into irradiated
308 immunocompetent C57Bl/6 recipients, sub-lethally irradiated immunodeficient NSG
309 mice, or non-conditioned NSGW41 hosts, but failed to detect signs of engraftment of
310 peripheral blood, spleen (Spl) or bone marrow (BM) at 12 days (including absence of
311 spleen colony formation), at 4-6 weeks (short-term engraftment), or beyond 8-10
312 weeks (long-term engraftment) by either flow cytometry detection of CD45.2 or by
313 genomic DNA (gDNA) PCR analysis of the *Flk1-GFP* locus. However, given the
314 absence of clear HSC-like transcriptional programs in 216h $CD45^{+}$ cells, as analysed
315 by scRNA-seq (**Fig. 2** and **S2**), and in light of the late specification of mesenchymal or

316 sympathetic neuronal cells (**Fig. 2A** and **S2D**) putatively supportive of HSC formation,
317 we hypothesized that hGx pre-HSC required additional maturation in a relevant
318 supportive niche. We thus tested implantation of undissociated hGx in the adrenal
319 gland of *Nude* mice (**Fig. 3E**), a topography commonly used to support tumor
320 development, and recently shown to support adult haematopoietic development
321 (Schyrr *et al.,* 2023). Given the coincidence of CD45 isoform between *Flk1-GFP*
322 mouse ES cells and *Nude* mice, we made use of the *Rosa26-BFP::Flk1-GFP* mES
323 cell line (**Fig. S3A**). We implanted 3 hGx each unilaterally in the adrenal gland of
324 unconditioned recipient mice. Control mice were injected with an equivalent volume of
325 PBS. Animals were sacrificed at 24 hours, 12 days, and 4 weeks after implantation,
326 with 5 experimental animals and 1 control collected at each timepoint. We checked
327 Spl and BM BFP engraftment at the different timepoints and analyzed the morphology
328 of adrenal glands receiving implants (**Fig. 3F**). We did not observe BFP engraftment,
329 spleen enlargement or colony development, or macroscopic changes in the adrenal
330 gland at the earlier timepoints. However, 3 of 5 adrenal glands were enlarged, and
331 their histology suggested *in situ* development of the hGx (**Fig. 3F**), particularly
332 mesodermal tissues – somite derivatives (**Fig. 3F** - left, red box), and lateral plate and
333 intermediate mesoderm derivatives, respectively blood cells and renal tubules (**Fig.**
334 **3F** - left, green box). The control, PBS-injected glands showed normal tissue
335 architecture (**Fig. 3F** - right). It should be noted that none of the specimens contained
336 glomeruli in addition to the renal tubules, pointing to an origin from the hGx, and not
337 from the adjacent kidney tissue. BFP engraftment of the Spl and BM by flow cytometry
338 was very low level albeit consistently above control (**Fig. S4E**), and we opted for
339 analysing engraftment by PCR analysis of gDNA using primers directed against the
340 BFP insert (**Fig. S4A**). We measured the sensitivity of the amplification through serial
341 dilution of *Flk1-GFP Rosa26-BFP* mES cell gDNA into that of human K562 cells, and
342 determined detection of 10pg/100ng gDNA, equivalent to 2-3 cells (**Fig. S4F**). PCR
343 analysis of spleen and BM gDNA detected *BFP* amplicon in both tissues in 3 of the 5
344 samples (#11, 12 and 15) (**Fig. 3G**). We sampled the gDNA of each recipient 4-5 times
345 and detected the presence of the amplicon in only a fraction of the replicate samples
346 of the positive animals, reflecting the low-level detection by flow cytometry. Replicate
347 sampling of the control was consistently negative (**Fig. 3G**). Sorting of erythroid

348 Ter119^{+} and lympho-myeloid CD45^{+} cells (**Fig. S4G**) showed PCR engraftment in both
349 lineage fractions (**Fig. S4H**).

350 Altogether, the data suggest that 216h hGx generate AGM-like pre-HSC capable of at
351 least short-term multilineage engraftment upon maturation. The low-level production
352 of engrafting cells recapitulates their rarity *in vivo*, in agreement with the embryo-like
353 qualities of the gastruloid system, which accommodates modular qualitative bias, but
354 not quantitative changes in tissue specification.

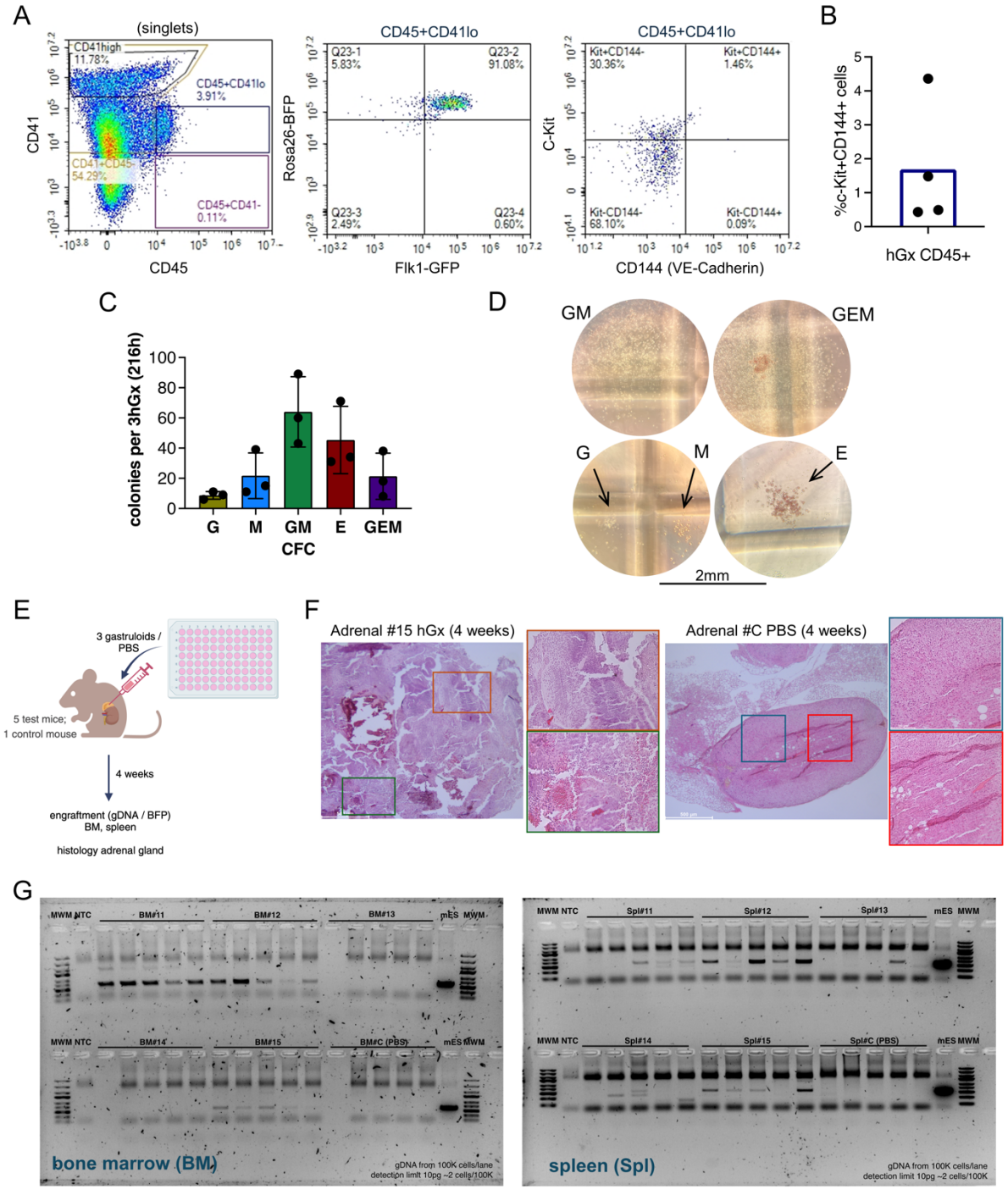
A
(singlets)
CD41high
11.78%
CD45+CD41lo
3.91%
CD41+CD45-
54.29%
CD45+CD41-
0.11%
CD41
CD45
CD45+CD41lo
Q23-1
5.83%
Q23-2
91.08%
Q23-3
2.49%
Q23-4
0.60%
Rosa26-BFP
Flk1-GFP
CD45+CD41lo
Kit+CD144-
30.36%
Kit+CD144+
1.46%
Kit-CD144-
68.10%
Kit-CD144+
0.09%
C-Kit
CD144 (VE-Cadherin)
B
%c-Kit+CD144+ cells
hGx CD45+
C
colonies per 3hGx (216h)
G
M
GM
E
GEM
CFC
D
GM
GEM
G
M
E
2mm
E
3 gastruloids / PBS
5 test mice;
1 control mouse
4 weeks
engraftment (gDNA / BFP)
BM, spleen
histology adrenal gland
F
Adrenal #15 hGx (4 weeks)
Adrenal #C PBS (4 weeks)
G
MWM NTC
BM#11
BM#12
BM#13
mES MWM
BM#14
BM#15
BM#C (PBS)
bone marrow (BM)
gDNA from 100K cells/lane
detection limit 10pg ~2 cells/100K
Spl#11
Spl#12
Spl#13
Spl#14
Spl#15
Spl#C (PBS)
spleen (Spl)

Figure 3

357

358 **HGx resolve the early hemogenic ontogeny targeted by infant AML gene *MNX1***

359 The ability of hGx to reconstitute YS and AGM-like HE and hemopoietic progenitors
360 with temporal coherence lends itself as an attractive model to understand
361 developmental leukemia. In particular, forms of AML unique to the first 2 years of life
362 – infant (inf)AML – have been shown to transform fetal liver (FL) but not adult
363 haematopoietic cells (Waraky *et al.,* 2024; Mercher *et al.,* 2009; Chen, W. *et al.,* 2011;
364 Ragusa *et al.,* 2023), compatible with targeting of a transient embryonic cell. In the
365 case of t(7;12) AML (herein MNX1-*r*) (Ragusa *et al.,* 2023), which results in a
366 transformation-driving ectopic expression of *MNX1* (Waraky *et al.,* 2024), analysis of
367 patient transcriptional signatures (Ragusa *et al.,* 2022) for over-representation of cell
368 atlas-defined affiliations, indicates enrichment in cardiomyogenic, endothelial,
369 HSC/progenitor and mast cells (**Fig. 4A**). These cell enrichments are compatible with
370 an early hemogenic cell derivative, capturable in the hGx model. Other cell type-
371 enrichments reflect MNX1 functions in pancreatic and neuronal development (Ragusa
372 *et al.,* 2022) (**Fig. S5A**).
373 We overexpressed *MNX1* (MNX1-OE) in *Flk1-GFP* mouse ES cells by lentiviral
374 transduction (**Fig. 4B**) and analyzed progression of hGx generation (**Fig. 4C**). We
375 used human *MNX1* cDNA to distinguish from the endogenous gene, but the degree of
376 homology is nevertheless high (84%), supporting functional equivalence. We
377 confirmed that *MNX1* overexpression in hGx is maintained at endpoint (**Fig S5B**).
378 MNX1-OE hGx activated polarized *Flk1-GFP* expression and elongated with similar
379 kinetics to empty-vector (EV) control (**Fig. 4C**), but consistently produced larger hGx
380 (**Fig. 4D**) denoting increased cellularity (**Fig. S5C**). From 192h onwards, MNX1-OE
381 hGx had a higher frequency of spontaneously contractile structures (**Fig. S5D**),
382 compatible with the cardiogenic cell association of patients' transcriptomes (**Fig. 4A**).
383 We interrogated the time-dependent hemogenic cell composition of hGx by flow
384 cytometry, quantifying C-Kit, CD41 and CD45 (**Fig. 4E-G** and **S5E-F**). We observed a
385 significant expansion of the C-Kit$^+$ compartment specifically at 144h (**Fig. 4E** and **S5E**).
386 All cell markers were relatively unchanged at later timepoints, although absolute cell
387 numbers were higher in MNX1-OE hGx (**Fig. S5C**).

388 To better understand the consequences of MNX1-OE on hemogenic development, we
389 performed RNA-sequencing (RNA-seq) of hGx at 144h and 216h, matching YS and
390 AGM-like hemopoietic waves, respectively (**Fig. 4H**). We confirmed expression of the
391 human *MNX1* transgene in the sequenced reads (**Fig. S6A**); in contrast, we could not
392 detect endogenous mouse *Mnx1* in any of the samples (**Fig. S6A**), confirming that it
393 does not normally play a role in hemogenic development. DEGs between MNX1-OE
394 and EV (**Supplemental File S3**) configured distinct MNX1-driven programs at 144h
395 and 216h, with >80% genes timepoint-specific (**Fig. 4H**). Both programs include up-
396 regulation of hematopoietic-associated genes, capturing EHT and progenitor
397 regulators (e.g. *Runx1*, and *Epor* and *Zfpm1*, respectively) at the earlier timepoint, and
398 more differentiated effectors, e.g. *Hba-a1*, *Hbb-bh1* and *Il3ra* at 216h, compatible with
399 selective targeting and/or expansion of newly specified progenitors at the YS-like
400 stage (**Fig. 4H**). An overview of transcriptional regulatory programs at both timepoints
401 denotes a clearer hemogenic enrichment at 144h, with over-representation of GATA2,
402 GATA1 and RUNX1 binding targets (**Fig. 4I**), putatively capturing an expansion of the
403 HE-to-EMP transition, also represented in enriched GO categories related to VEGF
404 signalling and epithelial-to-mesenchymal transition (EMT) (**Fig. 4J**). Compatible with
405 the increase in contractile foci in MNX1-OE late hGx, there was an enrichment in
406 cardiogenic gene ontologies (GO) at the 216h timepoint (**Fig. 4J**). No hemogenic
407 transcription factors (TF)-ChIP or GO categories were enriched in MNX1-OE hGx at
408 216h (**Fig. 4J-I**).
409 We attempted to deconvolute the cellular heterogeneity underlying bulk RNA-seq
410 DEGs by interrogating the cluster signatures obtained from the single-cell analysis of
411 hGx development (**Fig. S6B-C**; refer to **Fig. 2A**). In support of enhanced YS-stage
412 hemato-endothelial specification, MNX1-OE transcriptomes showed enrichment of
413 hGx HE cluster 5 and 0 signatures, and of the EMP-like cluster 4 (**Fig. S6B**). MNX1-
414 OE DEGs at both timepoints also enriched for the relatively less differentiated cluster
415 10, which predominantly captures unfractionated hGx cells at 120h (**Fig. S6B-C**; refer
416 to **Fig. 2A**). Significantly, clusters 5 and 0 HE signatures were enriched in *MNX1*-r
417 infAML patient signatures (Ragusa et al., 2022) (**Fig. S5B-C**), positioning MNX1-OE
418 at early stages of hemogenic specification. MNX1-OE DEGs at 216h show enrichment
419 in late-stage clusters 7 and 8 (**Fig. S6B-C**), denoting enhancement of hGx

420 development, including MLP/MPP in cluster 8, but diverging from *MNX1*-r infAML
421 signatures. Projection of infAML subtype-specific gene signatures onto hGx
422 development separates the lineage affiliations of *MNX1*-r from, for example, the most
423 frequent infAML driver – *KMT2A (MLL)* rearrangements (**Fig. S6B-C**), which carries
424 myelo-monocytic and/or mixed-lineage affiliation, highlighting the utility of the hGx
425 system for interrogating infAML biology.

