Dynamic Tracking of Native Polyclonal Hematopoiesis in Adult Mice

Suying Liu[1,2], Sarah E. Adams[1,2], Haotian Zheng[3], Juliana Ehnot[1,2], Seul K. Jung[1,2], Greer Jeffrey[1,2], Theresa Menna[1,2], Louise E. Purton[4,5], Hongzhe Lee[3], and Peter Kurre[1,2*]

[1] Comprehensive Bone Marrow Failure Center, Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA

[2]Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

[3]Department of Biostatistics, Epidemiology and informatics, University of Pennsylvania, Philadelphia, PA 19104, USA

[4]Stem Cell Regulation Unit, St. Vincent's Institute of Medical Research, Fitzroy, VIC 3065, Australia

[5]Department of Medicine, The University of Melbourne, Parkville, VIC 3052, Australia

*Correspondence: kurrep@chop.edu

Abstract

Hematopoietic dysfunction has been associated with a reduction in the number of active precursors. However, precursor quantification at homeostasis and under diseased conditions is constrained by the scarcity of available methods. To address this issue, we optimized a method for quantifying a wide range of hematopoietic precursors. Assuming the random induction of a stable label in precursors following a binomial distribution, the estimation depends on the inverse correlation between precursor numbers and the variance of precursor labeling among independent samples. Experimentally validated to cover the full dynamic range of hematopoietic precursors in mice (1 to 10^5), we utilized this approach to demonstrate that thousands of precursors, which emerge after modest expansion during fetal-to-adult transition, contribute to native and perturbed hematopoiesis. We further estimated the number of precursors in a mouse model of Fanconi Anemia, showcasing how repopulation deficits can be segregated into autologous (cell proliferation) and non-autologous causes (lack of precursor). Our results support an accessible and reliable approach for precursor quantification, emphasizing the contemporary perspective that native hematopoiesis is highly polyclonal.

Introduction

The continuous self-renewal and differentiation of hematopoietic stem and progenitor cells (HSPCs) are fundamental to blood production. Even though rare cases exist where a single HSPC clone supports hematopoiesis, the HSPC population contributing to homeostatic hematopoiesis is usually highly polyclonal [1]. For example, the number of HSPCs actively participating in white blood cell production was estimated to range from 20,000 to 200,000 in humans [2,3]. Conversely, a decline in the number of active hematopoietic precursors has been linked to hematopoietic dysfunctions. For example, in humans older than 70 years of age, hematopoietic dysfunction coincides with an abrupt reduction in the HSPCs population actively contributing to blood production [3].

These observations underscore the association between a low number of active hematopoietic precursors and hematological dysfunction. Yet, few methods are suitable for quantifying active hematopoietic precursors in a native environment [2–7]. In mice, methods utilizing *in situ* barcodes often suffer from barcode homoplasy, preventing total precursor number estimation [8,9]. In humans, estimations based on somatic mutations and computational modeling carry high uncertainty [2,3]. The absence of precise precursor quantification methods in their natural environment hampered the study of precursor numbers across different conditions and their potential usage as predictive functional markers.

In mice, the quantification of developmental hematopoietic precursors has been achieved using mice with a Confetti cassette [10]. This cassette can randomly recombine and express one of four fluorescence proteins (FPs, being RFP, CFP, YFP, or GFP) upon Cre induction, resulting in variability in Confetti expression pattern in mice inversely correlating with precursor numbers (Figure 1A) [10]. While suitable for developmental hematopoiesis, this method has limited linear range (50-2500), impeding its potential application to adult hematopoiesis or clonal hematopoiesis. To measure hematopoietic precursors in various conditions, expansion of the detection range is desired.

Interestingly, the random induction of Confetti colors in HSPCs bears similarity with X-chromosome inactivation (XCI). In XCI, one of the X chromosomes is randomly inactivated in precursor cells, a process that adheres to a binomial distribution. Formula derivation from binomial distribution implied that higher XCI variance among individuals correlates with a smaller number of precursors at the time of inactivation (Figure 1B). As XCI is faithfully maintained in the progeny, this correlation has been used to estimate precursor numbers from mature blood cells in human and mouse, leading to the discovery of the first mutation (Tet2) that contributes to clonal hematopoiesis [6,11–13]. Nonetheless, XCI occurs exclusively in females, limiting its applicability in males. Additionally, XCI takes place during early development when few precursors are present, resulting in low sensitivity in adulthood [14].

Building on insights from Confetti mice and XCI studies, we investigated whether random induction of Confetti FPs in precursor cells can be similarly modeled by a binomial distribution, thus the variance of FPs inversely correlates with precursor numbers. Based on this relationship, we asked if we could broaden the measurable range of precursor number, which would allow us to measure the numbers outside the existing correlation range.

Examining the premises of binomial distribution and setting some assumptions, we established that the random induction of FPs among a group of mice or cells could be modeled by a binomial distribution. Experimental validation established a broader linear range, covering the full spectrum of precursor numbers (1~10^5) and overcoming the prior range limit. We leveraged this correlation to probe the number of hematopoietic precursors at homeostasis, post-myeloablation, during developmental expansion, and in a mouse model of inherited bone marrow failure.

Results

Binomial distribution underlies the inverse linear correlation between FP% variance and the precursor numbers

To broaden the limited correlation range, we aimed to elucidate the mathematical relationship between variance and precursor numbers [10]. Inspired by XCI, we asked how the stochastic induction of Confetti FPs in precursors may satisfy the premises of binomial distribution [15,16]. We found that when we assume the number of precursors remain the same for the condition tested, then induction of FPs meet the required premises : (1) the number of precursors, denoted as *n*, remains constant within a group of mice or cells; (2) each observation (e.g. each precursor) is independent; (3) each observation yields one of two outcomes (e.g., it signifies the presence or absence of RFP in a precursor); (4) the probability of success, represented as *p* (e.g., the likelihood of a precursor expressing RFP upon Confetti induction), remains consistent across all observations.

Should the induction of a FP adhere to a binomial distribution, precursor numbers can be estimated using the following equation:

$$\lg(\hat{n}) = \lg\left[\frac{(1-\widehat{FP_{mean}}\%)}{\widehat{FP_{mean}}\%}\right] - 2 \times \lg\left(\widehat{CV_{FP}}\right) \qquad \text{(equation 1, see Methods)}$$

where $\hat{n}$ signifies the estimated number of precursors, $\widehat{FP_{mean}}\%$ represents the estimated probability of a precursor being one given FP (e.g., $\widehat{RFP_{mean}}\%$), and $\widehat{CV_{FP}}$ denotes the estimated coefficient of

variance (CV) of one given FP (e.g., $\widehat{CV_{RFP}}$). CV is the standard deviation divided by the mean, utilized by the empirical correlation formula.

The equation proposed a direct linear relationship between the precursor number and the CVs of individual FP, supporting the empirical correlation formula [10]. Indeed, by analyzing a published dataset providing precursor numbers and the corresponding distribution of FPs, we found that the correlation between *n* and individual CVs of FP exhibits superior fit (>0.93), higher than the empirical correlation equation based on all three CVs (0.75) (Figure S1A) [10]. Moreover, according to equation 1, the y-intercept of correlation between precursor number and CV was influenced by FP_{mean}%, while the slope was the same (-2) regardless of FP_{mean}%. Given the unequal value of RFP_{mean}%, CFP_{mean}%, and YFP_{mean}% in the published dataset, their correlation equations showed significantly distinct y-intercepts ($p<0.0001$) (Figure S1B). In contrast, the slopes of these correlation equations were similar ($p=0.49$, with all 95% confidence intervals encompassing -2, the theoretical value).

In conclusion, the FP induction in precursor cells can be modeled by a binomial distribution with the assumption that precursor numbers are constant among a group of mice. This sets the mathematical basis for the inverse linear correlation between variance of FP% and the number of precursors.

<u>A broad range of precursor numbers correlates with variance of FP%</u>

While the correlation formula derived from binomial distribution does not impose any range limitation, experimental errors may confound the measurement of variance. We next aimed to experimentally confirm this correlation. To streamline the calculation, we used standard deviation instead of CV to compute variance. The correlation between standard deviation and precursor numbers is expressed through the following equation:

$$\lg(\hat{n}) = \lg\left[\widehat{FP_{mean}}\% \times (1 - \widehat{FP_{mean}}\%)\right] - 2 \times \lg(\widehat{\sigma_{FP}}) \quad \text{(equation 2, see Methods)}$$

where $\widehat{\sigma_{FP}}$ denotes the estimated standard deviation of a given FP% (e.g. $\widehat{\sigma_{RFP}}$).

Given this equation, $\lg(\widehat{\sigma_{FP}})$ exhibits a linear and inverse correlation with $\lg(\hat{n})$. Indeed, $\lg(\widehat{\sigma_{RFP}})$, $\lg(\widehat{\sigma_{CFP}})$ and $\lg(\widehat{\sigma_{YFP}})$ exhibited linear and inverse correlation with $\lg(\hat{n})$ in the published data (Figure S1C) [10].

To simplify the validation, we used a two-color cell model HL-60 bearing one of the two FPs (BFP and GFP, GFP represents non-BFP) (Figure 1C). Although the Confetti cassette recombination can

142 generate one of four colors, our estimation is based on a given FP (e.g. a precursor expresses RFP
143 or not), making this simplification justifiable.
144
145 We first proved that neither BFP nor GFP HL-60 cells had a competitive growth advantage over the
146 other, ensuring that HL-60 progenies mirrored the seeding population (Figure S1D). We next sorted
147 one to 10^5 cells into individual wells and allowed them to proliferate for at least three generations
148 before assessing BFP% at the end of the culture (Figure 1C). We seeded up to 10^5 cells as $1.4x10^5$
149 non-HSC LSK and $5.2x10^3$ active HSCs were previously estimated in an adult female mouse, which
150 corresponded to the highest possible precursor numbers in mice [17]. We considered the survival rate
151 of cells in wells receiving a single HL-60 cell (53% to 60%), to ensure the accuracy of the number of
152 cells seeded.
153
154 Consistent with the proposed correlation, as seeded number increased, BFP% standard deviation
155 decreased (Figure 1D and E). The variance of BFP% displayed an inverse correlation with the seeding
156 numbers (R^2 = 0.996) (Figure 1F). After normalizing for cell survival rates, the calculated numbers
157 closely aligned with the expected numbers (Figure 1G). We had anticipated that measurement of small
158 variance in high precursor numbers may be confounded by experimental errors. Yet, at 10^5 seeded
159 cells, the estimated number of precursors did not plateau, suggesting minimal impact from
160 experimental error even at 10^5 cells.
161
162 Experimental practices for accurate precursor number measurement
163 It is crucial to ensure the accuracy of variance measured because variance of FP% is solely used for
164 estimation. Based on the data from the two-color HL-60 model, we concluded that there were at least
165 three experimental practices required for accurate cell estimates: (1) exclusion of outliers; (2) sufficient
166 flow cytometry recorded events; (3) sufficient sample size per group. Both outliers and insufficient
167 recorded events inaccurately inflate sample variance, leading to underestimation of precursor
168 numbers (Figure S1E and F). We found that the minimum recorded events increase in tandem with
169 the seeded number, in contrary to the reported 500-event threshold (Figure S1F) [10].
170
171 While having a small sample size per biological replicate is possible, the variability of estimations
172 among replicates would be high (Figure S1G). Practically, achieving a large sample size (>20) is
173 cumbersome in mice. To determine a feasible sample size for relatively accurate estimates, we
174 resampled datapoints to calculate the amount of variability reduction as sample size increased (Figure
175 S1H). Our analyses revealed that five samples per group were sufficient to dramatically reduce error,
176 where additional increase in sample size led to marginal error reduction. We therefore used at least
177 five mice per biological replicates for our precursor estimations.

178

179 Inducible HSC-Scl-CreER^T-targeted Confetti labeling in the active blood precursors at the adult stage

180 To label hematopoietic precursors and quantify their number *in vivo*, we considered the Cre mouse
181 line to be crossed with Confetti. Historically, hematopoietic precursors were thought to be
182 hematopoietic stem cells (HSCs) capable of long-term repopulation. However, recent studies indicate
183 that multipotent progenitors (MPPs) also contribute alongside HSCs in maintaining blood production
184 [18–21]. Hence, both HSCs and MPPs should be labeled.

185

186 Among limited Cre mouse lines capable of labeling HSCs and MPPs simultaneously at the adult stage
187 (such as Rosa26^{CreERT2}, Mx1-Cre, and HSC-Scl-CreER^T), we chose HSC-Scl-CreER^T because of its
188 preference for labeling HSPCs (immunophenotypically defined as Lin$^-$Sca-1$^+$cKit$^+$ (LSK)) [22]. To
189 validate the specificity of HSC-Scl-CreER^T, we generated mice possessing a single Confetti allele and
190 homozygous HSC-Scl-CreER^T alleles. We then examined Confetti expression one day after two-day
191 tamoxifen administration. We observed Confetti expression in the LSK population, consistent with prior
192 reports of HSC-Scl-CreER^T activity (Figure 2A, Figure S2A-C). Confetti expression was additionally
193 detected in T cells, NK cells, CD41$^+$ cells, non-inflammatory monocytes (SSC-A^{low}Ly6C$^-$Ly6G$^-$CD11b$^+$),
194 common myeloid progenitors (CMP, Lin$^-$Sca-1$^-$cKit$^+$CD16/32$^{-/low}$CD34$^+$), and granulocyte-monocyte
195 progenitors (GMP, Lin$^-$Sca-1$^-$cKit$^+$CD16/32$^+$CD34$^+$) [20] (Figure S2D-G). Considering the relatively
196 shorter lifespan of these cells compared to HSC/MPP, the labeling in these cells should minimally
197 impact precursor calculations, especially with a chase period post-labeling. While the Confetti
198 induction in the relatively long-lived T cells precludes estimation based on mature T cells, we
199 concluded that HSC-Scl-CreER^T remained a practical choice for labeling active hematopoietic
200 precursors.

201

202 The stability of Confetti labeling after induction is required for binomial distribution. Therefore, the
203 absence of background Cre activity after Confetti induction is critical. We determined the background
204 Cre activity in HSC-Scl-CreER^T/Confetti animals by two approaches: (1) detecting Confetti expression
205 in non-induced animals; (2) identifying cells co-expressing two Confetti FPs months after tamoxifen
206 induction due to cassette "flipping" from background Cre activity. We observed no Confetti-expressing
207 cells without tamoxifen treatment in mice up to 50 weeks old and minimal co-expression of two Confetti
208 FPs in induced HSC-Scl-CreER^T/Confetti animals (Figure S3A-B). Conversely, Vav-Cre/Confetti
209 animals constitutively express Cre, resulting in ~15% cells co-expressing two Confetti colors in
210 peripheral blood (PB) T cells (Figure S3C-D). Both pieces of evidence support minimal Cre
211 background activity in HSC-Scl-CreER^T and its feasibility to be used as a Cre driver for precursor
212 number calculation.