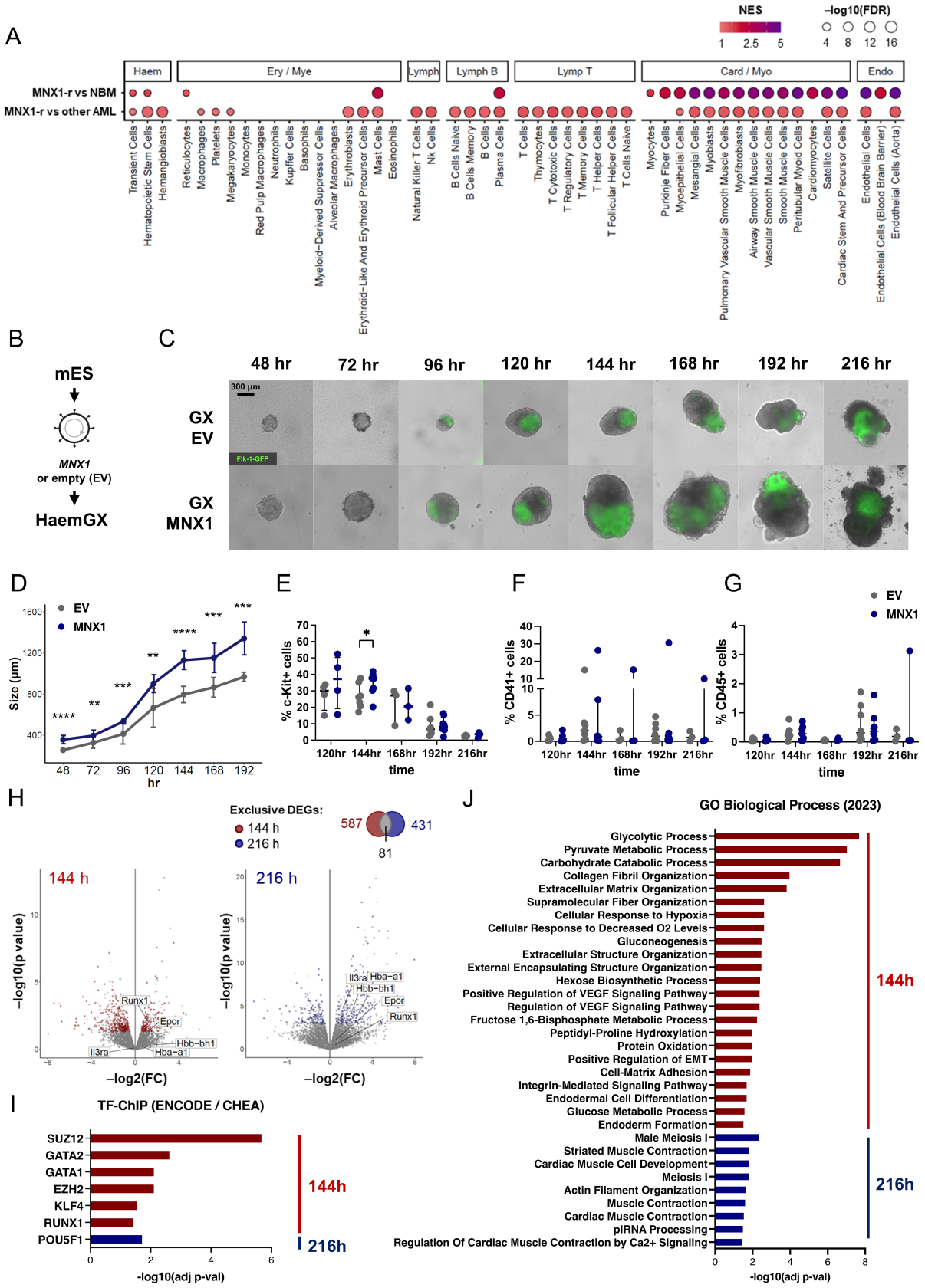
A
NES
1 2.5 5
−log10(FDR)
4 8 12 16
Haem
Ery / Mye
Lymph
Lymph B
Lymp T
Card / Myo
Endo
MNX1-r vs NBM
MNX1-r vs other AML
Transient Cells
Hematopoietic Stem Cells
Hemangioblasts
Reticulocytes
Macrophages
Platelets
Megakaryocytes
Monocytes
Red Pulp Macrophages
Neutrophils
Kupffer Cells
Basophils
Myeloid−Derived Suppressor Cells
Alveolar Macrophages
Erythroblasts
Erythroid−Like And Erythroid Precursor Cells
Mast Cells
Eosinophils
Natural Killer T Cells
Nk Cells
B Cells Naive
B Cells Memory
B Cells
Plasma Cells
T Cells
Thymocytes
T Cytotoxic Cells
T Regulatory Cells
T Memory Cells
T Helper Cells
T Follicular Helper Cells
T Cells Naive
Myocytes
Purkinje Fiber Cells
Myoepithelial Cells
Mesangial Cells
Myoblasts
Pulmonary Vascular Smooth Muscle Cells
Myofibroblasts
Airway Smooth Muscle Cells
Vascular Smooth Muscle Cells
Smooth Muscle Cells
Peritubular Myoid Cells
Cardiomyocytes
Satellite Cells
Cardiac Stem And Precursor Cells
Endothelial Cells
Endothelial Cells (Blood Brain Barrier)
Endothelial Cells (Aorta)
B
mES
MNX1
or empty (EV)
HaemGX
C
48 hr
72 hr
96 hr
120 hr
144 hr
168 hr
192 hr
216 hr
GX EV
GX MNX1
300 µm
Flk-1-GFP
D
EV
MNX1
Size (µm)
hr
E
% c-Kit+ cells
time
F
% CD41+ cells
time
G
% CD45+ cells
time
H
Exclusive DEGs:
144 h
216 h
587
431
81
144 h
216 h
−log10(p value)
−log2(FC)
Runx1
Epor
Hbb−bh1
Hba−a1
Il3ra
I
TF-ChIP (ENCODE / CHEA)
SUZ12
GATA2
GATA1
EZH2
KLF4
RUNX1
POU5F1
144h
216h
-log10(adj p-val)
J
GO Biological Process (2023)
Glycolytic Process
Pyruvate Metabolic Process
Carbohydrate Catabolic Process
Collagen Fibril Organization
Extracellular Matrix Organization
Supramolecular Fiber Organization
Cellular Response to Hypoxia
Cellular Response to Decreased O2 Levels
Gluconeogenesis
Extracellular Structure Organization
External Encapsulating Structure Organization
Hexose Biosynthetic Process
Positive Regulation of VEGF Signaling Pathway
Regulation of VEGF Signaling Pathway
Fructose 1,6-Bisphosphate Metabolic Process
Peptidyl-Proline Hydroxylation
Protein Oxidation
Positive Regulation of EMT
Cell-Matrix Adhesion
Integrin-Mediated Signaling Pathway
Endodermal Cell Differentiation
Glucose Metabolic Process
Endoderm Formation
Male Meiosis I
Striated Muscle Contraction
Cardiac Muscle Cell Development
Meiosis I
Actin Filament Organization
Muscle Contraction
Cardiac Muscle Contraction
piRNA Processing
Regulation Of Cardiac Muscle Contraction by Ca2+ Signaling
144h
216h
-log10(adj p-val)

Figure 4

429 **Serial replating of MNX1-OE hGx cells identifies a candidate C-Kit$^+$ leukemia-**
430 **propagating cell more susceptible to transformation in the YS period**

431 We further interrogated the *MNX1*-r-like leukemia-initiating potential of MNX1-OE hGx
432 through serial replating of CFC assays, a classical *in vitro* assay of leukemia
433 transformation. We compared 144h and 216h hGx to match functional and
434 transcriptome data, and plated unfractionated hGx, given the transient nature of C-Kit$^+$
435 enrichment. Both 144h (**Fig. 5A-B**) and 216h (**Fig. 5C-D**) MNX1-OE hGx replated for
436 at least 5 plates, with significant differences to control (EV) (**Fig. 5B,D**). EV-hGx
437 replating could not be sustained beyond plate 3 in most experiments started from 216h
438 (**Fig. 5C**); 144h EV-hGx exhibited some low-level replating until plate 5 (**Fig. 5A**), but
439 the colonies were formed by few scattered cells clearly distinct from the well-
440 demarcated colonies obtained from MNX1-OE hGx. We analyzed early-stage (plate
441 1) and late-stage (plate 5) colonies by flow cytometry to identify the nature of the
442 replating cells. Cells selected through replating were C-Kit$^+$ (**Fig. 5E**), matching the
443 cell phenotype transiently enriched at 144h. Early replating of MNX1-OE hGx also
444 enriched for erythroid-affiliated Ter119$^+$ cells particularly from the 144h time-point (**Fig.**
445 **5F**), reflecting the differential erythroid potential of EMP and MLP/MPP at 144h and
446 216h, respectively. This enrichment is not sustained upon replating, compatible with
447 the undifferentiated nature of *MNX1*-r infAML, which is also reflected by the low levels
448 of CD11b, a marker affiliated with the myelo-monocytic lineages (**Fig. 5E**). Notably,
449 the cells obtained are not CD45$^+$ (**Fig. 5F**), suggesting perpetuation of an early
450 hemogenic state.

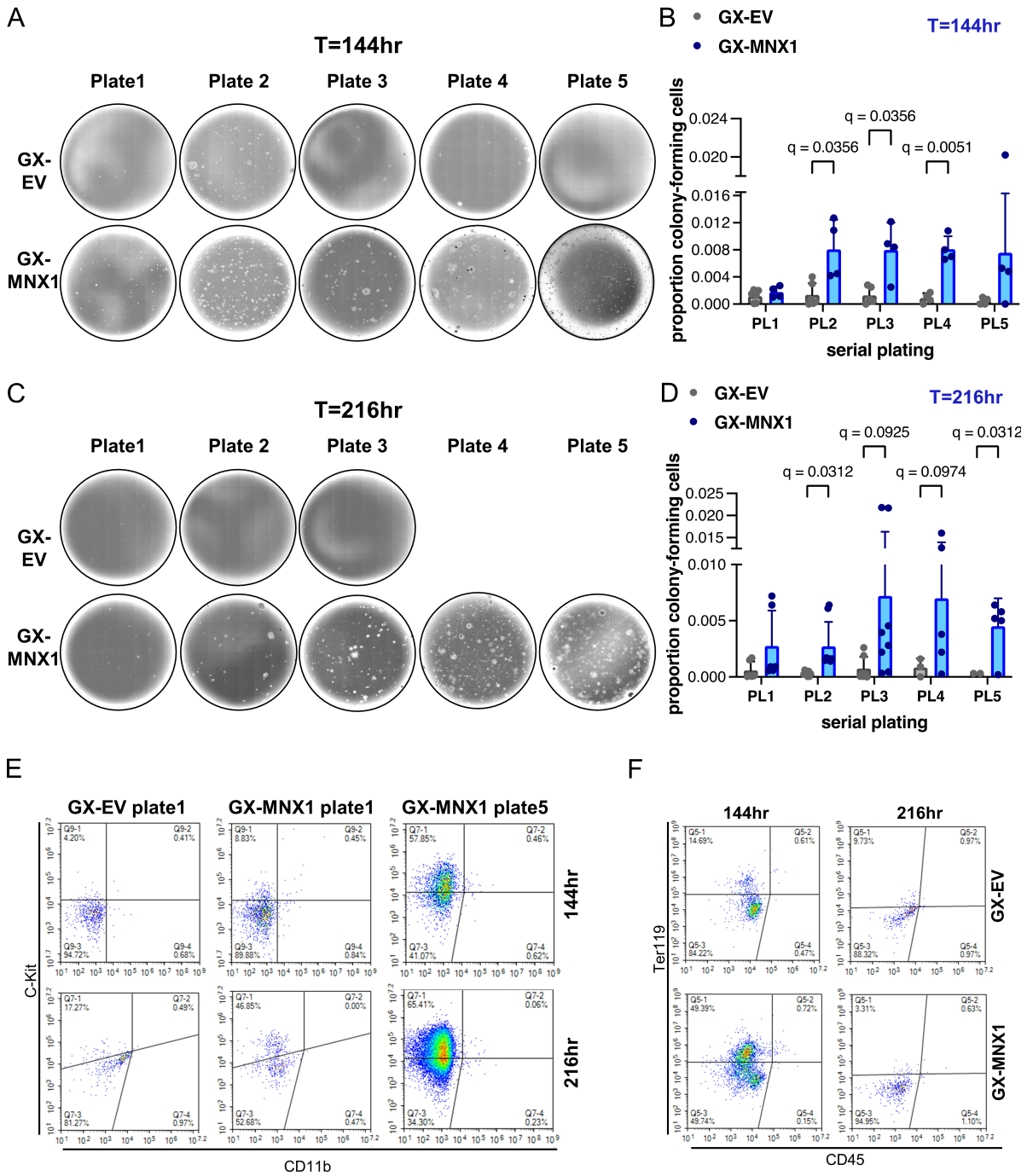
A
T=144hr
Plate1
Plate 2
Plate 3
Plate 4
Plate 5
GX-EV
GX-MNX1
B
GX-EV
GX-MNX1
T=144hr
proportion colony-forming cells
q = 0.0356
q = 0.0356
q = 0.0051
PL1
PL2
PL3
PL4
PL5
serial plating
C
T=216hr
Plate1
Plate 2
Plate 3
Plate 4
Plate 5
GX-EV
GX-MNX1
D
GX-EV
GX-MNX1
T=216hr
proportion colony-forming cells
q = 0.0925
q = 0.0312
q = 0.0312
q = 0.0974
PL1
PL2
PL3
PL4
PL5
serial plating
E
GX-EV plate1
GX-MNX1 plate1
GX-MNX1 plate5
144hr
216hr
C-Kit
CD11b
F
144hr
216hr
GX-EV
GX-MNX1
Ter119
CD45

Figure 5

453
454
455 We performed RNA-seq analysis of MNX1-OE hGx CFC cells obtained through re-
456 plating and integrated the data with the 144h and 216h hGx time-points to identify
457 genes enriched through the process of transformation. We observed relative
458 enrichment in Hs*MNX1* reads in MNX1-OE CFC samples, supporting selective
459 expansion of transduced cells (**Fig. 6A**). Hierarchical clustering of all DEGs in pairwise
460 comparisons (**Fig. 6B**), identified a subgroup of genes up-regulated upon CFC
461 replating which was also up-regulated at MNX1-OE hGx at 144h (K14), thus matching
462 the putative time-dependent enrichment of the MNX1-OE propagating population. To
463 understand the relevance of these genes in *MNX1*-r leukemia, we performed Gene
464 Set Enrichment Analysis (GSEA) of *MNX1*-r patients in comparison to other AML in
465 the same infant (0-2) age group. K14-genes were enriched in MNX1-*r* all comparisons
466 (**Fig. 6C**), with considerable overlap between leading-edge (i.e. significantly-enriched)
467 genes (**Fig. S6D**). Interrogation of cell-type-specific signatures within the K14 MNX1-
468 *r* leading edge genes (**Fig. 6D**) against the single-cell PanglaoDB atlas captured
469 similarities to hemogenic and non-hemogenic cell types, including interneuron, and
470 other neural, signatures, which reflect MNX1 biological functions (Harrison *et al.,* 1999;
471 Thaler *et al.,* 1999; Ragusa *et al.,* 2022). The most highly significant hemogenic
472 signatures capture early developmental components, such as 'endothelial cells', which
473 were specifically up-regulated within MNX1-OE hGx (**Fig. S6B**). 'Erythroid cell
474 precursors' also reflect MNX1-OE hGx effects (**Fig. 4E-J**). We were intrigued by the
475 'mast cells' affiliation of *MNX1*-r AML signatures within the hGx K14 cluster, which was
476 more specifically aligned with MNX1-OE hGx isolated at 144h (**Fig. S6E**). Given the
477 characteristic morphology of these cells, we analysed Giemsa-Wright-stained
478 cytospins of CFC replating assays (**Fig. 6E** and **S6F**). Consecutive plates of 144h
479 MNX1-OE-initiated *in vitro* transformation showed a shift from blast-like at plate 3 to
480 precursor cell morphology with some differentiated mast cells at plate 5 (**Fig. S6F**).
481 However, the extent of differentiation was greater, and apparent earlier, in 216h-
482 MNX1-OE initiated colonies (**Fig. 6E**). This suggests that the MNX1-OE-responsive
483 hemogenic C-Kit^{+} progenitor may experience maturation between 144h and 216h, and
484 is more likely targeted for transformation at the YS EMP stage. The data support the

utility of the hGx in dissecting ontogeny of embryonic origin in hematological cancers, allowing us to specifically narrow down the initiating cellular targets of MNX1-*r* AML.

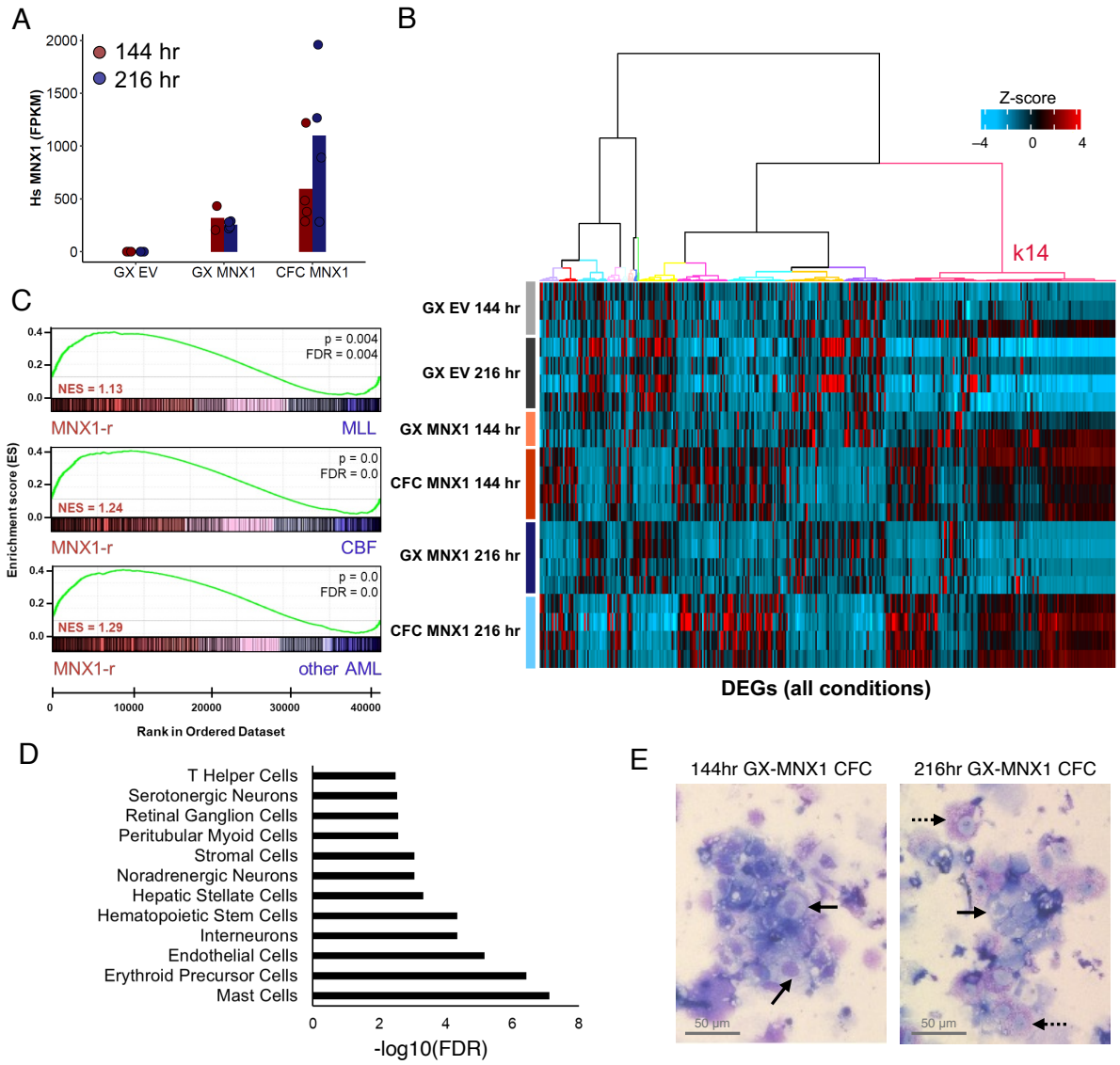

Figure 6

492 **DISCUSSION**

493 In this study, we developed and characterized a hGx model from mouse ES cells,
494 which successfully recapitulates YS-like and AGM-like waves of definitive blood
495 progenitor specification, and can engraft adult hematopoietic tissues upon *in vivo*
496 maturation. We took advantage of the hGx's nearly full recapitulation of embryonic
497 hemogenic specification, including critical early stages of HE production and EHT to
498 EMP and MLP/MPPs, to interrogate the biology of *MNX1*-r AML, a poorly
499 characterized form of leukemia exclusive to the infant age group (Ragusa *et al.,* 2023).
500 We showed that MNX1-OE transiently expanded a C-Kit$^+$ cell with hemato-endothelial
501 characteristics at the YS-like stage, and could selectively sustain the phenotype
502 through serial replating of CFC assays. Replating CFCs could be initiated beyond the
503 YS-like stage, albeit at the relative expense of an early differentiation state and of the
504 similarity to patient transcriptome, putatively defining a window of susceptibility to
505 MNX1-driven transformation. The hGx MNX1-OE propagating cell bears
506 transcriptional resemblance to patient *MNX1*-r AML, overall positioning the hGx as an
507 attractive tractable model in which to study developmental leukemias.

508 *MNX1*-r has been difficult to characterize due to a combination of rarity, diagnostic
509 difficulty (the underlying t(7;12) can be missed if not specifically looked for) (Tosi *et*
510 *al.,* 2015), and the lack of a fusion protein that can be ectopically expressed and
511 tracked through development, as is the case for other translocations unique to infAML
512 (Alexander and Mullighan, 2021). However, a recent paper using MNX1-OE has
513 shown that MNX1 can initiate a transplantable leukemia in immunodeficient mice by
514 targeting FL, but not BM, cells (Waraky *et al.,* 2024). This suggests that ectopic
515 expression of *MNX1* transforms an embryonic cell, rather than depends on an
516 embryonic niche. It also suggests that MNX1-OE can cause leukemia as a single
517 genetic hit, although it cannot be excluded that other mutations are required for, and
518 may indeed have contributed to, full transformation *in vivo*, as observed in a recent
519 report confirming the *in utero* origin of the leukemia (Bousquets-Muñoz *et al.,* 2024).
520 However, the mouse model does not address the nature of the embryonic cell targeted
521 by MNX1-OE. Like the MNX1-OE hGx cells selected through serial replating, MNX1-
522 OE mouse leukemia are also C-Kit$^+$ (Waraky *et al.,* 2024). Interestingly, clusters of
523 genes up-regulated in the mouse MNX1-OE leukemia (**Fig. S7A** – K1, 2, 5, 6, and 9)

524 do not robustly separate *MNX1*-r from other infAML by GSEA analysis (**Fig. S7B**), and
525 have diverged more from the patient data in their cell-type signature analysis than the
526 original MNX1-OE cells in the FL (**Fig. S7C**), putatively selecting a more generic
527 program of leukemia transformation. We propose that the divergence from *MNX1*-r
528 infAML signatures increases from MNX1-OE hGx 144h (YS-like), to 216h (AGM-like)
529 to FL, positioning the origin of the AML at the YS stage, likely at the HE-to-EMP
530 transition. This cell may persist in the hGx culture system to the 216h AGM-like time-
531 point, and migrate to the FL in the embryo, eventually extinguishing through
532 differentiation, as suggested by the MNX1-OE hGx CFC assays. Persistence of the
533 original MNX1-OE target cell will determine the sensitivity to MNX1-OE, and can be
534 addressed in the hGx system through lineage tracing. Whether additional mutations,
535 namely a reported association between *MNX1*-r and tri(19) (Espersen *et al.,* 2018;
536 Ragusa *et al.,* 2023), contribute to prolonging the undifferentiated state of the
537 candidate YS target cell with increased probability of transformation, or otherwise
538 expand the targeted cell in a more classical second-hit fashion, can potentially also be
539 screened in the hGx system, in combination with transplantation experiments. Finally,
540 although our study has positioned the MNX1-OE target cell within the YS-EMP stage,
541 the exact nature of the cell requires better definition. The mast cell affiliation,
542 transcriptional and by CFC morphology, is more likely to reflect the period of
543 susceptibility to MNX1 and the multilineage potential of the target cell, than the lineage
544 commitment of the latter. Indeed, blast cells in the leukemia mouse model are
545 morphologically undifferentiated (Waraky *et al.,* 2024) and a mast cell morphology has
546 not been described in patients (Ragusa *et al.,* 2023; Espersen *et al.,* 2018). Although
547 adult HSC have mast cell potential *in vitro*, this is not employed *in vivo* under
548 physiological conditions (Chia *et al.,* 2023). Tissue resident mast cells are formed in
549 the embryo at 2 periods: the YS EMP stage, which we propose to have targeted with
550 MNX1-OE in the hGx model; and a transient AGM FL-HSC population which does not
551 persist in post-natal life (Chia *et al.,* 2023; Yoshimoto *et al.,* 2022), and that may be
552 incipiently specified at the end of the current hGx protocol. Prospective isolation of
553 these cell types from embryos and over-expression of *MNX1* followed by
554 transformation *in vitro* and *in vivo*, will precisely define the embryonic time-window of
555 susceptibility to *MNX1*-r and the nature the leukemia-initiating cell. Harnessing the

556 high-throughput nature of the hGx culture system will help identifying therapeutic
557 vulnerabilities that can improve the prognosis of this often-fatal disease.
558 The hGx model furthermore constitutes an advanced platform to probe physiological
559 development of the hematopoietic system. Currently, *in vitro* platforms do not
560 recapitulate *in vivo* blood formation sufficiently faithfully (Dijkhuis *et al.,* 2023; Fidanza
561 and Forrester, 2021). Establishment of a tractable *in vitro* model of hematopoietic
562 development which faithfully recapitulates time-dependent emergence of
563 hematopoietic progenitors in their physiological niche would propel translational
564 applications by enabling high-throughput studies of HSC emergence. Most ESC-
565 based models of blood specification do not achieve concomitant recapitulation of the
566 surrounding microenvironment. Protocols typically remove ESC-derived hemogenic
567 cells from an early embryonic-like niche to direct differentiation through genetic
568 manipulation (Sugimura *et al.,* 2017) and/or empirically-determined growth factor
569 addition (Lis *et al.,* 2017; Vo and Daley, 2015; Fowler *et al.,* 2024; Nafria *et al.,* 2020;
570 Sroczynska *et al.,* 2009; Piau *et al.,* 2023), with minimal probing of intermediate steps.
571 When hematopoietic waves have been specifically inspected, emergence of YS and
572 AGM-like stages were conflated in time (Pearson *et al.,* 2015; Murry and Keller, 2008),
573 suggesting a bias towards intrinsic hematopoietic cell potential, at the expense of cell
574 fate modulation by time-coherent extrinsic physiological cues, affecting the nature of
575 cellular output. The present hGx system, in contrast, successfully deconvolutes YS-
576 like and AGM-like stages of blood production, including capturing differences in the
577 HE generation, suggesting a level of organisation amenable to further development
578 into an integrated hemogenic niche. Although the system is shy of achieving HSC
579 maturation *in situ*, it can mature in a supportive microenvironment, in this case
580 provided by the adrenal gland (Schyrr *et al.,* 2023), to produce engrafting multilineage
581 progenitors, and potentially HSCs. This suggests that access to the relevant
582 hematopoietic molecular programs is preserved. Incipient specification of non-
583 hematopoietic tissues shown to support HSC emergence, including PDGFRA^{+}
584 mesenchymal (Chandrakanthan *et al.,* 2022) and neuro-mesenchymal (Miladinovic *et*
585 *al.,* 2024) cells, and neural-crest derivatives (Kapeni *et al.,* 2022; Fitch *et al.,* 2012),
586 occurs in sync with development of AGM-like hematopoiesis, to indicate the presence

587 of cues that with further maturation may facilitate HSPC specification at the
588 appropriate developmental time.
589 We have observed that the nature of the symmetry-breaking pulse affects the
590 specification of hemato-endothelial programs, with a more robust *Flk1*/*Kdr* activation
591 in the presence of Activin A, potentially balancing mesodermal derivatives and the
592 anterior-posterior axis. The hematopoietic output of cardiogenic gastruloids (Rossi *et*
593 *al.,* 2022) does not seem to require Activin A, but their blood formation is of lower
594 efficiency, at least at the comparable timepoints of 144 and 168h, and apparently
595 biased towards erythroid formation, reflecting its YS, and potentially primitive, nature.
596 In the absence of additional cytokines after the CHI99021 pulse, standard gastruloid
597 protocols (Beccari *et al.,* 2018) capture a similar transcriptional erythrogenic wave.
598 Although Rossi *et al.* (2022) reported the formation of spleen colonies, these are
599 difficult to ascertain from the data published, and the engraftment capacity of
600 cardiogenic gastruloid-derived blood remains uncertain. Our hGx system has more
601 robust CD41^{+} cell production at 144-168h, which may reflect distinct tissue induction
602 by the CHI99021+Activin A pulse. Interestingly, through extending the hGx culture,
603 and in addition to AGM-like hematopoiesis, we also captured the potential to form renal
604 tubular structures upon *in vivo* maturation, suggesting induction of lateral plate and
605 intermediate mesoderm from which the AGM region derives, and putative
606 reconstitution of the intra-embryonic hematopoietic niche. An additional challenge is
607 tissue organisation: non-hematopoietic gastruloid models benefitted from mechanical,
608 or chemo-mechanical, support of artificial ECM embedding (Veenvliet *et al.,* 2020; van
609 den Brink, Susanne C *et al.,* 2020), to match temporal coherence and tissue
610 polarisation, with detailed tissue architecture illustrated by somite formation. Prior to
611 HSC emergence in the AGM, critical signals such as SCF and Shh, which we have
612 delivered to the cultures, are secreted by different cells on opposite dorsal-ventral
613 locations (Souilhol *et al.,* 2016), indicating the potential relevance of tissue
614 organisation to improve and potentially extend hematopoietic formation. Moreover,
615 HSPC formation in the AGM is facilitated by circulatory shear stress (North *et al.,* 2009;
616 Diaz *et al.,* 2015; Lundin *et al.,* 2020), suggesting that introduction of flow may improve
617 hemato-endothelial organization and boost blood production.