213

214 Variance of FP% inversely correlates with precursor numbers *in vivo*

215 To ascertain the inverse correlation between variance of FP% and the number of hematopoietic
216 precursors *in vivo*, we first investigated the variance of FP% in mice hosting various numbers of
217 hematopoietic precursors. We generated these mice by non-competitively transplanting $0.25x10^6$ or
218 $4x10^6$ induced $CD45.2^+$ HSC-Scl-$CreER^T$/Confetti mouse bone marrow (BM) cells into $CD45.1^+$
219 recipient mice (Figure 2B). Given the constant frequency of precursors in the donor BM ($1/10^5$) and
220 the varying doses of donor BM cell received, these mice received approximately 2.5 ("low precursor
221 number") or 40 ("high precursor number") precursors [15,23–25]. While similar experiments
222 transplanting defined numbers of precursors have been reported, the estimation for precursor
223 numbers lower than 50 has not been explored.

224

225 Due to the scarcity of hematopoietic precursors in the BM, the precursor number seeded in recipients
226 followed Poisson distribution instead of being constant, violating premise (1) of binomial distribution.
227 Nonetheless, our simulations showed that the seeding number variation among recipient mice was
228 relatively small, therefore the variance resulting from precursor number differences still dominated and
229 inversely correlated with precursor numbers. A minor problem was that the estimated numbers were
230 not at one-to-one ratio to expected numbers at low precursor number range (<10, see Methods)
231 (Figure S4A).

232

233 As cells were not 100% labeled with Confetti, for clarity, we use "FP_{PB}%" or "FP_{BM}%" to represent the
234 frequency of a given FP in the PB or BM, and use "FP_{Conf}%" to represent the frequency of a given FP
235 in the $Confetti^+$ population. For total precursor number calculations, we normalized the number
236 calculated from the Confetti-labeled population to Confetti labeling efficiency.

237

238 As expected, we observed higher RFP_{Conf}% variance across all PB cell types in mice belonging to the
239 "low precursor number" groups (Figure 2C and Figure S4B). Since variance of any FP following a
240 binomial distribution inversely correlates with precursor numbers, we also observed higher CFP_{Conf}%
241 and YFP_{Conf}% variance in the "low precursor number" groups (Figure S4D-E). Following the same
242 principle, we observed higher RFP_{PB}%, CFP_{PB}%, YFP_{PB}% and Confetti% variance in "low precursor
243 number" mice (Figure S4F).

244

245 Applying equation 2, we found that the precursor numbers of B cells and myeloid cells were noticeably
246 higher in the first two months than at four months post-transplantation, suggesting transient progenitor
247 contribution at early time points (Figure 2D). At four months post-transplantation, the estimated
248 precursor numbers aligned with the expected values (in myeloid cells, 94 for "high precursor number"
249 groups, three for "low precursor number" groups, Figure 2D). Again, for any FP following binomial

250 distribution, variance of FP% inversely correlates with precursor numbers and thus can be leveraged
251 to calculate precursor numbers (Methods). Since different FPs measured the same population of
252 precursors, we expected the estimations from different FPs to be similar. Indeed, precursor numbers
253 derived from RFP_{Conf}% highly correlated with those estimated from YFP_{Conf}% or CFP_{Conf}% (R^2 = 0.885
254 for CFP% in Confetti$^+$, R^2 = 0.769 for YFP% in Confetti$^+$), as well as those from RFP_{PB}%, YFP_{PB}% and
255 CFP_{PB}% (Figure 2E, Figure S4G).
256
257 Consistent with PB data, higher variance of RFP_{Conf}% was also observed in the BM HSPCs of "low
258 precursor number" mice (Figure S5A). The precursor numbers estimated from various HSPC
259 subpopulations aligned well with each other and were consistent within the same group (Figure 2F).
260 However, in "high precursor number" groups, estimates from BM HSPCs were lower than those
261 derived from PB myeloid cells (27 for BM HSPC, 94 for PB myeloid cells, Figure 2F). This discrepancy
262 may reflect uneven seeding of precursors to the BM throughout the body after transplantation, and the
263 fact that we only sampled a part of the BM (femur, tibia, and pelvis). In summary, we validated the
264 inverse correlation between variance of FP% and precursor numbers *in vivo*. Moreover, the
265 quantification of precursor numbers was feasible even for precursor numbers outside the empirical
266 correlation range.
267
268 Cell frequency measurements fail to reflect the differences in precursor numbers
269 Given that transplantation was performed non-competitively and the recipients were lethally irradiated,
270 we expected to see minimal differences in donor chimerism between two groups, despite drastic
271 differences in donor precursor numbers. Indeed, although we observed substantially lower donor
272 chimerism in mice belonging to "low precursor number" groups during the first two months post-
273 transplantation, the chimerism differences were very small by four months post-transplantation
274 (5.5±2.8%, Figure 2G and Figure S4B). The initial differences in donor precursor numbers did not
275 affect PB cell frequencies except for the first month post-transplantation, when higher PB myeloid
276 frequencies was observed in the "low precursor number" groups (Figure 2H). This suggests
277 myelopoiesis is preferred when very few precursors are available after irradiation-mediated injury
278 (Figure 2C).
279
280 In BM, nucleated cell counts, donor chimerism and HSPC frequencies were mostly similar regardless
281 of donor precursor numbers, suggesting that even extremely low numbers of hematopoietic precursors
282 can still repopulate effectively a non-competitive environment (Figure S5B-D). Therefore, in cases
283 where hematopoietic precursors expand or attrite without competition, stem cell frequencies are less
284 informative to study precursor activity. Measuring precursor numbers is more meaningful as low
285 number of active precursors may be hindered by compensatory proliferation.

286

287 Thousands of hematopoietic precursors contribute to native hematopoiesis

288 While the empirical formula, which measures 50-2500 precursors, supports quantification of
289 precursors labeled at fetal stages, it may not be applicable to measure native hematopoiesis as the
290 total precursor number may exceed 2500. Having validated the feasibility for measuring precursor
291 numbers *in vivo* by Confetti animals, we next sought to investigate the number of active precursors
292 contributing to native hematopoiesis. In HSC-Scl-CreER^T/Confetti animals, we opted to label a smaller
293 portion of HSPCs with two-day treatment despite the potential to label up to 60% HSPCs with Confetti
294 (CFP+YFP+RFP) by 14-day tamoxifen treatment (Figure 3A, Figure S6A). This precautionary
295 measure aimed to mitigate potential toxicity arising from prolonged tamoxifen treatment.

296

297 Unexpectedly, we observed a successive decline of CFP_{Conf}% and CFP_{PB}% in some animals, which
298 distorted the distribution of CFP_{Conf}% (Figure S6B and C). A similar decline was not observed in
299 RFP_{Conf}% or YFP_{Conf}%, nor the $RFP_{RPF+YFP}$% (calculated by RFP_{PB}%/(RFP_{PB}%+YFP_{PB}%)) (Figure
300 S6D and E). Declining CFP_{Conf}% was also not observed in animals where Confetti was induced during
301 fetal development, suggesting a potential immune response to CFP in adult-induced animals (Figure
302 S6F). Since only ~2% of myeloid cells were labeled with CFP, we reasoned that the decline of CFP in
303 some animals was unlikely to affect the calculation of total precursor numbers based on other FPs.
304 Therefore, we solely relied on the variance of $RFP_{RPF+YFP}$% for adult precursor calculations. In
305 transplantation studies, precursor numbers calculated with variance of $RFP_{RFP+YFP}$% linearly
306 correlated with those calculated with variance of RFP_{Conf}% (Figure S4G).

307

308 To ensure accuracy, we focused on PB myeloid cells, as (1) Confetti labeling of T cells resulted from
309 direct induction by HSC-Scl-CreER^T, but not differentiation from HSPCs (Figure S2D); (2) Confetti
310 labeling (RFP and YFP) of B cells did not plateau at seven months post-induction (Figure 3B); (3)
311 myeloid cells are of shorter lifespan. In myeloid cells, we focused on precursor numbers after four
312 months post-induction, when their labeling reached a plateau (Figure 3B). We also examined average
313 $RFP_{RFP+YFP}$% (Figure 3C). Its stability suggested an equal contribution of RFP^+ and YFP^+ cells to the
314 PB, a prerequisite for estimating precursor numbers from FP% distributions in the progenies (PB
315 myeloid cells). The average $RFP_{RFP+YFP}$% in BM HSPC subpopulations mirrored those observed in
316 PB, supporting non-biased differentiation of RFP^+ and YFP^+ precursor cells (Figure 3C).

317

318 Fitting data to equation 2, we estimated an average of 2667 precursors contributing to native
319 myelopoiesis (Figure 3D, average of numbers at five to seven months post-induction after myeloid cell
320 labeling plateaued). This number closely aligns with the clone number estimated by transposon-based
321 barcodes and statistic method in granulocytes (831/30% = 2770) [1].

322

323 The average myeloid precursor number at the time of BM analysis (1958) matched the average
324 precursor number calculated from BM myeloid progenitors (MP, Lin$^-$Sca-1$^-$cKit$^+$) and HSPCs (1773
325 and 1917), but it was five-fold higher than that of LT-HSC (Figure 3E). Although LT-HSCs were
326 conventionally believed to sustain steady-state hematopoiesis, recent studies suggest that ST-HSC
327 and MPPs persist long-term and are the primary contributors to adult hematopoiesis [1,18,20,26]. The
328 observed discrepancy between the number of precursors contributing to LT-HSC and those
329 contributing to MPPs indicates that at least some MPPs were not replaced by progenitors directly
330 differentiated from LT-HSC at seven months post-induction, affirming the long life of MPPs (Figure
331 3E). Furthermore, the fact that PB myeloid and BM MP precursor numbers were closer to those of
332 MPPs than LT-HSC confirm active and long-term (at least seven months) contributions from MPPs to
333 steady-state myelopoiesis.
334

335 Of note, the quantity of precursor does not necessarily correlate with the frequency of cell type in the
336 BM. For example, the cell frequency of MPP2 (Lineage$^-$cKit$^+$Sca-1$^+$Flk2$^-$CD48$^+$CD150$^+$) in BM was
337 very low, but MPP2 comprised a modest number of precursors, making its frequency-to-precursor-
338 number ratio the lowest among the HSPC subtypes (Figure 3F and S6G). The frequency-to-precursor-
339 number ratio reflects how well precursors of a particular type proliferate to expand their absolute cell
340 count. The low frequency-to-precursor-number ratio of MPP2 suggests that it expands more poorly
341 than other HSPC subtypes.
342

343 In summary, we detected thousands of hematopoietic precursors contributing to adult hematopoiesis.
344 At the time of BM analysis, the number of PB myeloid precursors was comparable to those observed
345 in BM MP and HSPCs.
346

347 <u>Precursor numbers determined by FP% variance confirmed reduced clonality of progenitors after</u>
348 <u>myeloablation</u>

349 Myeloablation through 5-fluorouracil (5-FU) treatment depletes most actively cycling cells, forcing
350 quiescent stem cells to proliferate [27]. Previous studies have indicated a significant reduction in the
351 number of clones detected within the c-Kit$^+$ population (consisting of MP and HSPC) of the BM
352 following a single dose of 5-FU treatment [8]. However, questions arise regarding whether this
353 observed reduction is an artifact stemming from potential under-sampling of the highly expanded c-
354 Kit$^+$ population after 5-FU treatment by single-cell sequencing. Given the quantitative estimation of a
355 wide range of precursor numbers through equation 2, we aimed to investigate whether precursor
356 numbers within progenitor populations indeed reduce ten days post-5-FU treatment (Figure 4A).
357

358 To compare the precursor changes post-5-FU treatment, we used the animals described in Figure 3A
359 as untreated (UT) benchmark, which were collected at the same age. At ten days post 5-FU injection,
360 the efficacy of 5-FU treatment was validated by lower PB myeloid cell frequency post-treatment and
361 higher frequencies of BM progenitors compared to UT animals (Figure S7A and B). The high
362 progenitor frequencies in the BM corresponded to the depletion of cycling cells in the BM and the
363 enhanced proliferation of HSPCs following 5-FU treatment.
364
365 Compared to UT animals, the precursor numbers of MP (including CMP, GMP, MEP) and HSPC
366 significantly decreased, confirming a reduction in clonality in the c-Kit$^+$ population (Figure 4B) [8]. By
367 contrast, the precursor numbers of LT-HSC did not show a decreasing trend. While transplantation
368 studies supported unchanged clonality of primitive stem cells after a single dose of 5-FU, similar
369 investigation in a native environment has not been assessed [28]. Our findings support the notion that
370 native LT-HSC clonality remains unaltered following one-dose 5-FU treatment. Together,
371 myeloablation treatment reinforces how the dynamics of precursor numbers can be tracked through
372 assessing Confetti pattern variations.
373
374 Active lifelong precursors expand poorly in development ontogeny
375 HSCs are thought to undergo substantial expansion in the fetal liver during fetal development [29–31].
376 However, a recent study employing the empirical formula challenged this notion by quantifying
377 endogenous lifelong hematopoietic precursors labeled at various developmental stages [32]. It
378 revealed limited expansion of hematopoietic precursors during the fetal liver stage (from E10 to
379 E15,1.8- to 2.7-fold) as well as a gradual and moderate increase from the fetal liver to the post-natal
380 stage (2.4- to 10-fold) [32]. Although intriguing, most post-natal measurements in this study fell outside
381 the empirical linear range, leaving the genuine degree of post-natal precursor expansion unknown. As
382 our formula quantitatively assesses a wide range of precursor numbers, we set out to directly compare
383 the numbers of precursor labeled at various development stages (E11.5 and E14.5, by one dose of
384 tamoxifen; adult-stage, benchmark shown in in Figure 3A, all analyzed at the same age) (Figure 4C).
385
386 For accuracy, we first examined the dynamics of Confetti labeling. In E14.5-induced animals, the
387 Confetti labeling of T cells almost doubled from one month to four months of age (4.7% at one month,
388 9.3% at four months), suggesting Confetti$^+$-labeled precursors contributed more to post-natal T cell
389 than non-labeled precursors (Figure 4D). Therefore, for T cells in E14.5-induced animals, we focused
390 on precursor numbers generated after three months of age. The Confetti labeling for E11.5-induced
391 animals remained stable for all timepoints examined (Figure S7C).
392