618 Over the last decade, the effort to achieve robust hematopoietic production from PSC,
619 mouse and human, has been centred on the generation of HSC, a critical translational
620 need (Chang *et al.,* 2023; Dijkhuis *et al.,* 2023; Fidanza and Forrester, 2021). There
621 are good indications that this is achievable, both with (Vo and Daley, 2015; Dang *et*
622 *al.,* 2002) and without (Piau *et al.,* 2023) genetic manipulation. However, success
623 remains episodic, mainly because mechanistic understanding of processes
624 underpinning blood generation during development is limiting, exacerbated by limiting
625 access to human embryonic and fetal material. On-going optimization of gastruloid
626 models and the success at integration of organoid cultures in microfluidic systems
627 (Saorin, Caligiuri and Rizzolio, 2023; Quintard *et al.,* 2024) will allow us to further
628 optimize hGx culture. In conclusion, the hGx represents a faithful *in vitro* reference to
629 replace and complement embryo studies, holding promise to understand the biology
630 of healthy and malignant developmental haematopoiesis and advance blood
631 production for clinical use.
632
633
634 **EXPERIMENTAL PROCEDURES**
635 **MATERIALS**

	SOURCE	IDENTIFIER
Antibodies		
Anti-mouse c-Kit (CD117) clone 2B8 - APC-Cy7	Biolegend	Cat. #105825
Anti-mouse CD41 clone MWReg30 - PE-Dazzle594	Biolegend	Cat. #133935
Anti-mouse CD45 clone 30F11 - APC	Biolegend	Cat. #103111
Anti-mouse CD144 (VE-cadherin) clone BV13 - APC	Biolegend	Cat. #138011
Anti-mouse CD43 clone S11 - PE	Biolegend	Cat. #143205
Anti-mouse Ter119 – PC7	Biolegend	Cat. #116223
CD11b anti-mouse/human clone M1/70 - biotin	Biolegend	Cat. #101203
Mouse Anti-Mouse CD45.2 clone 104 - PE	BD Biosciences	Cat. #560695
Rat Anti-Mouse CD31 clone MEC13.3 - biotin	BD Biosciences	Cat. #553371
Goat Anti-Mouse c-Kit/CD117	R&D Systems	Cat. #AF1356SP
Bacterial and virus strains		
E. coli: NEB 5-alpha Competent *E. coli*	New England Biolabs (NEB)	Cat. #C2987H
Chemicals, peptides, and recombinant proteins		
Murine LIF	Peprotech	Cat. #250-02
StemMACS PD0325901	Miltenyi Biotec	Cat. #130-106-5411
Chiron (CHIR99021)	Biogems	Cat. #2520691
Activin A PLUS	QKine	Cat. #Qk005
Murine VEGF 165	Peprotech	Cat. #450-32
Murine FGF-basic	Peprotech	Cat. #450-33

Murine Sonic Hedgehog (Shh)	Peprotech	Cat. #315-22
Murine TPO	Peprotech	Cat. #315-14
Murine Flt3-Ligand	Peprotech	Cat. #250-31L
Murine SCF	Peprotech	Cat. #250-03
N2B27 medium (NDiff 227)	Takara Bio	Cat. #Y40002
DMEM/F-12	Gibco	Cat. #11320033
Neurobasal medium	Gibco	Cat. #21103049
N-2 Supplement	Gibco	Cat. #17502048
B-27 Supplement	Gibco	Cat. #17504044
Critical commercial assays		
Mouse Methylcellulose Complete Media	R&D Systems	Cat. #HSC007
Deposited data		
Single cell RNA-seq of hemogenic gastruloids at timepoints	This paper	ArrayExpress: E-MTAB-12148
Bulk RNA-seq of hemogenic gastruloids with MNX1 overexpression at 144h and 216 h	This paper	ArrayExpress: E-MTAB-12173
Experimental models: Cell lines		
Mus musculus: Flk-1-GFP mouse embryonic stem cell line	Alfonso Martinez Arias Lab	Jakobsson et al., 2010
Mus musculus: Sox17-GFP mouse embryonic stem cell line	Alfonso Martinez Arias Lab	Niakan et al., 2010
Mus musculus: T/Bra::GFP mouse embryonic stem cell line	Alfonso Martinez Arias Lab	Fehling et al., 2003
Mus musculus: E14Tg2a mouse embryonic stem cell line	MMRRC, University of California Davis, US	Hooper et al., 1987
Mus musculus: Flk-1-GFP::Rosa26-BFP mouse embryonic stem cell line	This paper	n/a
Homo sapiens: HEK293T	ATCC	CRL-3216
Oligonucleotides		
MNX1 (forward 5-GTTCAAGCTCAACAAGTACC-3; reverse 5- GGTTCTGGAACCAAATCTTC-3)	Gulino et al., 2021	n/a
Ppia (forward 5- TTACCCATCAAACCATTCCTTCTG-3; reverse 5- AACCCAAAGAACTTCAGTGAGAGC-3)	Moris et al., 2018	n/a
BFP (forward 5-GCACCGTGGACAACCATCACTT-3; reverse 5-CAGTTTGCTAGGGAGGTCGC-3)	This paper	n/a
Recombinant DNA		
pWPT-LSSmOrange-PQR	Ghevaert Lab, WT-MRC Cambridge Stem Cell Institute, UK	Dalby et al., 2018
pWPT-LSSmOrange-PQR-MNX1	Cloned by Biomatik Corporation (Kitchener, Canada)	n/a
pcdna3-rosa26-rosa26-bfp	This paper	n/a
Software and algorithms		
GSEA software v4.2.3	Subramanian et al., 2005	https://www.gsea-msigdb.org/gsea/index.jsp
ImageJ	Schneider et al., 2012	https://imagej.nih.gov/ij/
Bowtie2	Langmead and Salzberg, 2012	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
Samtools	Li et al., 2009	http://samtools.sourceforge.net/

Tophat2	Kim et al., 2013	https://ccb.jhu.edu/software/tophat/index.shtml
HTseq	Anders et al., 2015	https://htseq.readthedocs.io/en/master/
DEseq2	Love et al., 2014	https://bioconductor.org/packages/release/bioc/html/DESeq2.html
ExpressAnalyst	n/a	https://www.expressanalyst.ca/
PANTHER	Thomas et al., 2022	http://www.pantherdb.org/
EnrichR	Chen et al., 2013; Kuleshov et al., 2016; Xie et al., 2021	https://maayanlab.cloud/Enrichr/
Panglao BP	Franzen et al., 2019	https://panglaodb.se
TARGET (Therapeutically Applicable Research to Generate Effective Treatments) Database	https://ocg.cancer.gov/programs/target https://xenabrowser.net/	GDC TARGET-AML
Single-cell RNA sequencing analysis scripts	DockerHub: dsblab/single_cell_analysis:0.5	https://github.com/dsb-lab/blood_gastruloids
Other		
CELLSTAR cell-repellent cell culture plate, 96 well, U-bottom	Greiner (Bio-One)	Cat. #650970

636
637

638 **METHODS**

639 **Cell culture**

640 Flk-1-GFP (Jakobsson *et al.,* 2010), Flk-1-GFP::Rosa26-BFP, Sox17-GFP (Niakan *et*
641 *al.,* 2010), T/Bra::GFP (Fehling *et al.,* 2003), and E14Tg2A (Hooper *et al.,* 1987)
642 mouse embryonic stem cells (mES) lines were cultured in ES+LIF medium in
643 gelatinized (0.1% gelatin) with daily medium change, as previously described (Turner
644 *et al.,* 2017). The ES+LIF medium contained 500 ml Glasgow MEM BHK-21 (Gibco),
645 50 ml of fetal bovine serum (FBS, Embryonic Stem Cells tested, biosera, Nuaillé,
646 France), 5 ml GlutaMAX supplement (Gibco), 5 ml MEM Non-Essential Amino Acids
647 (Gibco), 5 ml Sodium pyruvate solution (100 mM, Gibco), and 1 ml 2-Mercaptoethanol
648 (50mM, Gibco). Murine LIF (Peprotech) was added at 1000 U/ml. HEK293T cells were
649 grown in Dulbecco's modified Eagle medium (Gibco) supplemented with 10% FBS

650 (Gibco) and 1% penicillin/streptomycin (Gibco). All cultures were kept at 37°C and 5%
651 CO_2.

652

653 **Hemogenic gastruloids assembly, culture, and dissociation**

654 mES were maintained in ES+LIF medium and transferred to 2i+LIF (containing Chiron
655 and MEK inhibitor PD03) for 24 hours prior to the assembly into gastruloids. 250
656 untransduced mES cells (or 400 for GX-MNX1 and GX-EV) were seeded in each well
657 of a U-bottom, cell-repellent 96-well plate (Greiner Bio-One, Stonehouse, UK) in 40 µl
658 of N2B27 medium [Takara Bio or home-made as Cold Spring Harbour Protocols
659 'N2B27 Medium' (2017)]. The plate was centrifuged at 750 rpm for 2 minutes to
660 promote deposition and aggregation of the cells and was then incubated at 37°C, 5%
661 CO2 for 48 hours. After 48 hours, 150 µl of N2B27 medium supplemented with 100
662 ng/ml Activin A (QKine, Cambridge, UK) and 3 µM chiron (Peprotech) was added to
663 each well. At 72 hours, 150 µl of medium were removed, without disrupting the
664 gastruloids in the wells. 100 µl of N2B27 with 5 ng/ml of Vegf and Fgf2 each
665 (Peprotech) were added to each well. From 72 h to 144 h, each day 100 µl of medium
666 were removed and replaced with N2B27 + Vegf + Fgf2. At 144 h, the medium was
667 further supplemented with Shh at 20 ng/ml. From 168 h to 216 h, the medium was
668 N2B27 + 5 ng/ml Vegf, plus 20 ng/ml mTpo, 100 ng/ml mFlt3l, and 100 ng/ml mScf
669 (Peprotech). To dissociate cells from the gastruloid structures, medium was removed
670 and individual gastruloids were collected using a pipette and precipitated at the bottom
671 of a microcentrifuge tube. The remaining medium was aspirated and the bulk of
672 gastruloids was washed in PBS. 50 µl of TrypLE express was added to pelleted
673 gastruloids to be incubated at 37°C for 2 minutes to dissociate cells.

674

675 **Animals and implantation**

676 Female athymic nude-Foxn1^{nu} (*nu/nu*) mice (Envigo, Bresso, Italy) were housed under
677 pathogen-free conditions. In accordance with the "3Rs policy", experiments were
678 reviewed and approved by the Animal Welfare Body (OPBA) of IRCCS Ospedale
679 Policlinico San Martino and by the Italian Ministry of Health (n. 883/2020-PR).

680 Intact hGx (3 per mouse in 10 µL of PBS) were inoculated in the adrenal gland of five-
681 week-old mice. Briefly, mice were anesthetized with a mixture of xylazine (ROMPUN,
682 Bayer) and ketamine (LOBOTOR, Acme S.r.l.) and injected with hGx, after
683 laparotomy, in the capsule of the left adrenal gland, as previously described (Brignole
684 *et al.,* 2023; Pastorino *et al.,* 2003). All mice survived surgery. Control mice (CTR)
685 received vehicle only. Mice body weight and general physical conditions were
686 recorded daily. One month after hGx inoculation, mice were sacrificed and spleen and
687 bone marrow (BM) from femurs collected. BM cells and spleens, mechanically reduced
688 to cell suspension, were stored at -80°C until use. Adrenal glands were paraffin-
689 embedded for subsequent IHC studies.

690

691 **Methylcellulose colony forming assays (CFC)**

692 Disassembled gastruloids cells were plated on Mouse Methylcellulose Complete
693 Media (R&D Systems). Cells were first suspended in Iscove's Modified Dulbecco's
694 Medium (IMDM, Gibco) with 20% FBS (Gibco) before addition to the methylcellulose
695 medium. Cells were plated in duplicate 35 mm dishes with 2×10^5 cells/plate. Plates
696 were incubated at 37°C and 5% CO2 for 10 days, when colonies were scored. For
697 serial replating experiments, cells in methylcellulose were collected and washed in
698 phosphate buffer saline (PBS) to achieve single-cell suspensions and replated as
699 described above. Dissociated gastruloid cells from CFC were collected as cell
700 suspensions in PBS and centrifuged onto slides (Shandon Cytospin 2 Cytocentrifuge)
701 at 350rpm for 5-7min, with a PBS-only pre-spin at 1000rpm, 2min. Slides were stained
702 in Giemsa-Wright stain for 30 seconds followed by addition of phosphate buffer pH 6.5
703 for 5 min.

704

705 **Immunofluorescence staining**

706 Immunostaining of whole gastruloids was performed as described before (Baillie-
707 Johnson et al., 2015). Briefly, gastruloids were fixed in 4% paraformaldehyde (PFA)
708 dissolved in PBS for 4 hours at 4°C on orbital shaking and permeabilised in PBSFT
709 (10% FBS and 0.2% Triton X-100), followed by one hour blocking in PBSFT at 4°C on

710 orbital shaking. Antibody dilutions were made in PBSFT at 1:200 for primary and 1:500
711 for secondary antibodies. Antibody incubations were performed overnight at 4°C on
712 orbital shaking, and subjected to optical clearing overnight in ScaleS4 clearing
713 solution. Individual gastruloids were then mounted on glass coverslips by pipetting as
714 droplets in ScaleS4 and DAPI nuclear stain.