393 Similar to adult-induction, the distribution of average FP% was stable for fetal-induction, supporting
394 equal proliferation of RFP^+ and YFP^+ cells (Figure S7D). Unlike adult-induction, fetal-labeled precursor
395 numbers can be calculated with various combination of FP%, since the expression of CFP introduced
396 at the fetal stages was stable (Figure S6F). Nonetheless, for consistency, all precursor numbers were
397 calculated with variance of $RFP_{RFP+YFP}$%, the one employed in adult-induction animals.
398
399 The resulting precursor numbers estimated for E11.5 and E14.5-labeled cohorts were similar in all PB
400 cell types except at two and three months of age, echoing the previous study reporting limited lifelong
401 precursor expansion in the fetal liver (Figure 4E) [32]. Precursor numbers calculated from BM
402 subpopulations were also comparable between the two timepoints (Figure 4F). Although E9.5-labeled
403 clones were reported to seed non-uniformly across bones, the precursor numbers calculated from PB
404 myeloid cells did not significantly differ from those calculated from BM MP and HSPC in E14.5-labeled
405 animals (Figure 4F) [8].
406
407 For comparison between fetal- and adult-induced animals, we focused on PB myeloid cells, as adult-
408 induced animals had non-saturated labeling in B cells and non-HSPC-rooted Confetti labeling in T
409 cells (Figure 3B). We observed a relatively small increase of precursor numbers in adult-induced
410 animals in PB and BM compared to those labeled at the fetal stage (Figure 4E and F). The increase
411 between E11.5/E14.5- and adult-labeled precursors were less than ten-fold, a level similar to a
412 previous report (Figure 4G) [32].
413
414 Together, we detected minimal to no expansion of lifelong precursors in the fetal liver stage and minor
415 expansion of expansion of lifelong precursors from fetal liver to adult stage.
416
417 *Fancc*$^{-/-}$ mice have normal quantity of hematopoietic precursors at steady-state
418 Our approach provided an opportunity to investigate the number of active HSPC precursor in the native
419 environment of genetic mouse models. To showcase precursor quantification in genetic mouse models,
420 we focused on Fanconi Anemia (FA), the most common inherited bone marrow failure syndrome [33].
421 Current FA mouse models exhibit mostly normal adult hematopoiesis at steady state but demonstrate
422 reduced repopulation ability upon BM transplantation [34–39]. It remains unclear if their steady-state
423 blood production is sustained by a reduced number of precursors, predisposing them to repopulation
424 defects after transplantation.
425
426 To quantify precursor number in a mouse model of FA, we generated *Fancc*$^{+/+}$, *Fancc*$^{+/-}$ and *Fancc*$^{-/-}$
427 mice with a single Confetti allele and homozygous HSC-Scl-CreER^{T} alleles (Fancc$^{+/+}$Confetti$^{fl/+}$HSC-
428 Scl-CreER$^{T/T}$, hereafter *Fancc*$^{+/+}$; Fancc$^{+/-}$Confetti$^{fl/+}$HSC-Scl-CreER$^{T/T}$, hereafter *Fancc*$^{+/-}$; Fancc$^{-/-}$

429 Confetti$^{fl/+}$HSC-Scl-CreER$^{T/T}$, hereafter *Fancc*$^{-/-}$) [40]. Consistent with previous literature, we observed
430 similar PB cell frequencies and blood counts between *Fancc*$^{+/+}$ and *Fancc*$^{-/-}$ mice (Figure S8A-C) [39].
431 BM nucleated cell counts, as well as HSPC and MP frequency, were mostly identical, except for ST-
432 HSC frequencies, which were significantly reduced in *Fancc*$^{-/-}$ mice compared to *Fancc*$^{+/-}$ mice (Figure
433 S8D-E).
434
435 To label hematopoietic precursors in FA animals, Confetti expression was induced at two months of
436 age, and FP% was monitored over seven months (Figure 5A). The absence of *Fancc* did not affect
437 Confetti labeling efficiency, Confetti labeling dynamics or average FP% ($RFP_{RFP+YFP}$%), suggesting
438 precursor numbers in *Fancc*$^{-/-}$ can be similarly calculated (Figure 5B and Figure S8F). As in adult-
439 induction animals, $RFP_{RFP+YFP}$% was used to estimate precursor numbers. The resulting myeloid and
440 BM precursor numbers were comparable regardless of *Fancc* genotype, albeit *Fancc*$^{-/-}$ mice had a
441 slight reduction in HSPC precursor numbers (Figure 5C and D). These observations collectively
442 suggest that a normal number of precursors sustained blood production in *Fancc*$^{-/-}$ mice.
443
444 <u>The number of precursors remained unchanged in mice transplanted with young donor cells</u>
445 While *Fancc*$^{-/-}$ mice have similar number of precursors with their wildtype counterparts at homeostasis,
446 it is unknown whether their reduced repopulation ability post-transplantation stems from diminished
447 precursor numbers, reduced cell expansion, or a combination of both [35]. To understand the
448 underlying mechanism, we performed competitive transplantation using BM cells from three-month-
449 old Confetti-induced *Fancc*$^{+/+}$, *Fancc*$^{+/-}$ or *Fancc*$^{-/-}$ mice along with CD45.1^{+} competitor cells (Figure 5E,
450 Figure S9A). Consistent with previous studies, recipient mice of *Fancc*$^{-/-}$ cells showed significantly
451 lower donor chimerism in the PB and BM compared to those of *Fancc*$^{+/+}$ or *Fancc*$^{+/-}$ cells (Figure 5F,
452 Figure S9B). Despite lower donor chimerism, PB and BM precursor numbers were unaffected in
453 *Fancc*$^{-/-}$ recipient mice (Figure 5G, Figure S9C-E). This suggests that reduced *Fancc*$^{-/-}$ cell proliferation
454 instead of fewer active precursors was likely to be the cause for the reduced repopulation capacity
455 post-transplantation.
456
457 <u>Ageing *Fancc*$^{-/-}$ mice have reduced number of lymphoid hematopoietic precursors upon transplantation</u>
458 Ageing *Fancc*$^{-/-}$ mice have been reported to develop hematologic neoplasms, resulting in decreased
459 survival [41]. To determine if aging *Fancc*$^{-/-}$ cells also maintain similar precursor numbers post-
460 transplantation, we competitively transplanted BM cells from nine-month-old Confetti-induced *Fancc*$^{+/+}$
461 or *Fancc*$^{-/-}$ mice with CD45.1^{+} competitor cells, a stage when aging *Fancc*$^{-/-}$ mice started to show
462 decreased survival [41] (Figure S9F).
463

464 The diminished repopulation ability of *Fancc*$^{-/-}$ cells was reaffirmed by lower PB and BM chimerism
465 (Figure 5H, Figure S9I). While no differences in precursor numbers were noted during the initial post-
466 transplantation period, a slight yet consistent reduction in lymphoid precursors was observed in *Fancc*$^{-}$
467 $^{/-}$ recipient mice at three to five months post-transplantation (Figure 5I, Figure S9G and H). Changes
468 in precursor numbers in the BM of KO recipients were less clear as we observed high variance in
469 several HSPC subtypes (Figure S9J). Nonetheless, precursor numbers of MEP, LSK and ST-HSC
470 showed a consistent reduction. In conclusion, aging *Fancc*$^{-/-}$ mice showed a modest but consistent
471 loss of active PB lymphoid precursors post-transplantation, implying decreased lymphoid precursors
472 additionally contribute to reduced repopulation capacity as *Fancc*$^{-/-}$ mice age.
473
474 Discussion
475 The polyclonal nature of endogenous hematopoiesis imposes methodological problems on a robust
476 dynamic measurement. Inspired by the XCI studies, we successfully employed the correlation formula
477 informed by binominal distribution to quantify precursor number in native hematopoiesis. Using HSC-
478 Scl-CreER^{T}/Confetti animals, we estimate thousands of precursors contributing to adult native
479 hematopoiesis, a number comparable to the previous report [1]. This number resulted from a moderate
480 increase during fetal-to-adult transition and responded dynamically to myeloablation by 5-FU. Applying
481 to a mouse model of inherited bone marrow failure (Fancc$^{-/-}$ mice), we detected normal precursor
482 numbers at steady state and decreased lymphoid precursors upon transplantation of aging donors. In
483 sum, the study underscores the flexibility of the approach and the value in analyzing precursor
484 contributions *in situ*.
485
486 Although the linear relationship between variance of FP% and precursor numbers has been described,
487 its measurable range was limited [10]. Estimates outside this linear range may erroneously fall within
488 it, making it challenging to distinguish accurate measurements within the range from inaccurate ones
489 beyond it [10]. Based on binomial distribution, we expanded the measurable range to encompass the
490 possible dynamic range of hematopoietic precursors in mice, minimizing the likelihood of inaccuracies.
491 Futhermore, the linear correlation based on binomial distribution can accommodate any labeling
492 system that follows the underlying premises, alleviating the need for the intricate Confetti cassette.
493
494 Having established a quantitative measurement, we were initially surprised to confirm the moderate
495 number differences between fetal and adult-precursors seen in other report [32]. The maximum
496 increase observed (seven-fold observed in CMP) was substantially lower than the increase of
497 competitive repopulation units (CRU) from E12 to E16 determined by limited dilution assays (~38-fold)
498 (Figure 4G), as well as the escalation from immunophenotype-defined HSPC counts (28- to 47-fold,
499 from 3-5x10^3 at E14.5 to 1.4x10^5 at two months old) (Figure 4G) [17,30,42,43]. It is possible that we

500 overestimated E11.5 precursors, as the numbers of lifelong E11.5 precursors were substantially higher
501 (on average 753 PB myeloid precursor, 512 HSPC precursors) than the one to two repopulation units
502 estimated with transplantation [44]. Nonetheless, our results were comparable to the ~870
503 hematopoietic cluster cells observed at E11.5 [45]. On the other hand, we may have underestimated
504 precursor numbers in adult-labeled animals, as most adult HSCs are quiescent [27]. However, HSC
505 precursor numbers did not increase after induced proliferation by one dose 5-FU treatment (Figure
506 4B). In sum, our results suggest that the moderate precursor numbers difference between fetal- and
507 adult-stage are likely genuine, emphasizing the analysis of hematopoiesis in a native environment.
508 Future studies using different methods to investigate precursor activity locally is required to validate
509 this result.
510
511 Novel methods tracing clones *in situ* offer new opportunities to study native hematopoiesis, yet most
512 are challenging to apply to genetic mouse models [1,8,9,46,47]. Since only two mouse lines are
513 required (sometimes Cre is already used for conditional deletion of alleles), our approach is more
514 convenient to study hematopoiesis in mouse models of genetic disorders. Applying precursor
515 measurements to *Fancc*$^{-/-}$ mice, we observed a minor decrease in precursor numbers after
516 transplantation of aging donors. Although donor chimerism differences had been linked to difference
517 in precursor numbers, reduced proliferation capacity may also contribute to reduced competitive
518 repopulation capacity [25]. In this case, precursor number analysis is necessary to differentiate these
519 possibilities.
520
521 Currently, concerns regarding clonal restriction in the context of FA gene therapy arises, as the
522 engraftment of gene-corrected stem cells has yielded marginal clinical benefits [48]. For the first time,
523 we demonstrated that FA cell maintaining a normal precursor quantity post-transplantation, thereby
524 disproving clonal attrition, including putative homing deficits post-transplantation as a cause [49].
525 However, it is imperative to acknowledge that the majority of murine models of FA, including the *Fancc*$^{-}$
526 $^{/-}$ model utilized herein, do not manifest spontaneous bone marrow failure, raising uncertainty
527 regarding the extent to which FA mouse models recapitulate the pathophysiology observed in FA
528 patients. Future investigations in other FA mice and studies leveraging FA patient-derived materials
529 will be pivotal in corroborating and validating the findings presented here.
530
531 While we implemented careful data processing procedure, one pitfall of our analyses was that we
532 inferred precursor numbers from their progenies, assuming uniform and linear expansion from
533 precursor to progenies. A recent study has showed non-uniform precursor clone sizes, although the
534 level of non-uniformity is low [47]. In cases where the non-uniformity is high, according to mathematical
535 modeling, we measured the major contributors to hematopoiesis[50]. Another potential caveat is that

536 the relative contribution of Confetti-labeled precursors to blood production compared to non-labeled
537 precursors remains unknown. Future studies using different Cre drivers will help validate the precursor
538 numbers for steady-state hematopoiesis.
539
540 In summary, we established a wide applicable range of the correlation between variance of FP% and
541 the precursor based on binomial distribution. We discovered thousands of precursors contributing to
542 steady-state adult hematopoiesis and validated that fetal-to-adult precursor expansion were limited.
543 This analysis highlights active precursor number as an important metric in both normal and genetic
544 mouse models.
545
546 Methods
547 Variance modeling with two-color HL-60
548 HL-60 cells were cultured with IMDM containing 20% FBS (Gemini) and 5% Penicillin-Streptomycin
549 (Gibco) and maintained at 1 x 10^5 and 1 x 10^6 cells/ml. Mycoplasma tests (Lonza) were performed
550 routinely to ensure no mycoplasma contamination. BFP-HL-60 and GFP-HL-60 were generated with
551 lentivirus transduction of pGK-BFP (Genscript) and LeGO-V2 (a gift from Dr. Stefano Rivella lab).
552 For seeding of one to 10,000 cells, sorting of HL-60 mixture was performed using a BD FACS Aria
553 III. For seeding of 100,000 cells, cell counting and dilution was used. After expansion, HL-60 was
554 fixed with BD fixation buffer before proceeding to Aurora (Cytek) or Cytoflex (Beckman Coulter) for
555 analysis.
556
557 Mice
558 HSC-Scl-CreER^{T} mice **[22]** were crossed with R26R-Confetti mice (B6.129P2-*Gt(ROSA)26Sor*$^{tm1(CAG-}$
559 $^{Brainbow2.1)Cle}$/J) to generate HSC-Scl-CreER$^{T/T}$Confetti$^{fl/+}$ (HSC-Scl-CreER^{T}/Confetti) animals. HSC-Scl-
560 CreER^{T}/Confetti animals were crossed with *Fancc*$^{+/-}$ mice **[40]** to generate HSC-Scl-
561 CreER$^{T/T}$Confetti$^{fl/+}$*Fancc*$^{+/+}$ (*Fancc*$^{+/+}$), HSC-Scl-CreER$^{T/T}$Confetti$^{fl/+}$*Fancc*$^{+/-}$ (*Fancc*$^{+/-}$), and HSC-Scl-
562 CreER$^{T/T}$Confetti$^{fl/+}$Fancc$^{-/-}$ (*Fancc*$^{-/-}$) mice. Vav-Cre was generously offered by Dr. Wei Tong
563 (Children's Hospital of Philadelphia). Vav-Cre/Confetti animals were generated by crossing Vav-Cre
564 with R26R-Confetti mice. All animals are B6 CD45.2 background unless otherwise stated. Six- to
565 twelve-week-old females were used for timed pregnancies. Eight-week-old mice were used for adult
566 Confetti induction. Both female and male mice were used indistinctly **[10]**. Six- to eight-week-old old
567 female B6 CD45.1 mice (B6.SJL-*Ptprc*^{a} *Pepc*^{b}/BoyJ, Jackson laboratories) were used as competitors
568 and recipients for transplantation studies. All mice were maintained in the conventional small animal
569 facility at the Children's Hospital of Philadelphia (CHOP). All procedures involving animals were
570 approved by the Institutional Animal Care and Use Committee at the Children's Hospital of
571 Philadelphia.