715

716 **Histology and H&E staining**

717 Samples were fixed with 10% formalin at room temperature for 48 hours followed by
718 an ethanol series of ethanol and paraffin; ethanol 70% (90 minutes), ethanol 80% (120
719 minutes), ethanol 95% (180 minutes), 3 times ethanol 100% of 120, 120 and 180
720 minutes, 50% ethanol/paraffin (3x120 minutes) and 3 times paraffin (60, 90 and 120
721 minutes) before embedding in histological cassettes. Blocks were sectioned with a
722 HistoCore BIOCUT Microtome (Leica) at 10 mm and slides were stained with
723 Haematoxylin/Eosin.

724

725 **Imaging**

726 Images of cultured gastruloids and CFC plates were captured using the Cytation 5 Cell
727 Imaging Multi-Mode Reader (Biotek) plate reader using bright field and FITC channels.
728 ImageJ (Schneider et al., 2012) was used for gastruloid size quantification. Confocal
729 microscopy was performed on LSM700 on a Zeiss Axiovert 200 M with Zeiss EC Plan-
730 Neofluar 10x/0.30 M27 and Zeiss LD Plan-Neofluar 20x/0.4 M27 objective lens. Tissue
731 sections were imaged with a THUNDER Imager 3D microscope system (Leica).

732

733 **Lentiviral vector packaging and transduction**

734 The lentiviral overexpression vector pWPT-LSSmOrange-PQR was used to clone the
735 *MNX1* gene cDNA. The viral packaging vectors pCMV and pMD2.G, described in Pina
736 *et al*. (2008), were used to assemble lentiviral particles using HEK293T cells via
737 transfection using TurboFect Reagent (Invitrogen). Transduction of mES cells was
738 performed overnight by addition of lentivirus to culture medium and washed the

739 following day (Moris *et al.,* 2018). Transduced cells were sorted for positivity to
740 LSSmOrange.

741

742 **Flow cytometry**

743 Surface cell marker analysis was performed by staining using the antibodies listed in
744 Key Resources Table. Disassembled gastruloids cells were resuspended in PBS, 2%
745 FBS and 0.5 mM EDTA and stained at a dilution of 1:100 for primary antibodies for 20
746 minutes at 4°C. When indicated, streptavidin was added at a dilution of 1:200. Analysis
747 was performed on ACEA Novocyte (Agilent) or AttuneNxT (Thermo) analyzers, using
748 the respective software packages. Cell sorting was performed using a CS&T calibrated
749 BDFACS Aria III system (488nm 40 mW, 633nm 20 mW, 405nm 30 mW, and 561nm
750 50 mW), set with the 100µm nozzle at 20PSI and a 4-way purity mask or, in the case
751 of hGx adrenal implants, a Beckman Coulter CytoFlex SRT (488nm 50 mW, 633nm
752 100 mW, 405nm 90 mW, and 561nm 30 mW), set with 100µm nozzle at 15PSI, also
753 using 4-way purity. Single-cell deposition in 96-well plates was performed using single-
754 cell sorting mode. Intact cells were gated on FSC-A vs SSC-A plot, followed by doublet
755 exclusion on FSC-A vs FSC-H and SSC-A vs SSC-H, prior to gating on fluorescent
756 parameters for the markers described in the results.

757

758 **Single cell RNA sequencing of time-resolved hemogenic gastruloids**

759 Gastruloids were collected at different timepoints of the protocol, disassembled and
760 FACS deposited into 96-well plates, either as unsorted (global) or sorted by CD45,
761 CD41 and C-Kit/ScaI markers (sorted). RNA from the cells were extracted from single
762 cells using Smart-seq2 technology at 500Kb and depth of 151 bases. Sequencing
763 reads were quality-checked using FastQC (v0.11.9). We trimmed the samples using
764 trimGalore! (0.6.6) with a cutoff of 30, clipping 15 base pairs and retaining reads of
765 more than 100 bases. Alignment was performed on STAR (2.7.8a) and Mus musculus
766 annotations from GENCODE vM26. Aligned BAM files were annotated using
767 featureCounts (v2.0.1) and count matrices were computed in python by directly
768 accessing the BAM files with gtfparse packages and collapsing lanes.

769 For quality control, based on the histogram of counts and multimodality distributions,
770 we set a minimum count threshold of 200000 counts, a minimum threshold of 1000
771 expressed genes, and a maximum threshold of 20% mitochondrial fraction per cell.
772 We performed the same procedure of the second dataset with a more restrictive
773 threshold of 400000 counts and similar expressed genes and mitochondrial
774 thresholds. Cells that did not pass the quality control metrics were omitted from
775 analysis. We normalized the cells to the mean count number per dataset and applied
776 a plus-one-log transformation of the data before proceeding to the downstream
777 analysis.

778 Dimensionality reduction was performed on feature selection of the gene space using
779 the function scanpy.highly_varying_genes with default parameters. Selection of
780 principal components in principal component analysis (PCA) was performed by
781 heuristic elbow method. Nearest neighbor analysis was constructed by KNN graphs
782 using a correlation metric and 10 nearest neighbors. Data was projected for low
783 dimensional visualization using the UMAP algorithm with default parameters as
784 implemented in scanpy.tl.umap. We used the leiden algorithm as implemented in
785 scanpy.tl.leiden to partition the data into clusters. In order to assess the election of the
786 resolution parameter we used Newman-Girvan modularity as a metric of clustering
787 quality. Differential expressed genes were computed comparing each cluster against
788 the rest using the Wilcoxon test with Benjamini-Hochberg correction.

789

790 **Annotation and projection to additional datasets**

791 Raw counts matrices from Thambyrajah *et al.* (2024) were downloaded and processed
792 following the same pipeline as described above for scRNA-seq of gastruloids to
793 generate UMAP and clustering. We implemented the scmap algorithm to compare our
794 scRNA sequencing of gastruloids with the available datasets. We reduced the
795 dimensionality of the space by selecting highly varying genes from the annotated
796 dataset. Then, we constructed a KNN classifier with correlation metric and computed
797 the nearest neighbors of the target data. If neighbors with correlation metrics below
798 0.7 default standards, the projected cells were not projected onto any cell from the
799 annotated dataset. To visualize the cells over the UMAP plots of the other datasets,

800 we constructed a KNN regressor with three neighbors and a correlation metric.
801 Confusion matrices were used to visualize the overlap between gastruloid clusters and
802 the embryo dataset assigned clusters using scmap.

803

804 **Bulk RNA sequencing**

805 Total RNA was extracted from disassembled gastruloid cells at 216 hours. Sequencing
806 was performed on NovaSeq PE150 platform, at 20M paired-end reads per sample.
807 Tophat2 with Bowtie2 were used to map paired-end reads to the reference Mus
808 musculus genome build GRCm39. GENCODE Release M30 (Frankish *et al.,* 2019)
809 was used as the reference mouse genome annotation. Aligned reads were filtered by
810 quality using samtools (Li, H. *et al.,* 2009) with a minimum threshold set at 30 (q30).
811 Transcript assembly and quantification was achieved using htseq (Putri et al., 2022).
812 Differential expression between sample and control was performed by collapsing
813 technical replicates for each condition using Deseq2 (Love, Huber and Anders, 2014)
814 in R environment (Deseq2 library v 1.32.0). The differential expression was expressed
815 in the form of log2 fold change and filtered by false discovery rate (FDR) of 0.05.

816

817 **Gene list enrichment analyses**

818 Gene ontology (GO) analysis was performed in ExpressAnalyst (available at
819 www.expressanalyst.ca) using the PANTHER or GO Biological Process (BP)
820 repository. GO terms and pathways were filtered by false discovery rate (FDR) with a
821 cut-off of ≤ 0.1 for meaningful association. EnrichR (Xie *et al.,* 2021; Kuleshov *et al.,*
822 2016; Chen, E. Y. *et al.,* 2013) was used for cell type analysis using the Panglao DB
823 (Franzén, Gan and Björkegren, 2019) databases using a FDR threshold of ≤ 0.1 or p
824 value ≤ 0.01, where specified, as well as transcription factor (TF) binding site
825 enrichment using the ENCODE and ChEA Consensus TFs from ChIP database.

826

827 **Gene Set Enrichment Analysis (GSEA)**

828 Custom gene signatures were used as gene sets for GSEA analysis (Subramanian *et*
829 *al.,* 2005) on the GSEA software v4.2.3 on RNA sequencing expression values in
830 counts units. GSEA was ran in 10000 permutations on gene set using the weighted
831 Signal2Noise metric. Enrichment metrics are shown as normalized enrichment score
832 (NES) and filtered by FDR ≤ 0.05. Leading edge genes (LEGs) are genes with a "Yes"
833 values for core enrichment. For AML patient analysis, clinical phenotype and
834 expression data (in counts units) were extracted from the GDC TARGET-AML cohorts
835 in the Therapeutically Applicable Research to Generate Effective Treatments project
836 (TARGET, https://ocg.cancer.gov/programs/target), downloaded from the University
837 of California Santa Cruz (UCSC) Xena public repository (last accessed 31st August
838 2022). Patient samples were selected according to the reported karyotype to include
839 t(7;12), inv(16), MLL, normal karyotype, and t(8;21). GSEA was performed comparing
840 RNA sequencing counts of t(7;12) samples against pooled AML subtypes (inv(16),
841 MLL, normal karyotype and t(8;21)) as "other AML".

842

843 **Real-time polymerase chain reaction (qPCR)**

844 Extracted RNA was reverse-transcribed into complementary DNA (cDNA) using High-
845 Capacity RNA-to-cDNA Kit (Applied Biosystems). QPCR was performed using
846 FastGene 2x IC Green Universal qPCR Mix (Nippon Genetics, Duren, Germany) using
847 primers for human MNX1 (forward 5-GTTCAAGCTCAACAAGTACC-3; reverse 5-
848 GGTTCTGGAACCAAATCTTC-3) (Gulino et al., 2021) and Ppia for endogenous
849 control (forward 5-TTACCCATCAAACCATTCCTTCTG-3; reverse 5-
850 AACCCAAAGAACTTCAGTGAGAGC-3) (Moris *et al.,* 2018). Differential gene
851 expression was calculated using the delta delta Ct (ΔΔCt) method.

852

853 **Polymerase chain reaction (PCR)**

854 Genomic DNA was extracted using Monarch Genomic DNA Purification Kit (New
855 England Biolabs, Hitchin, UK) using manufacturer's instructions. PCR on genomic
856 DNA was performed using Phire PCR Master Mix (Thermo Scientific) to amplify the

857 BFP locus using forward primer 5'-GCACCGTGGACAACCATCACTT-3' and reverse
858 primer 5'-CAGTTTGCTAGGGAGGTCGC-3'.

859

860 **Statistical analysis**

861 Experiments were performed at least in triplicates, unless specified otherwise. Data
862 are plotted to include standard deviation (+/-SD) between replicates. Statistical
863 significance was set at a threshold of p value < 0.05. Statistical analysis was performed
864 in R environment (version 4.1.3) or using GraphPad Prism 8.0 software and is detailed
865 in respective figure legends.

866

867 **Data availability**

868 Raw data as well as processed count matrices and post-processed files from single-
869 cell RNA-seq for the time-resolved data is available at E-MTAB-12148. Bulk RNA-seq
870 of MNX1 overexpressing gastruloids is available at Array Express with accession code
871 E-MTAB-12173. The post-processing was performed in Python on DockerHub:
872 dsblab/single_cell_analysis:0.5. Scripts are available in https://github.com/dsb-
873 lab/blood_gastruloids and Zenodo (https://doi.org/10.5281/zenodo.7053423). The
874 results published here are partly based upon data generated by the Therapeutically
875 Applicable Research to Generate Effective Treatments (TARGET)
876 (https://ocg.cancer.gov/programs/target) initiative, of the Acute Myeloid Leukemia
877 (AML) cohort GDC TARGET-AML. The data used for this analysis are available at
878 https://portal.gdc.cancer.gov/projects and https://xenabrowser.net/.

879

880

881 **ACKNOWLEDGEMENTS**

882 This project was funded by a start-up grant and a BRIEF award from Brunel University
883 London to CP, by a Little Princess Trust (LPT) Project Grant (CCLGA 2023 22 Pina)
884 to CP, and by an ERC Advanced Grant (MiniEmbryoBlueprint 834580) to AMA. LPT

research is funded in partnership with CCLG through the CCLG Charity Research Network. DR received funding from the Royal Society of Biology (MRSB Travel Grant). GTC was funded by grant FPU18/05091 from the Spanish Ministry of Universities. AJ is funded by a Lady Tata Memorial Trust International Scholarship (2022). SvdB was funded by an EMBO Postdoctoral Fellowship (ALTF 195-2021) and is in receipt of a HFSP Postdoctoral Fellowship (LT0047/2022-L). Work in the CP lab was also funded by a KKLF Intermediate Fellowship (KL888), a Leuka John Goldman Fellowship for Future Science (2017-2019), and a Wellcome Trust / ISSF Bridge Funding award at the University of Cambridge (2019). JGO acknowledges financial support from the Spanish Ministry of Science and Innovation and FEDER (grant PGC2018-101251-B-I00), by the Maria de Maeztu Programme for Units of Excellence in R&D (grant CEX2018-000792-M), and by the Generalitat de Catalunya (ICREA Academia programme). MP acknowledges funding from the Italian Ministry of Health (Ricerca Finalizzata 5 per mille and Ricerca Corrente to MP). Library preparation and next-generation sequencing for single-cell RNA-seq analysis were performed by the Single Cell Genomics Group at the National Centre for Genomic Analysis – Centre for Genomic Regulation (CNAG-CRG), Barcelona. The Authors wish to acknowledge Tina Balayo, Ana Filipa Domingues, Shikha Gupta, Oliver Davies, Kristen Place and Remisha Gurung's technical support at different stages of the project.

AUTHOR CONTRIBUTIONS

Conceptualization: CP, AMA; Methodology: CWS, DR, GTC, FP, SvdB, MC,MP, KRK, VH-H, AB, JGO, AMA, CP; Software: GTC, JGO; Validation: CWS, DR, AJ, YC, SvdB, CP; Investigation: CWS, DR, GTC, FP, AJ, YC, LD, CB, G-AI, CP; Formal Analysis: GTC, DR, JGO, CP; Resources: GTC, JC, MP, JGO, AMA; Data curation: DR, GTC, JGO; Writing – Original Draft: CP, DR, GTC; Writing – Review and editing: CP, DR, GTC, AB, JGO, AMA; Visualisation: DR, GTC, CWS, CP; Supervision: CP; Project administration: CP, AMA, JGO; Funding acquisition: CP, AMA.