572

573 Animal identification

574 Tail snip DNA was extracted using KAPA Express Extract Kit (Roche). Genotyping PCR was
575 performed with HotStarTaq Master Mix (Qiagen) according to manufacturer's instruction. Genotyping
576 primers used are summarized in Table S3. To determine the zygosity of HSC-Scl-CreER^{T}, qPCR was
577 additionally performed with purified tail snip DNA using SYBR™ Green Universal Master Mix (Applied
578 Biosystems).

579

580 Animal procedures

581 For fetal inductions, timed matings of HSC-Scl-CreER$^{T/T}$Confetti$^{fl/fl}$ mice and HSC-Scl-CreER$^{T/T}$ mice
582 were set up. The mice were separated the next morning and the noon of the day of separation was
583 considered E0.5. Tamoxifen was delivered at 100mg/kg to the dam orally at E11.5 or E14.5. Pups
584 were C-sectioned and cross-fostered at E18.5 due to reported delivery difficulties caused by tamoxifen
585 treatment **[51]**. For mice used for defined number of transplantation (Figure 2), tamoxifen was
586 delivered at 70mg/kg orally once per day for 14 days at eight-week-old. For adult induction (Figure 3),
587 tamoxifen was delivered at 70mg/kg orally once per day for two days at eight-week-old. For one dose
588 5-fluorouracil (5-FU, Sigma) treatment, 5-FU was intraperitoneally injected once ten days before BM
589 harvest (37-week-old) at 150 mg/kg. To obtain peripheral blood for Confetti analysis, mice were
590 anesthetized using isoflurane and retro-orbitally bled or submandibularly bled for *Fancc*$^{-/-}$ mice with
591 occasional congenital eye defects**[52]** for 1 capillary of blood (50ul). For peripheral blood counts, blood
592 was collected in EDTA tubes using similar bleeding methods and was analyzed by the Translational
593 Core Lab (Children's Hospital of Philadelphia).

594

595 Bone marrow sample processing

596 To collect BM cells, mice were euthanized by CO_2 inhalation. Tibia, femur and pelvis were dissected,
597 and the bone marrow cells were flushed with 26-gauge needles. The single-cell suspension generated
598 with 18-gauge needles was then filtered through 70μm strainers. Erythrocytes in the bone marrow
599 cells were hemolyzed by RBC lysis buffer before antibody staining or transplantation. For stem cell
600 enrichment, bone marrow cells were further lineage depleted using EasySep™ Mouse Hematopoietic
601 Progenitor Cell Isolation Kit (STEMCELL Technologies).

602

603 Flow cytometry analysis

604 The peripheral blood and bone marrow cells were analyzed using Aurora (Cytek) or BD FACS Aria III
605 and the flow cytometry data was analyzed using FlowJo (Tree Star). The combinations of the following
606 cell surface markers were used to define the peripheral blood populations: myeloid cells: CD11b^{+} or
607 Gr-1^{+}; T-cell: CD3ε^{+}; B-cell: B220^{+}. The following combinations of cell surface markers were used to

define the bone marrow stem and progenitor cells (Lineage/Lin: CD11b, Gr-1, B220, CD3ε, Ter119): LTHSC: $Lin^-c\text{-}Kit^+Sca1^+Flk2^-CD150^+CD48^-$; MPP2: $Lin^-c\text{-}Kit^+Sca1^+Flk2^-CD150^+CD48^{+;}$ MPP3: $Lin^-c\text{-}Kit^+Sca1^+Flk2^-CD150^-CD48^+$; MPP4: $Lin^-c\text{-}Kit^+Sca1^+Flk2^+CD150^-CD48^+$; STHSC: $Lin^-c\text{-}Kit^+Sca1^+Flk2^-CD150^-CD48^-$; MEP: $Lin^-c\text{-}Kit^+Sca1^-CD34^-CD16/32^-$; CMP: $Lin^-c\text{-}Kit^+Sca1^-CD34^{mid}CD16/32^{mid}$; GMP: $Lin^-c\text{-}Kit^+Sca1^-CD34^+CD16/32^+$. For bone marrow stem and progenitor cell analysis, DAPI (Biolegend) was used to distinguish dead cells. Representative examples of flow cytometry gating can be found in Figure S2C. The antibodies were used at optimized dilutions listed in Table S1.

Bone marrow transplantation

The day before transplantation, female CD45.1 recipient mice were lethally irradiated (5.2Gy x 2, 3h apart) using an X-ray irradiator (Precision). On the day of transplantation, bone marrow cells from donor mice ($CD45.2^+$) were collected under sterile conditions as described, RBC-lysed and counted for cell number. For Figure 2 (transplantation of defined number of precursors), donor mice were induced to express Confetti as described; only donor bone marrow cells were injected into irradiated recipient mice; each matched high- and low- precursor number group received donor bone marrow cells pooled from three to five mice. For transplantation of young Fancc mice, donor mice were induced to express Confetti at E14.5; 1.5 x 10^6 donor bone marrow cells were mixed with 2.5 x 10^5 CD45.1 supporting bone marrow cells and injected into irradiated recipient mice via tail vein. For transplantation of aging Fancc mice, donor mice were induced to express Confetti at two-month-old; 2 x 10^6 donor bone marrow cells were mixed with 5 x 10^5 CD45.1 supporting bone marrow cells as donor population.

Derivation of the correlation between variance of confetti fp% and precursor numbers based on binomial distribution

When a random variable adheres to binomial distribution, studies have established that the following equation holds true:

$$n = \frac{p(1-p)}{\sigma^2}$$

where *n* signifies the number of precursors, *p* represents the probability of an individual being one of the FPs (e.g., RFP), and σ^2 denotes the variance of specific FP% (e.g. $\sigma_{RFP\%}{}^2$)[6]. In experiments, *p* will be estimated with the average FP% in the $Confetti^+$ cells (e.g. $RFP_{mean}\%$), and σ^2 will be estimated with the variance of FP% among a group of individual or mice (e.g. $\widehat{\sigma_{RFP}}^2$). Consequently, the estimation of precursors number n is calculated using the following equation:

$$\hat{n} = \frac{\widehat{FP_{mean}}\% \times (1 - \widehat{FP_{mean}}\%)}{\widehat{\sigma_{FP}}^2}$$

A logarithmic transformation can then be performed, allowing us to establish a linear relationship between the variance of FP% and the number of precursors:

$$\lg(\hat{n}) = \lg\left[\widehat{FP_{mean}}\% \times (1 - \widehat{FP_{mean}}\%)\right] - 2 \times \lg(\widehat{\sigma_{FP}}) \quad \text{(equation 2)}$$

In a previous study[10],

$$\sigma_{FP} = CV_{FP} \times FP_{mean}\%$$

Therefore,

$$\lg(\hat{n}) = \lg\left[\frac{(1-\widehat{FP_{mean}}\%)}{\widehat{FP_{mean}}\%}\right] - 2 \times \lg\left(\widehat{CV_{FP}}\right) \text{ (equation 1)}$$

<u>Resampling to determine sample size per replicate</u>

FP% data generated from HL-60 was used for re-sampling. For each seeded number, FP% was resampled for different sample sizes from all the FP% with replacement. Variance of estimated n was then calculated by the standard deviation of the estimated precursor numbers generated from the same sample size. Relative error was determined by dividing variance of estimated n with the average of the estimated n. Refer to “varying well numbers.rmd” for detailed R code.

<u>Simulation to determine the effect of varying FP_{mean}% on correlation between variance of FP% and precursor numbers</u>

Simulation of binomial distribution of varying FP_{mean}% (probability of being a FP) was performed in R, generating corresponding FP% values used for variance calculation. The correlation between variance and precursor number n was then determined by linear correlation, and the slopes and intercepts were compared to each other. Refer to “Simulation of varying FPmean_percent.rmd” for detailed R code.

<u>Simulation to determine the effect of precursor numbers following Poisson post-transplantation</u>

Precursor numbers in individual samples were simulated to follow a Poisson distribution, where the mean of precursor numbers is the expected precursor numbers (expected n). The induction of a FP in each precursor was then simulated by random assignment, where probability of being a FP was set to be FP_{mean}%. For each expected precursor numbers, variance of FP% among samples was then calculated, and equation 2 was used to estimate precursor numbers (estimated n). The correlation between expected n and estimated n was then plotted. Refer to “Double layer binomial simulation.rmd” for detailed R code.

<u>Data processing and normalization</u>

676 Peripheral blood cell subset frequencies were normalized to the total % of granulocytes and myeloid
677 cells, T-cell and B-cell, to avoid an underestimate due to incomplete RBC lysis. For transplant animals,
678 peripheral blood CD45.1^{+}% and CD45.2^{+}% are normalized to total CD45.1^{+}% and CD45.2^{+}% to avoid
679 an underestimate due to incomplete RBC lysis. After normalization of cell frequencies, the sum of
680 Confetti% and FP_{Conf}% (diving FP_{PB}% or FP_{BM}% by sum of Confetti%) were then calculated for each
681 cohort. At each step, potential outliers were removed based on Tukey method. The variance of FP_{Conf}%
682 and average FP_{Conf}% are then fitted into equation 2 to calculate precursor numbers if sample size is
683 at least five. The precursor number is then used to compared with the minimum flow cytometry
684 recorded event of samples. If the estimated number is higher than the minimum flow count of samples,
685 the sample with the minimum flow count will be excluded, and the calculation is performed again on
686 the rest of samples. After exclusion, the resulting precursor number will be compared again with
687 minimum flow count of samples, until it is lower than the minimum flow count or sample size is smaller
688 than five.
689
690 Statistical analysis
691 Statistical significance for precursor number estimates
692 As the distribution of precursor numbers is not predetermined, to compare mean precursor number
693 differences between two conditions from a limited number of biological replicates, we employed
694 permutation test. For unpaired permutation test, if there are three to five biological replicates per
695 condition, the smallest p values possible are 0.05 ~ 0.004 ($1/\binom{6}{3}$ ~ $1/\binom{5}{10}$). For paired permutation test,
696 if there are three to five biological replicates per condition, the smallest p value possible are 0.125 ~
697 0.03125 ($1/2^3$ ~ $1/2^5$). Thus, for three to five biological replicates per condition, even though some of
698 the p values may not reach the commonly used alpha level (0.05), it still represents substantial number
699 differences. For those p values that were lowest possible but did not reach the alpha level, we
700 specifically labeled with a "#" in the figures and legends.
701 Statistical analyses of other data
702 All other two-sample statistical analyses were performed using Student's t test, if the sample was
703 normally distributed, or Welch's t test, when the sample was not normally distributed (F test). For
704 multiple comparison, one-way ANOVA or two-way ANOVA was used.
705
706 Data and code availability
707 The original Confetti% measured in the peripheral blood and the bone marrow and R code to analyze
708 and reproduce all the results, numeric and figures can be found at
709 https://doi.org/10.5281/zenodo.8222789. The re-analyzed data from a previous study can be found in
710 the online version of papers[10].

References

1. Sun J, Ramos A, Chapman B, Johnnidis JB, Le L, Ho Y-J, et al. Clonal dynamics of native haematopoiesis. Nature. 2014;514: 322–327. doi:10.1038/nature13824

2. Lee-Six H, Øbro NF, Shepherd MS, Grossmann S, Dawson K, Belmonte M, et al. Population dynamics of normal human blood inferred from somatic mutations. Nature. 2018;561: 473–478. doi:10.1038/s41586-018-0497-0

3. Mitchell E, Chapman MS, Williams N, Dawson KJ, Mende N, Calderbank EF, et al. Clonal dynamics of haematopoiesis across the human lifespan. Nature. 2022;606: 343–350. doi:10.1038/s41586-022-04786-y

4. Breivik H. Haematopoietic stem cell content of murine bone marrow, spleen, and blood. Limiting dilution analysis of diffusion chamber cultures. J Cell Physiol. 1971;78: 73–78. doi:10.1002/jcp.1040780111

5. Campbell P, Chapman MS, Wilk M, Boettcher S, Mitchell E, Dawson K, et al. Clonal dynamics after allogeneic haematopoietic cell transplantation using genome-wide somatic mutations. 2023. doi:10.21203/rs.3.rs-2868644/v1

6. WALLIS VJ, LEUCHARS E, CHWALINSKI S, DAVIES AJS. ON THE SPARSE SEEDING OF BONE MARROW AND THYMUS IN RADIATION CHIMAERAS. Transplantation. 1975;19: 2–11. doi:10.1097/00007890-197501000-00002

7. Szilvassy SJ, Humphries RK, Lansdorp PM, Eaves AC, Eaves CJ. Quantitative assay for totipotent reconstituting hematopoietic stem cells by a competitive repopulation strategy. Proc National Acad Sci. 1990;87: 8736–8740. doi:10.1073/pnas.87.22.8736

8. Bowling S, Sritharan D, Osorio FG, Nguyen M, Cheung P, Rodriguez-Fraticelli A, et al. An Engineered CRISPR-Cas9 Mouse Line for Simultaneous Readout of Lineage Histories and Gene Expression Profiles in Single Cells. Cell. 2020;181: 1410-1422.e27. doi:10.1016/j.cell.2020.04.048

9. Pei W, Feyerabend TB, Rössler J, Wang X, Postrach D, Busch K, et al. Polylox barcoding reveals haematopoietic stem cell fates realized in vivo. Nature. 2017;548: 456–460. doi:10.1038/nature23653

10. Ganuza M, Hall T, Finkelstein D, Chabot A, Kang G, McKinney-Freeman S. Lifelong haematopoiesis is established by hundreds of precursors throughout mammalian ontogeny. Nat Cell Biol. 2017;19: 1153–1163. doi:10.1038/ncb3607

11. Buescher ES, Alling DW, Gallin JI. Use of an X-linked human neutrophil marker to estimate timing of lyonization and size of the dividing stem cell pool. J Clin Investig. 1985;76: 1581–1584. doi:10.1172/jci112140

12. Berger CN, Sturm KS. Estimation of the Number of Hematopoietic Precursor Cells During Fetal Mouse Development by Covariance Analysis. Blood. 1996;88: 2502–2509. doi:10.1182/blood.v88.7.2502.bloodjournal8872502

749 13. Busque L, Patel JP, Figueroa M, Vasanthakumar A, Provost S, Hamilou Z, et al. Recurrent
750 Somatic TET2 Mutations in Normal Elderly Individuals With Clonal Hematopoiesis. Nat Genet.
751 2012;44: 1179–1181. doi:10.1038/ng.2413

752 14. Gandini E, Gartler SM, Angioni G, Argiolas N, Dell'Acqua G. Developmental implications of
753 multiple tissue studies in glucose-6-phosphate dehydrogenase-deficient heterozygotes. Proc Natl
754 Acad Sci United States Am. 1968;61: 945–8. doi:10.1073/pnas.61.3.945

755 15. Micklem HS, Lennon JE, Ansell JD, Gray RA. Numbers and dispersion of repopulating
756 hematopoietic cell clones in radiation chimeras as functions of injected cell dose. Exp Hematol.
757 1987;15: 251–7.