DECLARATION OF INTERESTS

AMA and CP are co-inventors in the patent application PCT/GB2019/052668: Polarised Three-Dimensional Cellular Aggregates. The other authors have no interests to declare.

SUPPLEMENTAL INFORMATION

Supplemental Figures S1-S7

Supplemental Files S1-S3 (Excel files)

925 **FIGURE LEGENDS**

926

927 **Figure 1 – Hemogenic gastruloids (hGx) produced from mES cells promote hemato-**
928 **endothelial specification with spatio-temporally accurate ontogeny. (A)** Timeline of mES
929 cells assembly and culture into gastruloids over a 216h period with the addition of specified
930 factors for the promotion of hemato-endothelial specification. The 24-hour pre-treatment in
931 2i+LIF was omitted when mES cultures showed no signs of differentiation. **(B)** Imaging of hGx
932 over time from 72h to 216h at 10x magnification, showing the assembly and growth of the 3D
933 structures and the polarization of the Flk-1-GFP marker from 96h; scale bar: 100mm. (**C)** Flow
934 cytometry timecourse analysis of hGx for the presence of CD144 (VE-cadherin), C-Kit, CD41,
935 and CD45 markers; representative plots. **(D)** Quantification of analysis in (D). Violin plots with
936 median and interquartile range of up to 26 replicates; Kruskal-Wallis CD144 ($p=0.0755$), C-Kit
937 ($p<0.0001$), CD41 ($p<0.0342$), CD45 ($p<0.0001$). Shown are significant Dunn's multiple
938 comparisons tests; adjusted q-value. **(E)** Immunostaining of whole individual gastruloids at
939 216h showing the localized expression of CD31 (yellow), C-Kit (turquoise), and CD45 (pink)
940 markers; scale bar: 100μm.

941

942 **Figure 2 – Time-resolved analysis of scRNA-seq of hGxs captures successive waves of**
943 **hematopoietic specification. (A)** UMAP projection of all sequenced cells colored by
944 annotated clusters (left), by sorting markers C-Kit/ScaI, CD41, and CD45, or unfractionated
945 cells ('whole' corresponding to 'all' in panel A) (center), and by hGx culture time (right). **(B)**
946 Heatmap representation of the expression of endothelial and hematopoietic marker genes in
947 individual hGx C-Kit$^+$ cells sorted at 144 and 192h. Cells and genes ordered by unsupervised
948 hierarchical clustering; cell cluster dendrogram in Fig. S2E. **(C)** Heatmap representation of the
949 expression of endothelial and hematopoietic marker genes in individual hGx CD41$^+$ and
950 CD45$^+$ cells sorted at 144 and 192/216h, respectively. Cells and genes ordered by
951 unsupervised hierarchical clustering; cell cluster dendrogram in Fig. S2F. **(D)** Summary
952 representation of the proportion of hGx cells expressing endothelial (endo), erythroid/myeloid
953 (Ery+My), myeloid/lymphoid (My+Ly) and HSPC gene signatures. Endo: *Kdr*; Ery+My:
954 *Epor*±*Gata1*±*Klf1*±*Hbb*±*Hba*±*Hbg* + *Spi1*±*Mpo*±*Anpep*±*Csf1/2/3r*); My+Ly: *Spi1*±*Csf1/2a/3r* +
955 *Ikzf1*±*Ighm/d*±*Igk*); HSPC: *Ptprc* + *Myb*). **(E-G)** Cell type-associations of selected surface
956 markers and transcription factors affiliated with the HE-to-hematopoietic transition (*CD44*,
957 *Procr*, *Etv6*, *Gata2*, *Notch1*, *Tal1*) and with hematopoietic (*Gfi1b*, *Mllt3*, *Runx1*) and definitive
958 hematopoietic (*Ikzf1*, *Hlf*, *Hoxa9*, *Myb*) gene expression programs with **(E)** C-Kit$^+$ (endo-haem)
959 and CD41$^+$/CD45$^+$ (haem), **(F)** endo-haem cells obtained at 144h or 192h, and **(G)** CD41$^+$ or

960 CD45$^+$ cells. Gene expression analysed as present (non-zero expression) or absent (zero-
961 expression). Odds ratio with significance ($-\log_{10}p<1$) estimated by Fisher's test.
962
963 **Figure 3 – HGxs specify cells with pre-HSC characteristics. (A)** Flow cytometry analysis
964 of 216h *Rosa26-BFP::Flk1-GFP* hGx cells for CD41, CD45, c-Kit and CD144 (VE-Cadherin).
965 **(B)** Quantification of pre-HSC-like c-Kit$^+$CD144$^+$ cells within the CD45$^+$CD41^{lo} fraction of 216h-
966 hGx; mean of 4 independent replicate experiments. **(C)** Colony-forming cell (CFC) assay of
967 216h-hGx in multipotential methylcellulose-based medium. GEM: granulocyte-erythroid-
968 monocyte; GM: granulocyte-monocyte; M: monocyte; E: erythroid. Please note that the
969 medium does not include TPO and does not assess the presence of megakaryocytic
970 progenitors. CFC frequency in 3hGx, n=3, mean±SD. **(D)** Representative photographs of
971 colonies quantified in C; scale bar=2mm. **(E)** Schematic representation of the experimental
972 workflow for execution and analysis of hGx implantation in the adrenal gland of
973 immunodeficient mice. **(E)** Histology of implanted and control adrenal glands; paraffin sections
974 stained with hematoxylin & eosin (H&E). Inserts in implanted animal highlight neural (orange)
975 and hematopoietic and renal tubular epithelium (green); inserts in control animal highlight
976 normal cortical (blue) and medullar (red) adrenal architecture, absent in implanted adrenal
977 glands. **(F)** PCR detection of hGx genomic (g)DNA in the bone marrow (BM) and spleen (Spl)
978 of immunodeficient mice 1 month after unilateral adrenal implantation of 3 gastruloids/gland.
979 Analysis of 5 replicates of 100ng gDNA /recipient tissue; control animal was injected
980 unilaterally with PBS in an adrenal gland in parallel with experimental implantation. Reaction
981 positive control used 100ng of gDNA from *Rosa26-BFP*::*Flkj1-GFP* mES cells (mES) used to
982 generate hGxs. NTC: no template control. See Fig. S4F-H for additional details.
983
984 **Figure 4 – HGxs with MNX1 overexpression have increased cellularity and enhanced**
985 **hemogenic endothelial potential. (A)** Cell type enrichment analysis in MNX1-r AML
986 transcriptomes, compared to normal paediatric bone marrow (BM) or other paediatric AML
987 from the TARGET database. GSEA used representative cell type gene sets from the 2021 DB
988 database; bubble plot shows NES scores (color gradient) and statistical significance (–
989 log10(FDR), bubble size). **(B)** Schematic representation of generation of *MNX1*-
990 overexpressing and empty vector (EV) mES cells by lentiviral transduction, with subsequent
991 assembly into hGxs (haemGX). **(C)** Imaging of hGxs with *MNX1* overexpression and EV
992 controls at 10x magnification, showing appropriate assembly and polarization of the Flk1-GFP
993 marker. scale bar: 300mm. **(D)** Size of hGx at each timepoint, determined by the distance
994 between the furthest extreme points in μm. Mean ± SD of 3 replicate experiments; 2-tailed t-

995 test, p< 0.05 (*), 0.001 (**), 0.0001 (***), and 0.00001 (****). **(E-G)** Flow cytometry timecourse
996 analysis of **(E)** C-Kit$^+$, **(F)** CD41$^+$, and **(G)** CD45$^+$ cell abundance in MNX1 and EV hGxs (120-
997 216h). Mean ± standard deviation of 3-7 independent experiments; 2-way ANOVA and Sidak's
998 multiple comparison test significant at p<0.05 for C-Kit$^+$ cells only (construct contribution to
999 variance p=0.0191; 144h comparison *p=0.0190). **(H)** Volcano plots of differentially expressed
1000 genes (DEGs) by bulk RNA-seq between MNX1 and EV hGxs at 144h and 216h. Intersection
1001 of statistically significant DEGs between 144 h and 216 h (shown in Venn diagram on top right)
1002 identified unique DEGs for each condition (labelled in red for 144 h and blue for 216 h). **(J)**
1003 Gene Ontology (GO) term enrichment for differentially upregulated genes between MNX1 and
1004 EV hGx at 144h (red) and 216h (blue), computed in EnrichR using the GO Biological Process
1005 repository. **(I)** Transcription factor (TF) binding site enrichment on differentially upregulated
1006 genes between MNX1 and EV hGx at 144h (red) and 216 hr (blue) using the ENCODE and
1007 ChEA Consensus TFs from ChIP database on EnrichR.
1008
1009 **Figure 5 – MNX1 selects C-Kit$^+$ clonogenic cells and selectively transforms end-stage**
1010 **hGxs. (A)** Representative photographs of serial replating of colony-forming cell assays
1011 initiated from EV or MNX1 cells obtained at 144h of the hGx protocol. **(B)** Quantification of
1012 colony-replating efficiency of EV and MNX1 144h-hGx cells. Mean+SD of n>3 replicates;
1013 Kruskal-Wallis with Dunn's multiple comparison testing at significant q<0.05. **(C)**
1014 Representative photographs of serial replating of colony-forming cell assays initiated from EV
1015 or MNX1 cells obtained at 216h of the hGx protocol. **(D)** Quantification of colony-replating
1016 efficiency of EV and MNX1 216h-hGx cells. Mean+SD of n=5-8 replicates; Kruskal-Wallis with
1017 Dunn's multiple comparison testing at significant q<0.05. **(E)** Flow cytometry analysis of
1018 hemogenic progenitor C-Kit$^+$ cells and myeloid-affiliated CD11b$^+$ cells at first round of plating
1019 of 144h and 216h-hGxs initiated from EV and MNX1 cells; representative plots. **(F)**
1020 Representative flow cytometry plots of serially re-plating MNX1 cells from 144h and 216h-
1021 gastruloids (plate 5 cells).
1022
1023 **Figure 6 – HGxs with *MNX1* overexpression recapitulate MNX1-Acute Myeloid Leukemia**
1024 **patient signatures. (A)** Quantification of human *MNX1* transcripts in FPKM units from RNA-
1025 seq of GX-EV, GX-MNX1, and CFC at 144 hr and 216 hr. Bars represent mean of replicates.
1026 **(B)** Heatmap comparing the expression of all differentially expressed genes (DEGs) between
1027 all conditions (GX-MNX1 144 hr vs GX-EV 144 hr; GX-MNX1 216 hr vs GX-EV 216 hr; CFC
1028 MNX1 144 hr vs GX MNX1 144 hr; CFC MNX1 216 hr vs GX MNX1 216 hr) as Z score.
1029 Hierarchical clustering by Ward D method on Euclidean distances identifies 14 clusters (k).

1030 **(C)** GSEA plots for gene set extracted from k14 from **(B)** against MNX1-r patients RNA-seq
1031 counts vs MLL, core-binding factors (CBF), or other pediatric AML. **(D)** Cell type analysis of
1032 the intersect of leading edge genes LEGs (n=476) from Fig. S5D in MNX1-r patients (k14
1033 LEGs) **(B)** using the Panglao DB 2021 database. **(E)** Representative Giemsa-Wright stained
1034 dissociated CFC replating MNX1 hGx cells from 144h and 216h. Solid black arrows, mast cell
1035 precursors; dashed black arrows, mast cells.

1036 **REFERENCES**

1037 Alexander, T.B. and Mullighan, C.G. (2021) 'Molecular biology of childhood leukemia',
1038 *Annual Review of Cancer Biology,* 5, pp. 95-117.

1039 Ballabio, E., Cantarella, C.D., Federico, C., Di Mare, P., Hall, G., Harbott, J., Hughes, J.,
1040 Saccone, S. and Tosi, S. (2009) 'Ectopic expression of the HLXB9 gene is associated with
1041 an altered nuclear position in t (7; 12) leukaemias', *Leukemia,* 23(6), pp. 1179-1183.

1042 Beccari, L., Moris, N., Girgin, M., Turner, D.A., Baillie-Johnson, P., Cossy, A., Lutolf, M.P.,
1043 Duboule, D. and Arias, A.M. (2018) 'Multi-axial self-organization properties of mouse
1044 embryonic stem cells into gastruloids', *Nature,* 562(7726), pp. 272-276.

1045 Boisset, J., van Cappellen, W., Andrieu-Soler, C., Galjart, N., Dzierzak, E. and Robin, C.
1046 (2010) 'In vivo imaging of haematopoietic cells emerging from the mouse aortic
1047 endothelium', *Nature,* 464(7285), pp. 116-120.

1048 Bolouri, H., Farrar, J.E., Triche Jr, T., Ries, R.E., Lim, E.L., Alonzo, T.A., Ma, Y., Moore, R.,
1049 Mungall, A.J. and Marra, M.A. (2018) 'The molecular landscape of pediatric acute myeloid
1050 leukemia reveals recurrent structural alterations and age-specific mutational interactions',
1051 *Nature medicine,* 24(1), pp. 103.

1052 Bousquets-Muñoz, P., Molina, O., Varela, I., Álvarez-Eguiluz, Á, Fernández-Mateos, J.,
1053 Gómez, A., Sánchez, E.G., Balbín, M., Ruano, D. and Ramírez-Orellana, M. (2024)
1054 'Backtracking NOM1:: ETV6 fusion to neonatal pathogenesis of t (7; 12)(q36; p13) infant
1055 AML', *Leukemia,* , pp. 1-5.

1056 Bowie, M.B., McKnight, K.D., Kent, D.G., McCaffrey, L., Hoodless, P.A. and Eaves, C.J.
1057 (2006) 'Hematopoietic stem cells proliferate until after birth and show a reversible phase-
1058 specific engraftment defect', *The Journal of clinical investigation,* 116(10), pp. 2808-2816.

1059 Brignole, C., Calarco, E., Bensa, V., Giusto, E., Perri, P., Ciampi, E., Corrias, M.V.,
1060 Astigiano, S., Cilli, M. and Loo, D. (2023) 'Antitumor activity of the investigational B7-H3
1061 antibody-drug conjugate, vobramitamab duocarmazine, in preclinical models of
1062 neuroblastoma', *Journal for Immunotherapy of Cancer,* 11(9).

1063 Calvanese, V., Nguyen, A.T., Bolan, T.J., Vavilina, A., Su, T., Lee, L.K., Wang, Y., Lay, F.D.,
1064 Magnusson, M. and Crooks, G.M. (2019) 'MLLT3 governs human haematopoietic stem-cell
1065 self-renewal and engraftment', *Nature,* 576(7786), pp. 281-286.