758 16. Harrison DE, Lerner C, Hoppe PC, Carlson GA, Alling D. Large Numbers of Primitive Stem Cells
759 Are Active Simultaneously in Aggregated Embryo Chimeric Mice. Blood. 1987;69: 773–777.
760 doi:10.1182/blood.v69.3.773.773

761 17. Cosgrove J, Hustin LSP, Boer RJ de, Perié L. Hematopoiesis in numbers. Trends Immunol.
762 2021;42: 1100–1112. doi:10.1016/j.it.2021.10.006

763 18. Fanti A-K, Busch K, Greco A, Wang X, Cirovic B, Shang F, et al. Flt3- and Tie2-Cre tracing
764 identifies regeneration in sepsis from multipotent progenitors but not hematopoietic stem cells. Cell
765 Stem Cell. 2023;30: 207-218.e7. doi:10.1016/j.stem.2022.12.014

766 19. Patel SH, Christodoulou C, Weinreb C, Yu Q, Rocha EL da, Pepe-Mooney BJ, et al. Lifelong
767 multilineage contribution by embryonic-born blood progenitors. Nature. 2022; 1–7.
768 doi:10.1038/s41586-022-04804-z

769 20. Schoedel KB, Morcos MNF, Zerjatke T, Roeder I, Grinenko T, Voehringer D, et al. The bulk of
770 the hematopoietic stem cell population is dispensable for murine steady-state and stress
771 hematopoiesis. Blood. 2016;128: 2285–2296. doi:10.1182/blood-2016-03-706010

772 21. Solomon M, Song B, Govindarajah V, Good S, Arasu A, Hinton EB, et al. Slow cycling and
773 durable Flt3+ progenitors contribute to hematopoiesis under native conditions. J Exp Med. 2023;221:
774 e20231035. doi:10.1084/jem.20231035

775 22. Göthert JR, Gustin SE, Hall MA, Green AR, Göttgens B, Izon DJ, et al. In vivo fate-tracing
776 studies using the Scl stem cell enhancer: embryonic hematopoietic stem cells significantly contribute
777 to adult hematopoiesis. Blood. 2005;105: 2724–2732. doi:10.1182/blood-2004-08-3037

778 23. Harrison DE, Astle CM, Lerner C. Number and continuous proliferative pattern of transplanted
779 primitive immunohematopoietic stem cells. Proc Natl Acad Sci. 1988;85: 822–826.
780 doi:10.1073/pnas.85.3.822

781 24. Harrison DE, Stone M, Astle CM. Effects of transplantation on the primitive
782 immunohematopoietic stem cell. J Exp Med. 1990;172: 431–437. doi:10.1084/jem.172.2.431

783 25. Harrison DE, Jordan CT, Zhong RK, Astle CM. Primitive hemopoietic stem cells: direct assay of
784 most productive populations by competitive repopulation with simple binomial, correlation and
785 covariance calculations. Exp Hematol. 1993;21: 206–19.

786 26. Busch K, Klapproth K, Barile M, Flossdorf M, Holland-Letz T, Schlenner SM, et al. Fundamental
787 properties of unperturbed haematopoiesis from stem cells in vivo. Nature. 2015;518: 542–546.
788 doi:10.1038/nature14242

789 27. Wilson A, Laurenti E, Oser G, Wath RC van der, Blanco-Bose W, Jaworski M, et al.
790 Hematopoietic Stem Cells Reversibly Switch from Dormancy to Self-Renewal during Homeostasis
791 and Repair. Cell. 2008;135: 1118–1129. doi:10.1016/j.cell.2008.10.048

792 28. Lerner C, Harrison DE. 5-Fluorouracil spares hemopoietic stem cells responsible for long-term
793 repopulation. Exp Hematol. 1990;18: 114–8.

794 29. Morrison SJ, Hemmati HD, Wandycz AM, Weissman IL. The purification and characterization of
795 fetal liver hematopoietic stem cells. Proc National Acad Sci. 1995;92: 10302–10306.
796 doi:10.1073/pnas.92.22.10302

797 30. Ema H, Nakauchi H. Expansion of hematopoietic stem cells in the developing liver of a mouse
798 embryo. Blood. 2000;95: 2284–2288. doi:10.1182/blood.v95.7.2284

799 31. Rybtsov S, Ivanovs A, Zhao S, Medvinsky A. Concealed expansion of immature precursors
800 underpins acute burst of adult HSC activity in foetal liver. Dev (Camb, Engl). 2016;143: 1284–1289.
801 doi:10.1242/dev.131193

802 32. Ganuza M, Hall T, Myers J, Nevitt C, Sánchez-Lanzas R, Chabot A, et al. Murine foetal liver
803 supports limited detectable expansion of life-long haematopoietic progenitors. Nat Cell Biol. 2022;24:
804 1475–1486. doi:10.1038/s41556-022-00999-5

805 33. Ceccaldi R, Parmar K, Mouly E, Delord M, Kim JM, Regairaz M, et al. Bone Marrow Failure in
806 Fanconi Anemia Is Triggered by an Exacerbated p53/p21 DNA Damage Response that Impairs
807 Hematopoietic Stem and Progenitor Cells. Cell Stem Cell. 2012;11: 36–49.
808 doi:10.1016/j.stem.2012.05.013

809 34. Carreau M, Gan OI, Liu L, Doedens M, Dick JE, Buchwald M. Hematopoietic compartment of
810 Fanconi anemia group C null mice contains fewer lineage-negative CD34+ primitive hematopoietic
811 cells and shows reduced reconstitution ability. Exp Hematol. 1999;27: 1667–1674.
812 doi:10.1016/s0301-472x(99)00102-2

813 35. Haneline LS, Gobbett TA, Ramani R, Carreau M, Buchwald M, Yoder MC, et al. Loss of FancC
814 Function Results in Decreased Hematopoietic Stem Cell Repopulating Ability. Blood. 1999;94: 1–8.
815 doi:10.1182/blood.v94.1.1.413k03_1_8

816 36. Zhang Q-S, Marquez-Loza L, Eaton L, Duncan AW, Goldman DC, Anur P, et al. Fancd2−/− mice
817 have hematopoietic defects that can be partially corrected by resveratrol. Blood. 2010;116: 5140–
818 5148. doi:10.1182/blood-2010-04-278226

819 37. Du W, Amarachintha S, Erden O, Wilson A, Meetei AR, Andreassen PR, et al. Fancb deficiency
820 impairs hematopoietic stem cell function. Sci Rep-uk. 2015;5: 18127. doi:10.1038/srep18127

821 38. Dubois EL, Guitton-Sert L, Béliveau M, Parmar K, Chagraoui J, Vignard J, et al. A Fanci
822 knockout mouse model reveals common and distinct functions for FANCI and FANCD2. Nucleic
823 Acids Res. 2019;47: 7532–7547. doi:10.1093/nar/gkz514

824 39. Zha J, Kunselman L, Fan J-M, Olson TS. Bone marrow niches of germline FANCC/FANCG
825 deficient mice enable efficient and durable engraftment of hematopoietic stem cells after
826 transplantation. Haematologica. 2019;104: e284–e287. doi:10.3324/haematol.2018.202143

827 40. Chen M, Tomkins DJ, Auerbach W, McKerlie C, Youssoufian H, Liu L, et al. Inactivation of Fac in
828 mice produces inducible chromosomal instability and reduced fertility reminiscent of Fanconi
829 anaemia. Nat Genet. 1996;12: 448–451. doi:10.1038/ng0496-448

830 41. Cerabona D, Sun Z, Nalepa G. Leukemia and chromosomal instability in aged Fancc−/− mice.
831 Exp Hematol. 2016;44: 352–357. doi:10.1016/j.exphem.2016.01.009

832 42. Young I-C, Wu B, Andricovich J, Chuang S-T, Li R, Tzatsos A, et al. Differentiation of fetal
833 hematopoietic stem cells requires ARID4B to restrict autocrine KITLG/KIT-Src signaling. Cell Rep.
834 2021;37: 110036–110036. doi:10.1016/j.celrep.2021.110036

835 43. Mochizuki-Kashio M, Yoon Y me, Menna T, Grompe M, Kurre P. FANCD2 Alleviates Physiologic
836 Replication Stress in Fetal Liver HSC. Biorxiv. 2020; 2020.09.30.320796.
837 doi:10.1101/2020.09.30.320796

838 44. Kumaravelu P, Hook L, Morrison AM, Ure J, Zhao S, Zuyev S, et al. Quantitative developmental
839 anatomy of definitive haematopoietic stem cells/long-term repopulating units (HSC/RUs): role of the
840 aorta-gonad-mesonephros (AGM) region and the yolk sac in colonisation of the mouse embryonic
841 liver. Development. 2002;129: 4891–4899. doi:10.1242/dev.129.21.4891

842 45. Yokomizo T, Ideue T, Morino-Koga S, Tham CY, Sato T, Takeda N, et al. Independent origins of
843 fetal liver haematopoietic stem and progenitor cells. Nature. 2022;609: 779–784.
844 doi:10.1038/s41586-022-05203-0

845 46. Rodriguez-Fraticelli AE, Wolock SL, Weinreb CS, Panero R, Patel SH, Jankovic M, et al. Clonal
846 analysis of lineage fate in native haematopoiesis. Nature. 2018;553: 212–216.
847 doi:10.1038/nature25168

848 47. Li L, Bowling S, McGeary SE, Yu Q, Lemke B, Alcedo K, et al. A mouse model with high clonal
849 barcode diversity for joint lineage, transcriptomic, and epigenomic profiling in single cells. Cell.
850 2023;186: 5183-5199.e22. doi:10.1016/j.cell.2023.09.019

851 48. Río P, Navarro S, Wang W, Sánchez-Domínguez R, Pujol RM, Segovia JC, et al. Successful
852 engraftment of gene-corrected hematopoietic stem cells in non-conditioned patients with Fanconi
853 anemia. Nat Med. 2019;25: 1396–1401. doi:10.1038/s41591-019-0550-z

854 49. Zhang X, Shang X, Guo F, Murphy K, Kirby M, Kelly P, et al. Defective homing is associated with
855 altered Cdc42 activity in cells from patients with Fanconi anemia group A. Blood. 2008;112: 1683–
856 1686. doi:10.1182/blood-2008-03-147090

857 50. Stone M. A general statistical model for clone--tissue studies, using X-chromosome inactivation
858 data. Biometrics. 1983;39: 395–409.

859 51. Lizen B, Claus M, Jeannotte L, Rijli FM, Gofflot F. Perinatal induction of Cre recombination with
860 tamoxifen. Transgenic Res. 2015;24: 1065–1077. doi:10.1007/s11248-015-9905-5

861 52. Joly S, Rodriguez L, Mdzomba JB, Carreau M, Pernet V. The ocular manifestations of Fanconi
862 anemia in a genetic mouse model. Investigative Ophthalmology & Visual Science. 60: 4738–4738.

863

864 Acknowledgments

865 Work was supported by R01-HL150882. We thank Florin Tuluc (CHOP Flow Cytometry Core) and
866 Jessica Gucwa (Cytek bioscience) for assistance in flow cytometry; Kaosheng Lv (CHOP) for
867 preparation of reagents; Zilu Zhou (Penn) for help in statistical analysis; Nancy Speck (Penn), Wei
868 Tong (Penn), Julia Warren (Penn), Ding-wen (Roger) Chen (CHOP), Stephanie N Hurwitz (Penn), Hua
869 Qing (Genentech), Hui Chen (Penn) for in-depth discussion. Some figure panels were generated using
870 Biorender.

871

872 Author information

873 Authors and Affiliations

874 Comprehensive Bone Marrow Failure Center, Children's Hospital of Philadelphia, Philadelphia, PA,
875 USA

876 Suying Liu, Sarah Adams, Juliana Ehnot,Seul K. Jung, Greer Jeffery, Theresa Menna, Peter Kurre

877

878 Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

879 Suying Liu, Sarah Adams, Juliana Ehnot,Seul K. Jung, Greer Jeffery, Theresa Menna, Peter Kurre

880

881 Department of Biostatistics, Epidemiology and informatics, University of Pennsylvania, Philadelphia,
882 PA, USA

883 Haotian Zheng, Hongzhe Lee

884

885 Stem Cell Regulation Unit, St. Vincent's Institute of Medical Research, Fitzroy, VIC 3065, Australia

886 Louise E. Purton

887

888 Department of Medicine, The University of Melbourne, Parkville, VIC 3052, Australia

889 Louise E. Purton

890

891 Contributions

892 S.L designed, performed experiments, analyzed data and wrote the manuscript. S.A, J.E, S.J, G.J
893 and T.M performed experiments. H.Z and H.L designed statistical analysis. L.P contributed reagents,
894 guidance and edited the manuscript. P.K conceived and supervised the study and edited the
895 manuscript.