1066 Cazzola, A., Cazzaniga, G., Biondi, A., Meneveri, R., Brunelli, S. and Azzoni, E. (2021)
1067 'Prenatal origin of pediatric leukemia: lessons from hematopoietic development', *Frontiers in*
1068 *cell and developmental biology,* 8, pp. 618164.

1069 Chandrakanthan, V., Rorimpandey, P., Zanini, F., Chacon, D., Olivier, J., Joshi, S., Kang,
1070 Y.C., Knezevic, K., Huang, Y. and Qiao, Q. (2022) 'Mesoderm-derived PDGFRA cells
1071 regulate the emergence of hematopoietic stem cells in the dorsal aorta', *Nature cell biology,*
1072 24(8), pp. 1211-1225.

1073 Chang, Y., Hummel, S.N., Jung, J., Jin, G., Deng, Q. and Bao, X. (2023) 'Engineered
1074 hematopoietic and immune cells derived from human pluripotent stem cells', *Experimental*
1075 *hematology,* .

1076 Chaudhury, S., O'Connor, C., Cañete, A., Bittencourt-Silvestre, J., Sarrou, E., Prendergast,
1077 Á, Choi, J., Johnston, P., Wells, C.A. and Gibson, B. (2018) 'Age-specific biological and
1078 molecular profiling distinguishes paediatric from adult acute myeloid leukaemias', *Nature*
1079 *communications,* 9(1), pp. 5280.

1080 Chen, E.Y., Tan, C.M., Kou, Y., Duan, Q., Wang, Z., Meirelles, G.V. and Clark, N.R. (2013)
1081 'A. 1126 Ma'ayan, Enrichr: Interactive and collaborative HTML5 gene list enrichment
1082 analysis tool. BMC 1127', *Bioinformatics,* .

1083 Chen, W., O'Sullivan, M.G., Hudson, W. and Kersey, J. (2011) 'Modeling human infant MLL
1084 leukemia in mice: leukemia from fetal liver differs from that originating in postnatal marrow',
1085 *Blood,* 117(12), pp. 3474.

1086 Chia, S.L., Kapoor, S., Carvalho, C., Bajénoff, M. and Gentek, R. (2023) 'Mast cell ontogeny:
1087 From fetal development to life-long health and disease', *Immunological reviews,* 315(1), pp.
1088 31-53.

1089 Ciau-Uitz, A., Pinheiro, P., Gupta, R., Enver, T. and Patient, R. (2010) 'Tel1/ETV6 specifies
1090 blood stem cells through the agency of VEGF signaling', *Developmental cell,* 18(4), pp. 569-
1091 578.

1092 Ciriza, J., Thompson, H., Petrosian, R., Manilay, J.O. and García-Ojeda, M.E. (2013) 'The
1093 migration of hematopoietic progenitors from the fetal liver to the fetal bone marrow: lessons
1094 learned and possible clinical applications', *Experimental hematology,* 41(5), pp. 411-423.

1095 Costa, G., Kouskoff, V. and Lacaud, G. (2012) 'Origin of blood cells and HSC production in
1096 the embryo', *Trends in immunology,* 33(5), pp. 215-223.

1097 Dang, S.M., Kyba, M., Perlingeiro, R., Daley, G.Q. and Zandstra, P.W. (2002) 'Efficiency of
1098 embryoid body formation and hematopoietic development from embryonic stem cells in
1099 different culture systems', *Biotechnology and bioengineering,* 78(4), pp. 442-453.

1100 Deguchi, Y. and Kehrl, J.H. (1991) 'Selective expression of two homeobox genes in CD34-
1101 positive cells from human bone marrow', *Blood,* 78(2), pp. 323-328.

1102 Diaz, M.F., Li, N., Lee, H.J., Adamo, L., Evans, S.M., Willey, H.E., Arora, N., Torisawa, Y.,
1103 Vickers, D.A. and Morris, S.A. (2015) 'Biomechanical forces promote blood development
1104 through prostaglandin E2 and the cAMP–PKA signaling axis', *Journal of Experimental*
1105 *Medicine,* 212(5), pp. 665-680.

1106 Dijkhuis, L., Johns, A., Ragusa, D., van den Brink, Susanne C and Pina, C. (2023)
1107 'Haematopoietic development and HSC formation in vitro: promise and limitations of
1108 gastruloid models', *Emerging Topics in Life Sciences,* 7(4), pp. 439-454.

1109 Dong, Y., Shi, O., Zeng, Q., Lu, X., Wang, W., Li, Y. and Wang, Q. (2020) 'Leukemia
1110 incidence trends at the global, regional, and national level between 1990 and 2017',
1111 *Experimental hematology & oncology,* 9, pp. 1-11.

1112 Dzierzak, E. and Bigas, A. (2018) 'Blood development: hematopoietic stem cell dependence
1113 and independence', *Cell stem cell,* 22(5), pp. 639-651.

1114 Espersen, A.D.L., Noren-Nyström, U., Abrahamsson, J., Ha, S., Pronk, C.J., Jahnukainen,
1115 K., Jónsson, ÓG., Lausen, B., Palle, J. and Zeller, B. (2018) 'Acute myeloid leukemia (AML)
1116 with t (7; 12)(q36; p13) is associated with infancy and trisomy 19: Data from Nordic Society
1117 for Pediatric Hematology and Oncology (NOPHO-AML) and review of the literature', *Genes,*
1118 *Chromosomes and Cancer,* 57(7), pp. 359-365.

1119 Fehling, H.J., Lacaud, G., Kubo, A., Kennedy, M., Robertson, S., Keller, G. and Kouskoff, V.
1120 (2003) 'Tracking mesoderm induction and its specification to the hemangioblast during
1121 embryonic stem cell differentiation', .

1122 Fidanza, A. and Forrester, L.M. (2021) 'Progress in the production of haematopoietic stem
1123 and progenitor cells from human pluripotent stem cells', *Journal of Immunology and*
1124 *Regenerative Medicine,* 13, pp. 100050.

1125 Fitch, S.R., Kimber, G.M., Wilson, N.K., Parker, A., Mirshekar-Syahkal, B., Göttgens, B.,
1126 Medvinsky, A., Dzierzak, E. and Ottersbach, K. (2012) 'Signaling from the sympathetic
1127 nervous system regulates hematopoietic stem cell emergence during embryogenesis', *Cell*
1128 *stem cell,* 11(4), pp. 554-566.

1129 Fowler, J.L., Zheng, S.L., Nguyen, A., Chen, A., Xiong, X., Chai, T., Chen, J.Y., Karigane,
1130 D., Banuelos, A.M. and Niizuma, K. (2024) 'Lineage-tracing hematopoietic stem cell origins
1131 in vivo to efficiently make human HLF HOXA hematopoietic progenitors from pluripotent
1132 stem cells', *Developmental Cell,* 59(9), pp. 1110-1131. e22.

1133 Frankish, A., Diekhans, M., Ferreira, A., Johnson, R., Jungreis, I., Loveland, J., Mudge, J.M.,
1134 Sisu, C., Wright, J. and Armstrong, J. (2019) 'GENCODE reference annotation for the
1135 human and mouse genomes', *Nucleic acids research,* 47(D1), pp. D766-D773.

1136 Franzén, O., Gan, L. and Björkegren, J.L. (2019) 'PanglaoDB: a web server for exploration
1137 of mouse and human single-cell RNA sequencing data', *Database,* 2019, pp. baz046.

1138 Fujimoto, T., Ogawa, M., Minegishi, N., Yoshida, H., Yokomizo, T., Yamamoto, M. and
1139 Nishikawa, S. (2001) 'Step-wise divergence of primitive and definitive haematopoietic and
1140 endothelial cell lineages during embryonic stem cell differentiation', *Genes to Cells,* 6(12),
1141 pp. 1113-1127.

1142 Ganuza, M., Hall, T., Myers, J., Nevitt, C., Sánchez-Lanzas, R., Chabot, A., Ding, J.,
1143 Kooienga, E., Caprio, C. and Finkelstein, D. (2022) 'Murine foetal liver supports limited
1144 detectable expansion of life-long haematopoietic progenitors', *Nature cell biology,* 24(10),
1145 pp. 1475-1486.

1146 Ghiaur, G., Ferkowicz, M.J., Milsom, M.D., Bailey, J., Witte, D., Cancelas, J.A., Yoder, M.C.
1147 and Williams, D.A. (2008) 'Rac1 is essential for intraembryonic hematopoiesis and for the
1148 initial seeding of fetal liver with definitive hematopoietic progenitor cells', *Blood, The Journal*
1149 *of the American Society of Hematology,* 111(7), pp. 3313-3321.

1150 Harrison, K.A., Thaler, J., Pfaff, S.L., Gu, H. and Kehrl, J.H. (1999) 'Pancreas dorsal lobe
1151 agenesis and abnormal islets of Langerhans in Hlxb9-deficient mice', *Nature genetics,* 23(1),
1152 pp. 71-75.

1153 Hooper, M., Hardy, K., Handyside, A., Hunter, S. and Monk, M. (1987) 'HPRT-deficient
1154 (Lesch–Nyhan) mouse embryos derived from germline colonization by cultured cells',
1155 *Nature,* 326(6110), pp. 292-295.

1156 Ivanovs, A., Rybtsov, S., Welch, L., Anderson, R.A., Turner, M.L. and Medvinsky, A. (2011)
1157 'Highly potent human hematopoietic stem cells first emerge in the intraembryonic aorta-
1158 gonad-mesonephros region', *Journal of Experimental Medicine,* 208(12), pp. 2417-2427.

1159 Jakobsson, L., Franco, C.A., Bentley, K., Collins, R.T., Ponsioen, B., Aspalter, I.M.,
1160 Rosewell, I., Busse, M., Thurston, G. and Medvinsky, A. (2010) 'Endothelial cells
1161 dynamically compete for the tip cell position during angiogenic sprouting', *Nature cell*
1162 *biology,* 12(10), pp. 943-953.

1163 Kapeni, C., Nitsche, L., Kilpatrick, A.M., Wilson, N.K., Xia, K., Mirshekar-Syahkal, B.,
1164 Chandrakanthan, V., Malouf, C., Pimanda, J.E. and Göttgens, B. (2022) 'p57Kip2 regulates
1165 embryonic blood stem cells by controlling sympathoadrenal progenitor expansion', *Blood,*
1166 *The Journal of the American Society of Hematology,* 140(5), pp. 464-477.

1167 Kikuchi, K. and Kondo, M. (2006) 'Developmental switch of mouse hematopoietic stem cells
1168 from fetal to adult type occurs in bone marrow after birth', *Proceedings of the National*
1169 *Academy of Sciences,* 103(47), pp. 17852-17857.

1170 Kiselev, V.Y., Yiu, A. and Hemberg, M. (2018) 'scmap: projection of single-cell RNA-seq
1171 data across data sets', *Nature methods,* 15(5), pp. 359-362.

1172 Kuleshov, M.V., Jones, M.R., Rouillard, A.D., Fernandez, N.F., Duan, Q., Wang, Z., Koplev,
1173 S., Jenkins, S.L., Jagodnik, K.M. and Lachmann, A. (2016) 'Enrichr: a comprehensive gene
1174 set enrichment analysis web server 2016 update', *Nucleic acids research,* 44(W1), pp. W90-
1175 W97.

1176 Lacaud, G. and Kouskoff, V. (2017) 'Hemangioblast, hemogenic endothelium, and primitive
1177 versus definitive hematopoiesis', *Experimental hematology,* 49, pp. 19-24.

1178 Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis,
1179 G., Durbin, R. and 1000 Genome Project Data Processing Subgroup (2009) 'The sequence
1180 alignment/map format and SAMtools', *bioinformatics,* 25(16), pp. 2078-2079.

1181 Li, W., Ferkowicz, M.J., Johnson, S.A., Shelley, W.C. and Yoder, M.C. (2005) 'Endothelial
1182 cells in the early murine yolk sac give rise to CD41-expressing hematopoietic cells', *Stem*
1183 *cells and development,* 14(1), pp. 44-54.

1184 Lis, R., Karrasch, C.C., Poulos, M.G., Kunar, B., Redmond, D., Duran, J.G.B., Badwe, C.R.,
1185 Schachterle, W., Ginsberg, M. and Xiang, J. (2017) 'Conversion of adult endothelium to
1186 immunocompetent haematopoietic stem cells', *Nature,* 545(7655), pp. 439-445.

1187 Love, M., Huber, W. and Anders, S. (2014) 'No title', *Moderated estimation of fold change*
1188 *and dispersion for RNA-seq data with DESeq2 Genome Biology 15 (12), 550,* .

1189 Lundin, V., Sugden, W.W., Theodore, L.N., Sousa, P.M., Han, A., Chou, S., Wrighton, P.J.,
1190 Cox, A.G., Ingber, D.E. and Goessling, W. (2020) 'YAP regulates hematopoietic stem cell
1191 formation in response to the biomechanical forces of blood flow', *Developmental cell,* 52(4),
1192 pp. 446-460. e5.

1193 Masetti, R., Vendemini, F., Zama, D., Biagi, C., Pession, A. and Locatelli, F. (2015) 'Acute
1194 myeloid leukemia in infants: biology and treatment', *Frontiers in pediatrics,* 3, pp. 37.

1195 McGrath, K.E., Frame, J.M., Fegan, K.H., Bowen, J.R., Conway, S.J., Catherman, S.C.,
1196 Kingsley, P.D., Koniski, A.D. and Palis, J. (2015) 'Distinct sources of hematopoietic
1197 progenitors emerge before HSCs and provide functional blood cells in the mammalian
1198 embryo', *Cell reports,* 11(12), pp. 1892-1904.

1199 McGrath, K.E. and Palis, J. (2005) 'Hematopoiesis in the yolk sac: more than meets the eye',
1200 *Experimental hematology,* 33(9), pp. 1021-1028.

1201 Medvinsky, A. and Dzierzak, E. (1996) 'Definitive hematopoiesis is autonomously initiated by
1202 the AGM region', *Cell,* 86(6), pp. 897-906.

1203 Medvinsky, A., Rybtsov, S. and Taoudi, S. (2011) 'Embryonic origin of the adult
1204 hematopoietic system: advances and questions', *Development,* 138(6), pp. 1017-1031.

1205 Mercher, T., Raffel, G.D., Moore, S.A., Cornejo, M.G., Baudry-Bluteau, D., Cagnard, N.,
1206 Jesneck, J.L., Pikman, Y., Cullen, D. and Williams, I.R. (2009) 'The OTT-MAL fusion
1207 oncogene activates RBPJ-mediated transcription and induces acute megakaryoblastic
1208 leukemia in a knockin mouse model', *The Journal of clinical investigation,* 119(4), pp. 852-
1209 864.

1210 Miladinovic, O., Canto, P., Pouget, C., Piau, O., Radic, N., Freschu, P., Megherbi, A., Brujas
1211 Prats, C., Jacques, S. and Hirsinger, E. (2024) 'A multistep computational approach reveals
1212 a neuro-mesenchymal cell population in the embryonic hematopoietic stem cell niche',
1213 *Development,* 151(7), pp. dev202614.

1214 Moore, M.A. and Metcalf, D. (1970) 'Ontogeny of the haemopoietic system: yolk sac origin of
1215 in vivo and in vitro colony forming cells in the developing mouse embryo', *British journal of*
1216 *haematology,* 18(3), pp. 279-296.

1217 Moris, N., Edri, S., Seyres, D., Kulkarni, R., Domingues, A.F., Balayo, T., Frontini, M. and
1218 Pina, C. (2018) 'Histone acetyltransferase KAT2A stabilizes pluripotency with control of
1219 transcriptional heterogeneity', *Stem cells,* 36(12), pp. 1828-1838.

1220 Murry, C.E. and Keller, G. (2008) 'Differentiation of embryonic stem cells to clinically relevant
1221 populations: lessons from embryonic development', *Cell,* 132(4), pp. 661-680.

1222 Nafria, M., Bonifer, C., Stanley, E.G., Ng, E.S. and Elefanty, A.G. (2020) 'Protocol for the
1223 generation of definitive hematopoietic progenitors from human pluripotent stem cells', *STAR*
1224 *protocols,* 1(3), pp. 100130.

1225 Nagel, S., Scherr, M., Quentmeier, H., Kaufmann, M., Zaborski, M., Drexler, H.G. and
1226 MacLeod, R. (2005) 'HLXB9 activates IL6 in Hodgkin lymphoma cell lines and is regulated
1227 by PI3K signalling involving E2F3', *Leukemia,* 19(5), pp. 841.

1228 Niakan, K.K., Ji, H., Maehr, R., Vokes, S.A., Rodolfa, K.T., Sherwood, R.I., Yamaki, M.,
1229 Dimos, J.T., Chen, A.E. and Melton, D.A. (2010) 'Sox17 promotes differentiation in mouse
1230 embryonic stem cells by directly regulating extraembryonic gene expression and indirectly
1231 antagonizing self-renewal', *Genes & development,* 24(3), pp. 312-326.

1232 North, T.E., Goessling, W., Peeters, M., Li, P., Ceol, C., Lord, A.M., Weber, G.J., Harris, J.,
1233 Cutting, C.C. and Huang, P. (2009) 'Hematopoietic stem cell development is dependent on
1234 blood flow', *Cell,* 137(4), pp. 736-748.

1235 Palis, J., Robertson, S., Kennedy, M., Wall, C. and Keller, G. (1999) 'Development of
1236 erythroid and myeloid progenitors in the yolk sac and embryo proper of the mouse',
1237 *Development,* 126(22), pp. 5073-5084.

1238 Pastorino, F., Brignole, C., Marimpietri, D., Cilli, M., Gambini, C., Ribatti, D., Longhi, R.,
1239 Allen, T.M., Corti, A. and Ponzoni, M. (2003) 'Vascular damage and anti-angiogenic effects
1240 of tumor vessel-targeted liposomal chemotherapy', *Cancer research,* 63(21), pp. 7400-7409.

1241 Pearson, S., Cuvertino, S., Fleury, M., Lacaud, G. and Kouskoff, V. (2015) 'In vivo
1242 repopulating activity emerges at the onset of hematopoietic specification during embryonic
1243 stem cell differentiation', *Stem Cell Reports,* 4(3), pp. 431-444.

1244 Piau, O., Brunet-Manquat, M., L'homme, B., Petit, L., Birebent, B., Linard, C., Moeckes, L.,
1245 Zuliani, T., Lapillonne, H. and Benderitter, M. (2023) 'Generation of transgene-free
1246 hematopoietic stem cells from human induced pluripotent stem cells', *Cell Stem Cell,* 30(12),
1247 pp. 1610-1623. e7.

1248 Picelli, S., Faridani, O.R., Björklund, ÅK., Winberg, G., Sagasser, S. and Sandberg, R.
1249 (2014) 'Full-length RNA-seq from single cells using Smart-seq2', *Nature protocols,* 9(1), pp.
1250 171-181.

1251 Pina, C., May, G., Soneji, S., Hong, D. and Enver, T. (2008) 'MLLT3 regulates early human
1252 erythroid and megakaryocytic cell fate', *Cell stem cell,* 2(3), pp. 264-273.

1253 Quintard, C., Tubbs, E., Jonsson, G., Jiao, J., Wang, J., Werschler, N., Laporte, C., Pitaval,
1254 A., Bah, T. and Pomeranz, G. (2024) 'A microfluidic platform integrating functional
1255 vascularized organoids-on-chip', *Nature Communications,* 15(1), pp. 1452.

1256 Ragusa, D., Cicirò, Y., Federico, C., Saccone, S., Bruno, F., Saeedi, R., Sisu, C., Pina, C.,
1257 Sala, A. and Tosi, S. (2022) 'Engineered model of t (7; 12)(q36; p13) AML recapitulates
1258 patient-specific features and gene expression profiles', *Oncogenesis,* 11(1), pp. 50.

1259 Ragusa, D., Dijkhuis, L., Pina, C. and Tosi, S. (2023) 'Mechanisms associated with t (7; 12)
1260 acute myeloid leukaemia: from genetics to potential treatment targets', *Bioscience reports,*
1261 43(1), pp. BSR20220489.

1262 Rossi, G., Broguiere, N., Miyamoto, M., Boni, A., Guiet, R., Girgin, M., Kelly, R.G., Kwon, C.
1263 and Lutolf, M.P. (2021) 'Capturing cardiogenesis in gastruloids', *Cell stem cell,* 28(2), pp.
1264 230-240. e6.

1265 Rossi, G., Giger, S., Hübscher, T. and Lutolf, M.P. (2022) 'Gastruloids as in vitro models of
1266 embryonic blood development with spatial and temporal resolution', *Scientific Reports,* 12(1),
1267 pp. 13380.

1268 Rybtsov, S., Batsivari, A., Bilotkach, K., Paruzina, D., Senserrich, J., Nerushev, O. and
1269 Medvinsky, A. (2014) 'Tracing the origin of the HSC hierarchy reveals an SCF-dependent,
1270 IL-3-independent CD43– embryonic precursor', *Stem cell reports,* 3(3), pp. 489-501.

1271 Saorin, G., Caligiuri, I. and Rizzolio, F. (2023) *Microfluidic organoids-on-a-chip: The future of*
1272 *human models.* Elsevier, pp. 41.

1273 Schyrr, F., Alonso-Calleja, A., Vijaykumar, A., Gebhard, S., Sarkis, R., Lopes, S.F., Oggier,
1274 A., De Leval, L., Nombela-Arrieta, C. and Naveiras, O. (2023) 'Adrenal extramedullary
1275 hematopoiesis as an inducible model of the adult hematopoietic niche', *bioRxiv,* , pp.
1276 2023.03. 15.531679.

1277 Souilhol, C., Gonneau, C., Lendinez, J.G., Batsivari, A., Rybtsov, S., Wilson, H., Morgado-
1278 Palacin, L., Hills, D., Taoudi, S. and Antonchuk, J. (2016) 'Inductive interactions mediated by
1279 interplay of asymmetric signalling underlie development of adult haematopoietic stem cells',
1280 *Nature communications,* 7(1), pp. 10784.

1281 Sroczynska, P., Lancrin, C., Pearson, S., Kouskoff, V. and Lacaud, G. (2009) 'In vitro
1282 differentiation of embryonic stem cells as a model of early hematopoietic development',
1283 *Leukemia: Methods and Protocols,* , pp. 317-334.

1284 Steliarova-Foucher, E., Colombet, M., Ries, L.A., Moreno, F., Dolya, A., Bray, F., Hesseling,
1285 P., Shin, H.Y. and Stiller, C.A. (2017) 'International incidence of childhood cancer, 2001–10:
1286 a population-based registry study', *The Lancet Oncology,* 18(6), pp. 719-731.

1287 Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A.,
1288 Paulovich, A., Pomeroy, S.L., Golub, T.R. and Lander, E.S. (2005) 'Gene set enrichment
1289 analysis: a knowledge-based approach for interpreting genome-wide expression profiles',
1290 *Proceedings of the National Academy of Sciences,* 102(43), pp. 15545-15550.

1291 Sugimura, R., Jha, D.K., Han, A., Soria-Valles, C., Da Rocha, E.L., Lu, Y., Goettel, J.A.,
1292 Serrao, E., Rowe, R.G. and Malleshaiah, M. (2017) 'Haematopoietic stem and progenitor
1293 cells from human pluripotent stem cells', *Nature,* 545(7655), pp. 432-438.

1294 Taketani, T., Taki, T., Sako, M., Ishii, T., Yamaguchi, S. and Hayashi, Y. (2008) 'MNX1–
1295 ETV6 fusion gene in an acute megakaryoblastic leukemia and expression of the MNX1 gene
1296 in leukemia and normal B cell lines', *Cancer genetics and cytogenetics,* 186(2), pp. 115-119.

1297 Thaler, J., Harrison, K., Sharma, K., Lettieri, K., Kehrl, J. and Pfaff, S.L. (1999) 'Active
1298 suppression of interneuron programs within developing motor neurons revealed by analysis
1299 of homeodomain factor HB9', *Neuron,* 23(4), pp. 675-687.

1300 Thambyrajah, R., Maqueda, M., Neo, W.H., Imbach, K., Guillén, Y., Grases, D., Fadlullah,
1301 Z., Gambera, S., Matteini, F. and Wang, X. (2024) 'Cis inhibition of NOTCH1 through
1302 JAGGED1 sustains embryonic hematopoietic stem cell fate', *Nature Communications,* 15(1),
1303 pp. 1604.

1304 Tosi, S., Kamel, Y.M., Owoka, T., Federico, C., Truong, T.H. and Saccone, S. (2015)
1305 'Paediatric acute myeloid leukaemia with the t (7; 12)(q36; p13) rearrangement: a review of
1306 the biological and clinical management aspects', *Biomarker research,* 3(1), pp. 1.

1307 Turner, D.A., Girgin, M., Alonso-Crisostomo, L., Trivedi, V., Baillie-Johnson, P., Glodowski,
1308 C.R., Hayward, P.C., Collignon, J., Gustavsen, C. and Serup, P. (2017) 'Anteroposterior
1309 polarity and elongation in the absence of extra-embryonic tissues and of spatially localised
1310 signalling in gastruloids: mammalian embryonic organoids', *Development,* 144(21), pp.
1311 3894-3906.

1312 van den Brink, Susanne C, Alemany, A., van Batenburg, V., Moris, N., Blotenburg, M., Vivie,
1313 J., Baillie-Johnson, P., Nichols, J., Sonnen, K.F. and Martinez Arias, A. (2020) 'Single-cell
1314 and spatial transcriptomics reveal somitogenesis in gastruloids', *Nature,* 582(7812), pp. 405-
1315 409.

1316 Van den Brink, Susanne C, Baillie-Johnson, P., Balayo, T., Hadjantonakis, A., Nowotschin,
1317 S., Turner, D.A. and Martinez Arias, A. (2014) 'Symmetry breaking, germ layer specification
1318 and axial organisation in aggregates of mouse embryonic stem cells', *Development,* 141(22),
1319 pp. 4231-4242.

1320 van den Brink, Susanne C and van Oudenaarden, A. (2021) '3D gastruloids: A novel frontier
1321 in stem cell-based in vitro modeling of mammalian gastrulation', *Trends in cell biology,* 31(9),
1322 pp. 747-759.

1323 Veenvliet, J.V., Bolondi, A., Kretzmer, H., Haut, L., Scholze-Wittler, M., Schifferl, D., Koch,
1324 F., Guignard, L., Kumar, A.S. and Pustet, M. (2020) 'Mouse embryonic stem cells self-
1325 organize into trunk-like structures with neural tube and somites', *Science,* 370(6522), pp.
1326 eaba4937.

1327 Vo, L.T. and Daley, G.Q. (2015) 'De novo generation of HSCs from somatic and pluripotent
1328 stem cell sources', *Blood, The Journal of the American Society of Hematology,* 125(17), pp.
1329 2641-2648.

1330 Waraky, A., Östlund, A., Nilsson, T., Weichenhan, D., Lutsik, P., Bähr, M., Hey, J., Tunali,
1331 G., Adamsson, J. and Jacobsson, S. (2024) 'Aberrant MNX1 expression associated with t (7;
1332 12)(q36; p13) pediatric acute myeloid leukemia induces the disease through altering histone
1333 methylation', *Haematologica,* 109(3), pp. 725.

1334 Weichenhan, D., Riedel, A., Sollier, E., Toprak, U.H., Hey, J., Breuer, K., Wierzbinska, J.A.,
1335 Touzart, A., Lutsik, P. and Bähr, M. (2023) 'Altered enhancer-promoter interaction leads to
1336 MNX1 expression in pediatric acute myeloid leukemia with t (7; 12)(q36; p13)', *bioRxiv,* , pp.
1337 2023.09. 13.557546.

1338 Wildenhain, S., Ingenhag, D., Ruckert, C., Degistirici, O., Dugas, M., Meisel, R., Hauer, J.
1339 and Borkhardt, A. (2012) 'Homeobox protein HB9 binds to the prostaglandin E receptor 2
1340 promoter and inhibits intracellular cAMP mobilization in leukemic cells', *The Journal of*
1341 *biological chemistry,* 287(48), pp. 40703-40712 Available at: 10.1074/jbc.M111.308965 [doi].

1342 Xie, Z., Bailey, A., Kuleshov, M.V., Clarke, D.J., Evangelista, J.E., Jenkins, S.L., Lachmann,
1343 A., Wojciechowicz, M.L., Kropiwnicki, E. and Jagodnik, K.M. (2021) 'Gene set knowledge
1344 discovery with Enrichr', *Current protocols,* 1(3), pp. e90.

1345 Yokomizo, T., Ideue, T., Morino-Koga, S., Tham, C.Y., Sato, T., Takeda, N., Kubota, Y.,
1346 Kurokawa, M., Komatsu, N. and Ogawa, M. (2022) 'Independent origins of fetal liver
1347 haematopoietic stem and progenitor cells', *Nature,* 609(7928), pp. 779-784.

1348 Yoshimoto, M., Kosters, A., Cornelius, S., Valiente, N., Cheng, H., Latorre, A., Nishida, C.,
1349 Ghosn, E.E. and Kobayashi, M. (2022) 'Mast cell repopulating ability is lost during the
1350 transition from pre-HSC to FL HSC', *Frontiers in Immunology,* 13, pp. 896396.

1351 Zovein, A.C., Hofmann, J.J., Lynch, M., French, W.J., Turlo, K.A., Yang, Y., Becker, M.S.,
1352 Zanetta, L., Dejana, E. and Gasson, J.C. (2008) 'Fate tracing reveals the endothelial origin
1353 of hematopoietic stem cells', *Cell stem cell,* 3(6), pp. 625-636.

1354