896

Correspondence authors
Correspondence to Peter Kurre kurrep@chop.edu

Figure Legends

Figure 1. The Principle of Inverse Correlation between Variance of FP% and precursor numbers.

(A) Schematic of the Confetti cassette at the mouse *Rosa26* locus. Sequences of four fluorescence protein are interspaced by loxP sites. Upon Cre-expression, the Confetti cassette will recombine and express one of the four fluorescence proteins.

(B) The inverse relationship between variance of markers and precursor numbers. The distribution of FP% is determined in the progeny to estimate the number of precursors.

(C) Workflow to validate the correlation formula between variance of FP% and precursor numbers by a two-color cell system.

(D) BFP frequency in wells seeded with 1 to 100,000 HL-60. Each dot represents a well. Two replicates were shown. Each replicate consists of at least 24 wells per seeded number. Error bars represent mean ± SD.

(E) The standard deviation of BFP% in wells seeded with various numbers of HL-60. Error bars represent mean ± SD. N = 2 replicates.

(F) The correlation between seeded numbers and standard deviation of BFP%. N = 2 replicates.

(G) The correlation between seeded numbers and the numbers estimated with standard deviation of BFP% and equation 2. N = 2 replicates.

For (A-C), images were created with Biorender.

Figure 2. Transplantation of Defined Number of Precursors

(A) Representative flow plots showing Confetti induction by HSC-Scl-CreER^{T} in the LSK population.

(B) Schematic for transplantation of defined number of precursors. Briefly, $4x10^6$ Confetti-induced WBM for “high precursor number” groups, or $0.25x10^6$ WBM for “low precursor number” groups were transplanted into lethally irradiated recipient mice. Image was created with Biorender.

(C) RFP_{Conf}% distribution in PB myeloid cells. Each dot represents one animal. Recipient sample size = 7-14 mice per replicate, N = 3 replicates per group. Error bars represent mean ± SD.

(D) The precursor numbers of B cells, T cells, and myeloid cells in recipient mice. Dash lines mark the level of historically estimated transplantable clone numbers. Each dot represents a precursor number calculated from multiple mice. Error bars represent mean ± SD.

(E) The correlation of precursor numbers calculated from RFP_{Conf}%, YFP_{Conf}% or CFP_{Conf}%. Each dot represents a precursor number calculated from multiple mice.

931 (F) The precursor number of BM MP and HSPC in recipient mice. Dash lines represent historically
932 estimated transplantable clone numbers. Each dot represents a precursor number calculated from
933 multiple mice. Error bars represent mean ± SD.
934 (G) The donor chimerism of recipient mice, and the CD45.2$^+$ chimerism differences between high- and
935 low- precursor number groups. Each solid dot represents one animal. Each outlined dot represents
936 the average value of a replicate containing multiple animals. Recipient sample size = 7-14 mice
937 per replicate, N = 3 replicates for each precursor number group. Two-way ANOVA for paired
938 samples was performed.
939 (H) The frequency of myeloid cells in the peripheral blood of recipients. Each dot represents one
940 animal. Recipient sample size = 7-14 mice per replicate, N = 3 replicates for each precursor
941 number group. Two-way ANOVA for paired samples was performed.
942 ns, non-significant, * $p < 0.05$, ** $p < 0.001$.
943
944 Figure 3 Quantification of active hematopoietic precursor at steady-state
945 (A) Experiment schematic for Confetti induction in adult-induced animals. Image was created with
946 Biorender.
947 (B) Confetti labeling in B cells, myeloid cells and BM HSPCs. Statistics for comparisons between
948 labeling of month seven and other months were shown. One-way ANOVA for paired samples was
949 performed.
950 (C) RFP_{Conf}% in PB myeloid cells and BM HSPCs. Statistics for comparisons between labeling of
951 month seven and other months were shown. One-way ANOVA for paired samples was performed.
952 (D) The number of myeloid precursors and the average precursor number from month five to seven.
953 (E) The number of BM precursors. Paired permutation was performed.
954 (F) The frequency-to-clone ratio of bone marrow HSPCs. Paired permutation test was performed.
955 Each solid dot represents one animal. Each outlined dot represents the average value of a replicate
956 containing multiple animals. Error bars represent mean ± SD. PB sample size = 14-22 mice per
957 replicate; BM sample size = 8-10 mice per replicate; N = 4 replicates. ns, non-significant, * $p < 0.05$,
958 ** $p < 0.001$.
959
960 Figure 4. Number of precursor post-5-FU treatment and along developmental ontogeny
961 (A) Experiment schematic for one-dose 5-FU treatment. Image was created with Biorender.
962 (B) The number of BM precursors post-5FU treatment. Permutation test was performed. For UT,
963 sample size = 8-10 mice per replicate; N = 4 replicates; for 5-FU, sample size = 6-8 mice per
964 replicate; N = 5 replicates.
965 (C) Experiment schematic for Confetti induction at fetal stages. Image was created with Biorender.

966 (D) The Confetti labeling of B cells, T cells and myeloid cells. Each dot represents one animal. Each
967 outlined dot represents the average value of a replicate containing multiple animals. Sample size
968 = 4-11 mice per replicate; N = 5 replicates.
969 (E) The number of precursors in PB myeloid cells, B cells and T cells. Permutation test was performed.
970 For E14.5 and adult-induction, N = 4 replicates; for E11.5, n = 3 replicates.
971 (F) The number of BM precursors. Each dots represent one precursor number. Permutation test was
972 performed. For E14.5 and adult-induction, N = 4 replicates; for E11.5, N = 3 replicates.
973 (G) The relative fold change increase of precursor numbers from fetal- to adult-stage, as well as fold
974 change of CRU and cell counts.
975 Error bars represent mean ± SD. ns, non-significant, * $p < 0.05$, ** $p < 0.001$.
976
977 Figure 5. Precursor numbers in a mouse model of FA
978 (A) Experimental workflow to generate *Fancc*$^{+/+}$ and *Fancc*$^{-/-}$ mice, induce Confetti expression and
979 track precursor number dynamics. Image was created with Biorender.
980 (B) Confetti labeling in PB B cells and myeloid cells. Each dot represents one animal. Each outlined
981 dot represents the average value of a replicate containing multiple animals. Dashed lines connect
982 the values for the same replicate.
983 (C) The number of PB myeloid precursor.
984 (D) The number of BM precursor.
985 (E) Experiment schematic for competitive transplantation. Image was created with Biorender.
986 (F) PB donor chimerism in recipient mice of young FA donors. Each dot represents one animal.
987 Dashed lines connect the average donor chimerism of three replicates. For Fancc$^{+/+}$ or Fancc$^{+/-}$, n
988 = 10 per replicate, N = 3 replicates; for Fancc$^{-/-}$, N = 7-9 per replicate, N = 3 replicates. Two-way
989 ANOVA was performed.
990 (G) The number of PB myeloid precursor post-transplantation of young FA donors. Dashed lines
991 connect the average precursor numbers. Permutation test was performed. N = 3 for each group.
992 (H) PB donor chimerism recipient mice of aging FA donor cells. Each dot represents one animal.
993 Dashed lines connect the average donor chimerism of three replicates. For both genotype, N = 10
994 per replicate, N = 3 replicates. Two-way ANOVA was performed.
995 (I) The number of PB B precursor post-transplantation of aging FA donors. Dashed lines connect the
996 average precursor numbers. Permutation test was performed. N = 3 for each genotype.
997 For (A-C), Fancc$^{+/+}$, sample size = 11-17 mice per replicate, N =5 replicates; Fancc$^{+/-}$, sample size =
998 10-22 mice per replicate, N =7 replicates; Fancc$^{-/-}$, sample size = 9-17 mice per replicate, N =4
999 replicates. For (D), Fancc$^{+/+}$, sample size = 5-9 mice per replicate, N =5 replicates; Fancc$^{+/-}$, sample
1000 size = 6-12 mice per replicate, N =5 replicates; Fancc$^{-/-}$, sample size = 6-11 mice per replicate, N =4

1001 replicates. Error bars represent mean ± SD. nd, not determined; ns, non-significant; * $p < 0.05$, ** $p <$
1002 0.001. # represents lowest p value possible for permutation test.
1003
1004 Supplementary Figure Legend
1005 Figure S1: Establishing the correlation between variance of FP and precursor numbers.
1006 (A) The correlation between precursor numbers and CVs of individual FP or all CVs (original
1007 correlation formula), replotted from Ganuza et al., 2017.
1008 (B) FP_{mean}% in a previously published dataset, replotted from Ganuza et al., 2017.
1009 (C) The correlation between precursor numbers and standard deviation (σ) of individual FP, replotted
1010 from Ganuza et al., 2017.
1011 (D) The BFP% in HL-60 mix during two weeks of growth.
1012 (E) The effect of outliers on precursor number estimation. Left, BFP% in samples with/without outliers.
1013 Quartiles by Tukey method was shown. The outlier is highlighted by circle. Middle, standard
1014 deviation of BFP% with/without outliers. Right, estimated precursor numbers with/without outliers.
1015 (F) The changes of standard deviation when recorded events varied. "&" denotes the potential
1016 overestimated standard deviation due to low flow recorded events.
1017 (G) The variance of estimated precursor numbers when varying sample sizes were used.
1018 (H) The relative error of estimation as sample sizes increased. Each dot represents data from one
1019 seeded number. Error bars represent mean ± SD.
1020
1021 Figure S2: The induction specificity of HSC-Scl-CreER^T
1022 (A) Experiment schematic to examine the induction specificity of HSC-Scl-CreER^T. N =10 for
1023 induced animals, N =3 for non-induced animals. Image was created with Biorender.
1024 (B) Flow gating strategy for BM HSPC and MPs.
1025 (C) Confetti labeling in BM HSPCs. Error bars represent mean ± SD.
1026 (D) Flow gating strategy for BM lineage cells.
1027 (E) Confetti labeling in BM lineage cells. Error bars represent mean ± SD.
1028 (F) Flow gating strategy to examine spleen cells.
1029 (G) Confetti labeling in spleen cells. Error bars represent mean ± SD.
1030 See Methods and Table S1 for the detailed antibody, fluorophore and marker combination used for
1031 each cell type.
1032
1033 Figure S3: Analysis of HSC-Scl-CreER^T background Cre activity
1034 (A) Representative flow plot showing Confetti expression in PB T cells in 48-50-week-old HSC-Scl-
1035 CreER^T/Confetti animals compared to an animal induced with one-day tamoxifen treatment.
1036 (B) Quantification of RFP and YFP% in PB T cells. N =8 for non-induced 48-50-week-old animals.

1037 (C) Representative flow plot for Confetti labeling in PB T cells post HSC-Scl-CreER^T-mediated
1038 induction.
1039 (H) The level of CFP$^+$ and RFP$^+$ double positive cells in PB T cells of animals induced by HSC-Scl-
1040 CreER^T compared to Vav-Cre/Confetti animals analyzed at the same age. N =3 for Vav-Cre
1041 animals, N= 8 for HSC-Scl-CreER^T animals.
1042
1043
1044
1045 Figure S4: Variance of FP% inversely correlates with precursor numbers *in vivo*
1046 (A) The correlation between expected numbers and the estimated numbers calculated from FP% of
1047 precursors following Poisson distribution. See Methods for simulation details.
1048 (B) Flow gating strategy for PB B cell, T cell and myeloid cell. See Methods and Table S1 for the
1049 detailed antibody, fluorophore and marker combination used for each cell type.
1050 (C) RFP_{Conf}% distribution in B cell and T cell.
1051 (D) YFP_{Conf}% distribution in myeloid cell.
1052 (E) CFP_{Conf}% distribution in myeloid cell.
1053 (F) RFP_{PB}%, CFP_{PB}%, YFP_{PB}% and Confetti% distribution in myeloid cells.
1054 (G) Linear correlation of PB precursor numbers calculated from RFP_{Conf}%, RFP_{PB}%, CFP_{PB}%,
1055 YFP_{PB}%, Confetti% or $RFP_{RFP+YFP}$%, Each dot represents a precursor number calculated from
1056 multiple mice.
1057 For (C-F), each dot represents one animal. Recipient sample size = 7-14 mice per replicate, N =3
1058 replicates per group. Error bars represent mean ± SD.
1059
1060
1061 Figure S5 BM analysis for high- and low-precursor number groups
1062 (A) RFP_{Conf}% distribution in BM HSPCs. Each dot represents one animal
1063 (B) BM donor chimerism.
1064 (C) BM nucleated cell numbers. N =15 mice for high precursor number group, n= 14 for low precursor
1065 number group. Unpaired t-test was performed.
1066 (D) BM HSPC and MP frequency in donor CD45.2$^+$ population. Two-way ANOVA for paired samples
1067 was performed.
1068 For (B) and (D), each solid dot represents one animal; each outlined dot represents the average value
1069 of a replicate containing multiple animals. Except for (C), recipient sample size = 7-14 mice per
1070 replicate, N =3 replicates per group. For all graphs, Error bars represent mean ± SD. ns, non-significant,
1071 * $p < 0.05$.
1072

1073 Figure S6 Analysis of FP induction in adult-induced animals
1074 (A) Confetti labeling in LSK after 14days of tamoxifen treatment. N =5 animals. Error bars represent
1075 mean ± SD.
1076 (B) CFP_{Conf}% in PB myeloid cells.
1077 (C) CFP_{PB}% in PB myeloid cells.
1078 (D) $RFP_{RFP+YFP}$% in PB myeloid cells.
1079 (E) RFP_{PB}% and YFP_{PB}% in PB myeloid cells.
1080 (F) CFP_{Conf}% in PB myeloid cells of fetal-induced animals
1081 (G) BM HSPC frequencies. Error bars represent mean ± SD.
1082 For (B-E) sample size = 14-22 mice per replicate; for (F), sample size = 4-11 mice per replicate; for
1083 (G) sample size = 8-10 mice per replicate. For (B-E) and (G), N =4 replicates. For (F), N =5 replicates.
1084
1085 Figure S7: Analysis related to 5-FU treatment and Confetti fetal induction
1086 (A) PB myeloid cells frequencies before 5-FU treatment and ten days post-5-FU treatment. Dashed
1087 lines connect the values for the same replicate. Sample size = 6-8 mice per replicate, N =5
1088 replicates for 5-FU. Paired t test was performed.
1089 (B) The BM HSPC frequencies of UT and 5-FU treated groups ten days post-5-FU treatment. For UT,
1090 sample size = 8-10 mice per replicate, N =4 replicates; for 5-FU, sample size = 6-8 mice per
1091 replicate, N =5 replicates.
1092 (C) Confetti% in PB in E11.5-labeled animals. Dashed lines connect values for the same replicate.
1093 (D) The distribution of FPs in $Confetti^{+}$ population of B cells. Data showing average from all mice. N
1094 =24-34 for E14.5, N =7-24 for E11.5. Error bars represent mean ± SD.
1095 For (A-C), each solid dot represents one animal; each outlined dot represents the average value of a
1096 replicate containing multiple animals. ns, non-significant; * $p < 0.05$, ** $p < 0.001$.
1097
1098 Figure S8 Analysis of FA mice at steady-state
1099 (A) PB myeloid cell frequencies. Each dot represents one animal. Each outlined dot represents the
1100 average value of a replicate containing multiple animals.
1101 (B) PB B cell frequencies. Each dot represents one animal. Each outlined dot represents the average
1102 value of a replicate containing multiple animals.
1103 (C) Complete blood count analysis. Each dot represents one animal. N =13 for *Fancc*$^{+/+}$, N =16 for
1104 *Fancc*$^{-/-}$ mice. Unpaired t test was performed.
1105 (D) BM HSPC frequencies. Each dot represents one animal. Each outlined dot represents the average
1106 value of a replicate containing multiple animals. One-way ANOVA was performed.
1107 (E) BM nucleated cells counts. Each dot represents one animal. Unpaired t test was performed. N =12
1108 for *Fancc*$^{+/+}$and *Fancc*$^{-/-}$ mice.

1109 (F) $RFP_{RFP+YFP}$% in PB myeloid cells and B cells. Each dot represents one animal. Each outlined dot
1110 represents the average value of a replicate containing multiple animals.
1111 For (A-B) and (F), Fancc$^{+/+}$, sample size = 11-17 mice per replicate, N =5 replicates; Fancc$^{+/-}$, sample
1112 size = 10-22 mice per replicate, N =7 replicates; Fancc$^{-/-}$, sample size = 9-17 mice per replicate, N =4
1113 replicates. For (D), Fancc$^{+/+}$, sample size = 5-9 mice per replicate, N =5 replicates; Fancc$^{+/-}$, sample
1114 size = 6-12 mice per replicate, N =5 replicates; Fancc$^{-/-}$, sample size = 6-11 mice per replicate, N =4
1115 replicates. Error bars represent mean ± SD. ns, non-significant; * $p < 0.05$.
1116
1117 Figure S9 Analysis of FA recipient mice
1118 (A) Young *Fancc* donor BM HSPC frequency analysis. N = 5 for Fancc$^{+/+}$ or Fancc$^{+/-}$, N = 2 for
1119 Fancc$^{-/-}$.
1120 (B) Donor chimerism in recipient BM of young *Fancc* donor. Each dot represents one animal. Each
1121 outlined dot represents the average value of a replicate containing multiple animals. Two-way
1122 ANOVA was performed. Recipient sample size = 7-10 mice per replicate, N =3 replicates per
1123 group.
1124 (C) Number of PB B cell precursor in recipients of young FA donor.
1125 (D) Number of PB T cell precursor in recipients of young FA donor.
1126 (E) Number of BM HSPC precursor in recipients of young FA donor.
1127 (F) Ageing *Fancc* donor BM HSPC frequency analysis. N = 3 for Fancc$^{+/+}$, N = 3 for Fancc$^{-/-}$.
1128 (G) Number of PB myeloid cell precursor in recipients of aging FA donor.
1129 (H) Number of PB T cell precursor in recipients of aging FA donor.
1130 (I) Donor chimerism in recipient BM of ageing *Fancc* donor. Each dot represents one animal. Each
1131 outlined dot represents the average value of a replicate containing multiple animals. Two-way
1132 ANOVA was performed. Recipient sample size = 7-10 mice per replicate, N =3 replicates per
1133 group.
1134 (J) Number of BM HSPC precursor in recipients of aging FA donor.
1135 For (C-E), (G-H) and (J), N =3 replicates per group; permutation test was performed; dashed lines
1136 connect the average precursor numbers. Error bars represent mean ± SD. nd, not determined; ns,
1137 non-significant; * $p < 0.05$, ** $p < 0.001$. # represents lowest p value possible for permutation test.
1138

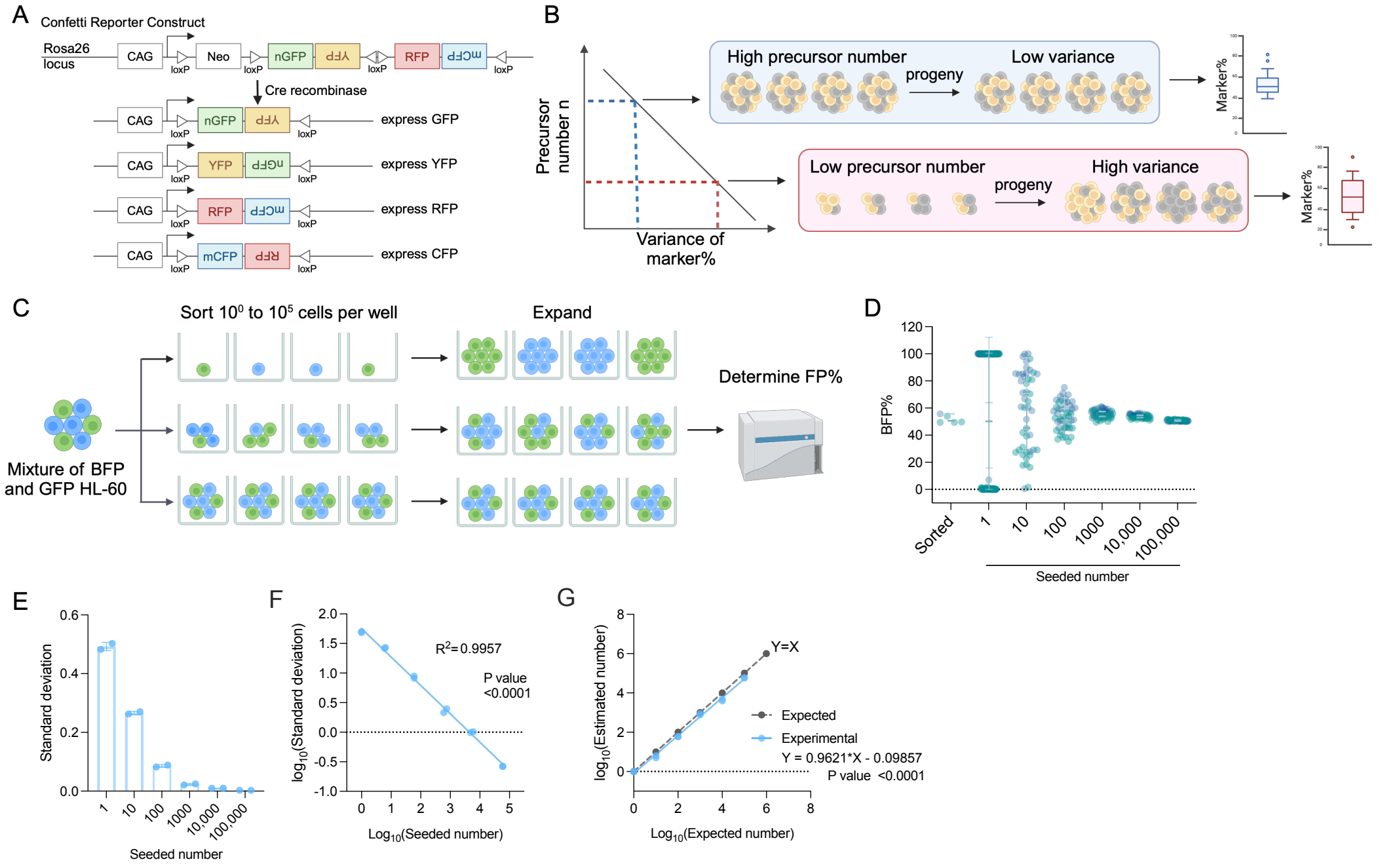
A
Confetti Reporter Construct
Rosa26 locus
CAG
loxP
Neo
nGFP
YFP
RFP
mCFP
Cre recombinase
express GFP
express YFP
express RFP
express CFP
B
Precursor number n
Variance of marker%
High precursor number
Low variance
progeny
Low precursor number
High variance
Marker%
C
Sort 10⁰ to 10⁵ cells per well
Expand
Determine FP%
Mixture of BFP and GFP HL-60
D
BFP%
Sorted
1
10
100
1000
10,000
100,000
Seeded number
E
Standard deviation
Seeded number
F
log10(Standard deviation)
R²=0.9957
P value <0.0001
Log10(Seeded number)
G
log10(Estimated number)
Y=X
Expected
Experimental
Y = 0.9621*X - 0.09857
P value <0.0001
Log10(Expected number)

Figure 1

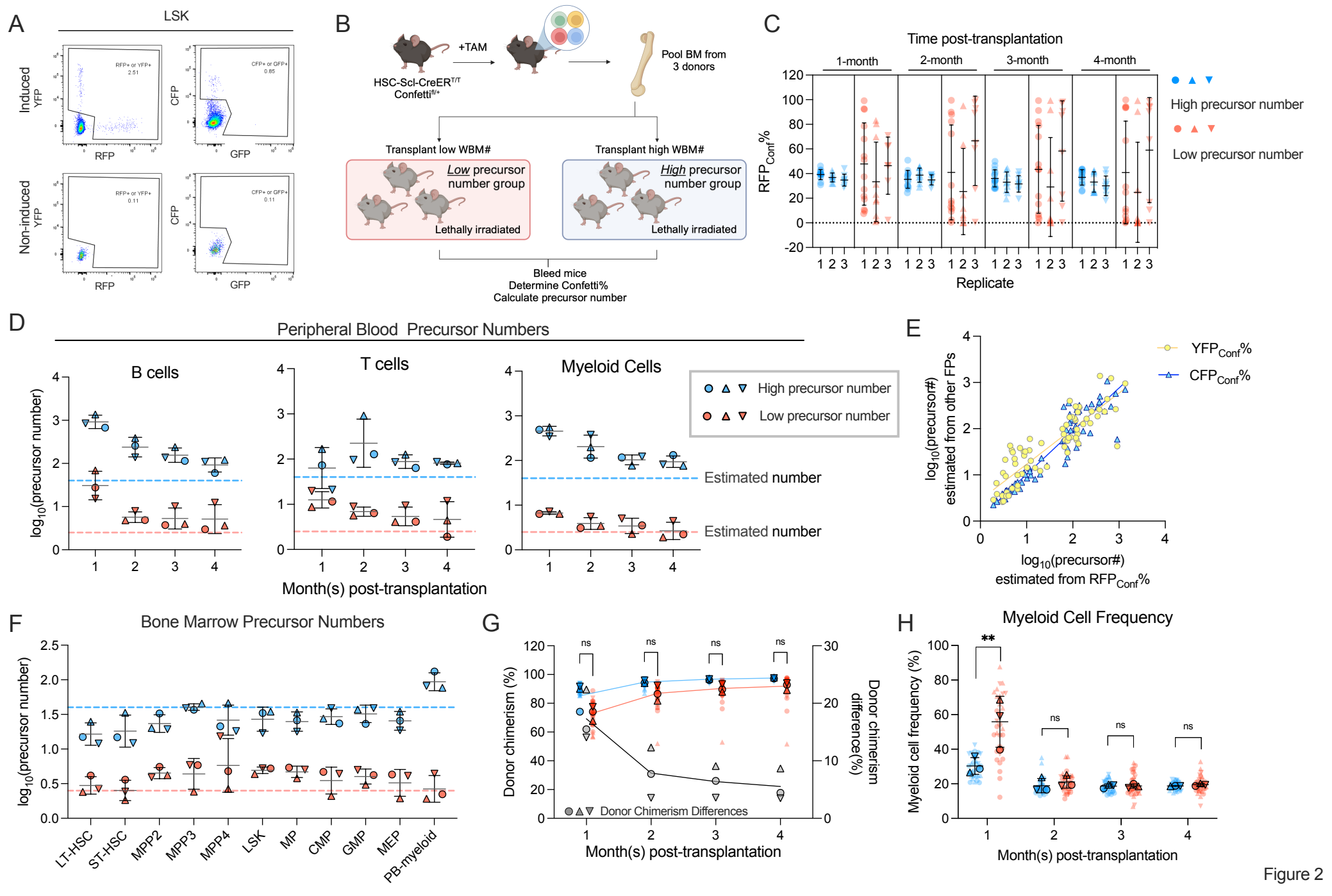
A
LSK
Induced
Non-induced
YFP
RFP
CFP
GFP
B
+TAM
HSC-Scl-CreER$^{T/T}$
Confetti$^{fl/+}$
Pool BM from
3 donors
Transplant low WBM#
Transplant high WBM#
Low precursor number group
High precursor number group
Lethally irradiated
Bleed mice
Determine Confetti%
Calculate precursor number
C
Time post-transplantation
1-month
2-month
3-month
4-month
RFP_{Conf}%
Replicate
High precursor number
Low precursor number
D
Peripheral Blood Precursor Numbers
B cells
T cells
Myeloid Cells
$\log_{10}$(precursor number)
Estimated number
Month(s) post-transplantation
E
YFP_{Conf}%
CFP_{Conf}%
$\log_{10}$(precursor#) estimated from other FPs
$\log_{10}$(precursor#) estimated from RFP_{Conf}%
F
Bone Marrow Precursor Numbers
LT-HSC
ST-HSC
MPP2
MPP3
MPP4
LSK
MP
CMP
GMP
MEP
PB-myeloid
G
Donor chimerism (%)
Donor chimerism difference (%)
ns
Donor Chimerism Differences
Month(s) post-transplantation
H
Myeloid Cell Frequency
**
Myeloid cell frequency (%)
Month(s) post-transplantation

Figure 2

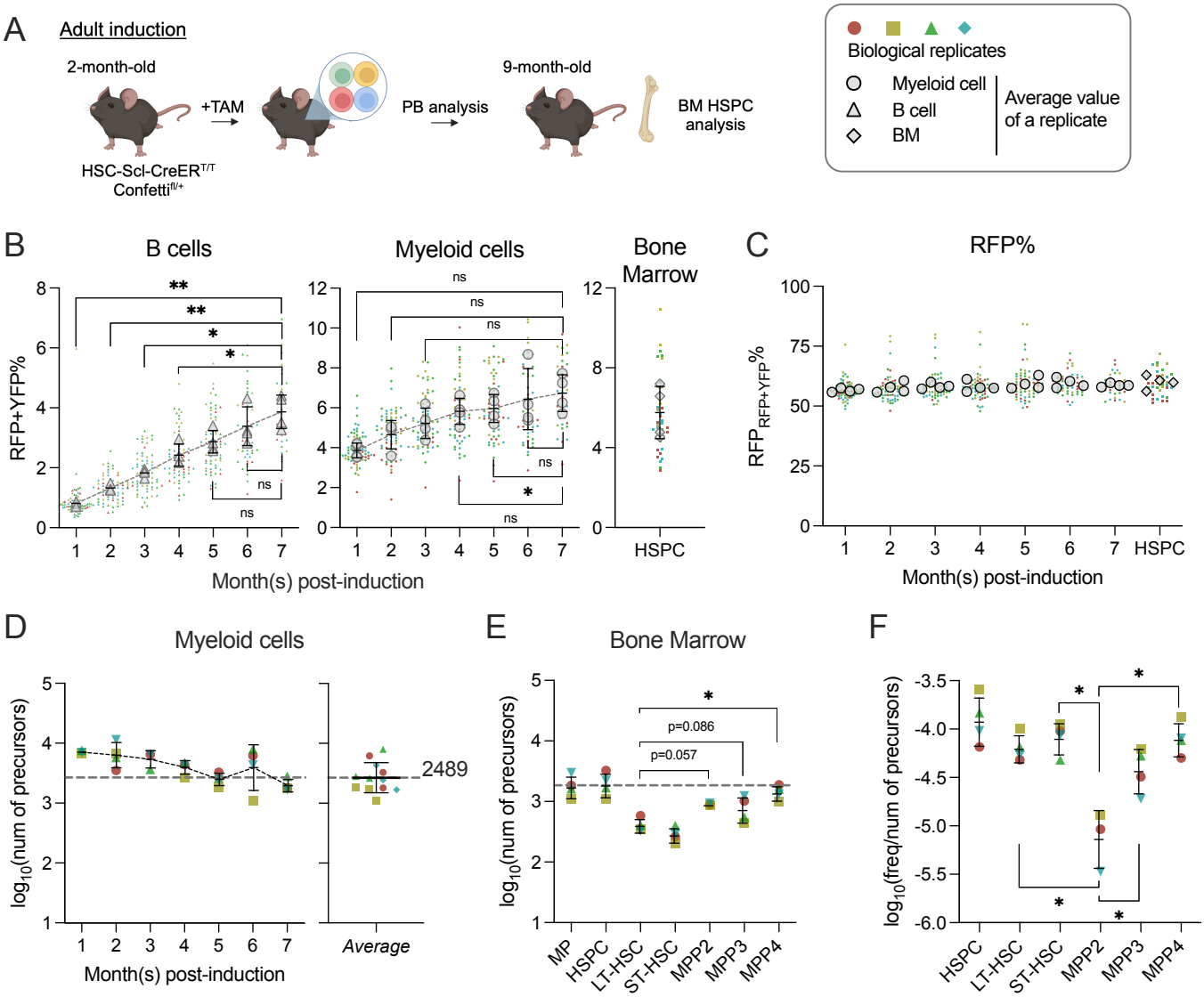
A
Adult induction
2-month-old
+TAM
PB analysis
9-month-old
BM HSPC analysis
HSC-Scl-CreER$^{T/T}$
Confetti$^{fl/+}$
Biological replicates
Myeloid cell
B cell
BM
Average value of a replicate
B
B cells
Myeloid cells
Bone Marrow
RFP+YFP%
Month(s) post-induction
HSPC
ns
C
RFP%
$RFP_{RFP+YFP}$%
D
Myeloid cells
log_{10}(num of precursors)
2489
Average
E
Bone Marrow
p=0.086
p=0.057
MP
HSPC
LT-HSC
ST-HSC
MPP2
MPP3
MPP4
F
log_{10}(freq/num of precursors)

Figure 3

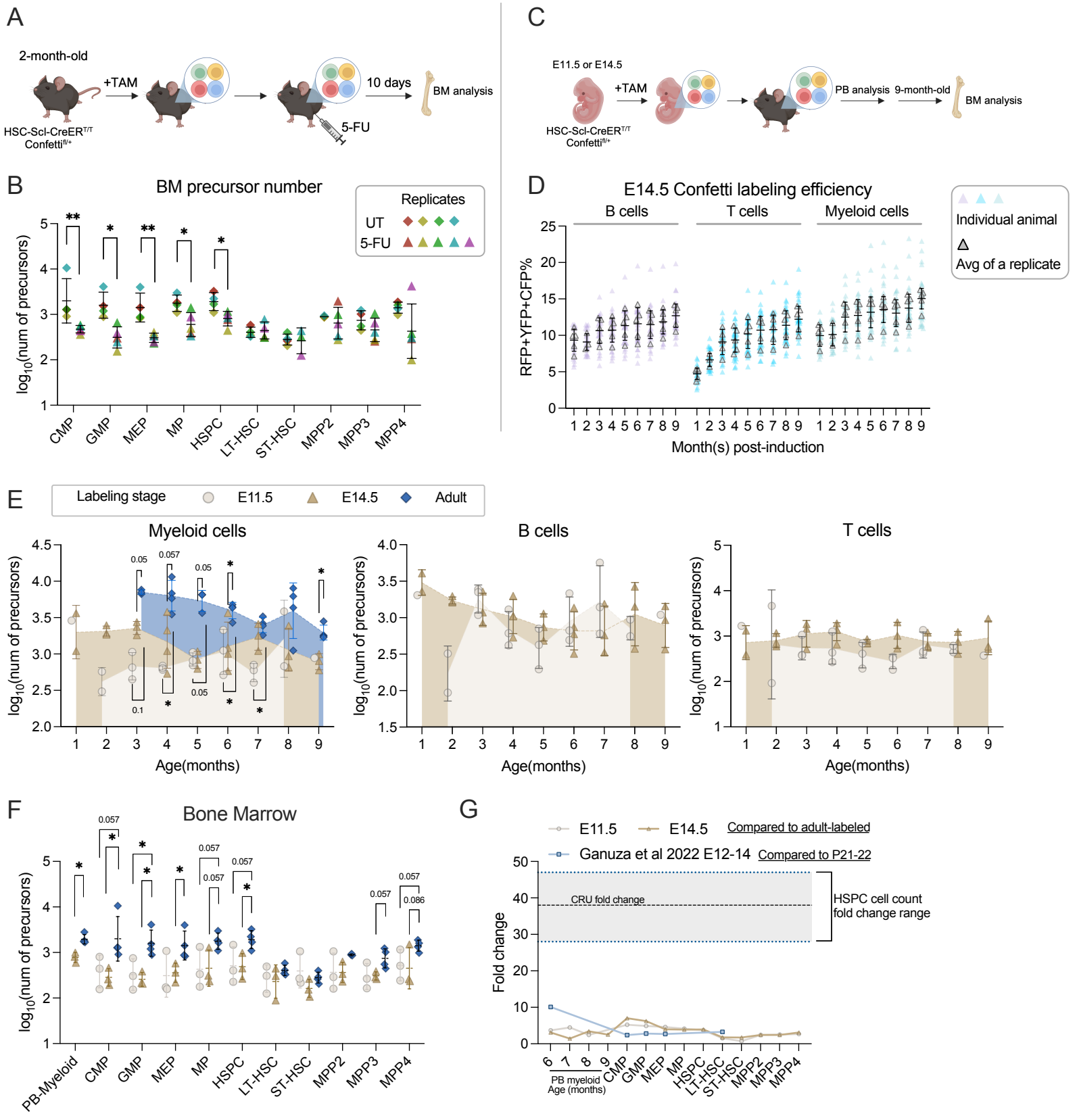
A
2-month-old
+TAM
10 days
BM analysis
5-FU
HSC-Scl-CreER$^{T/T}$
Confetti$^{fl/+}$
B
BM precursor number
Replicates
UT
5-FU
log$_{10}$(num of precursors)
CMP
GMP
MEP
MP
HSPC
LT-HSC
ST-HSC
MPP2
MPP3
MPP4
C
E11.5 or E14.5
+TAM
PB analysis
9-month-old
BM analysis
HSC-Scl-CreER$^{T/T}$
Confetti$^{fl/+}$
D
E14.5 Confetti labeling efficiency
B cells
T cells
Myeloid cells
Individual animal
Avg of a replicate
RFP+YFP+CFP%
Month(s) post-induction
E
Labeling stage
E11.5
E14.5
Adult
Myeloid cells
B cells
T cells
log$_{10}$(num of precursors)
Age(months)
F
Bone Marrow
log$_{10}$(num of precursors)
PB-Myeloid
CMP
GMP
MEP
MP
HSPC
LT-HSC
ST-HSC
MPP2
MPP3
MPP4
G
E11.5
E14.5
Compared to adult-labeled
Ganuza et al 2022 E12-14
Compared to P21-22
CRU fold change
HSPC cell count
fold change range
Fold change
PB myeloid
Age (months)

Figure 4

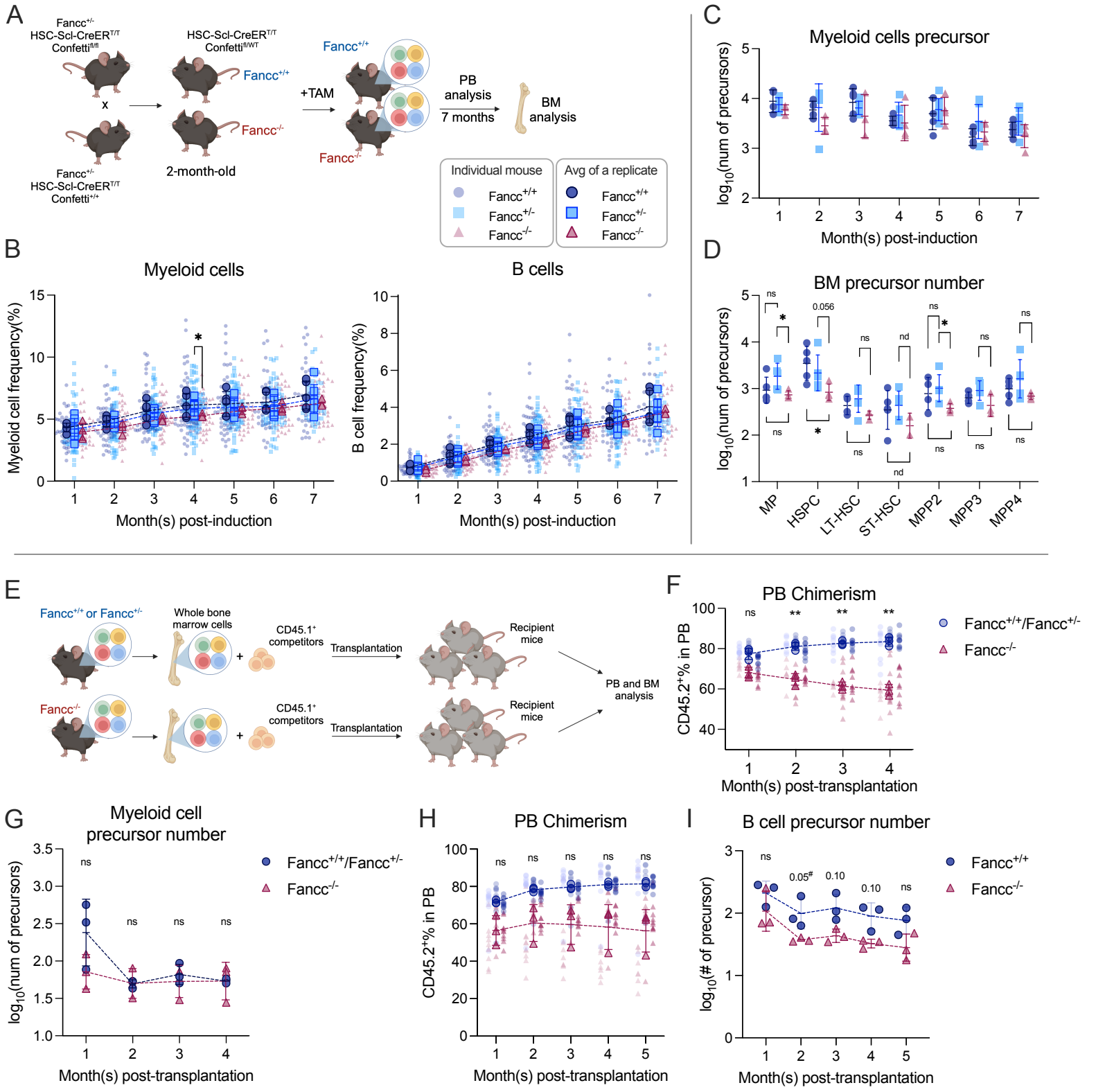
A
Fancc+/- HSC-Scl-CreERT/T Confettifl/fl
x
Fancc+/- HSC-Scl-CreERT/T Confetti+/+
HSC-Scl-CreERT/T Confettifl/WT
Fancc+/+
Fancc-/-
2-month-old
+TAM
Fancc+/+
Fancc-/-
PB analysis
7 months
BM analysis
Individual mouse
Fancc+/+
Fancc+/-
Fancc-/-
Avg of a replicate
Fancc+/+
Fancc+/-
Fancc-/-
B
Myeloid cells
Myeloid cell frequency(%)
Month(s) post-induction
B cells
B cell frequency(%)
Month(s) post-induction
C
Myeloid cells precursor
log10(num of precursors)
Month(s) post-induction
D
BM precursor number
log10(num of precursors)
MP
HSPC
LT-HSC
ST-HSC
MPP2
MPP3
MPP4
E
Fancc+/+ or Fancc+/-
Whole bone marrow cells
CD45.1+ competitors
Transplantation
Recipient mice
Fancc-/-
CD45.1+ competitors
Transplantation
Recipient mice
PB and BM analysis
F
PB Chimerism
CD45.2+% in PB
Month(s) post-transplantation
Fancc+/+/Fancc+/-
Fancc-/-
G
Myeloid cell precursor number
log10(num of precursors)
Month(s) post-transplantation
Fancc+/+/Fancc+/-
Fancc-/-
H
PB Chimerism
CD45.2+% in PB
Month(s) post-transplantation
I
B cell precursor number
log10(# of precursor)
Month(s) post-transplantation
Fancc+/+
Fancc-/-

Figure 5

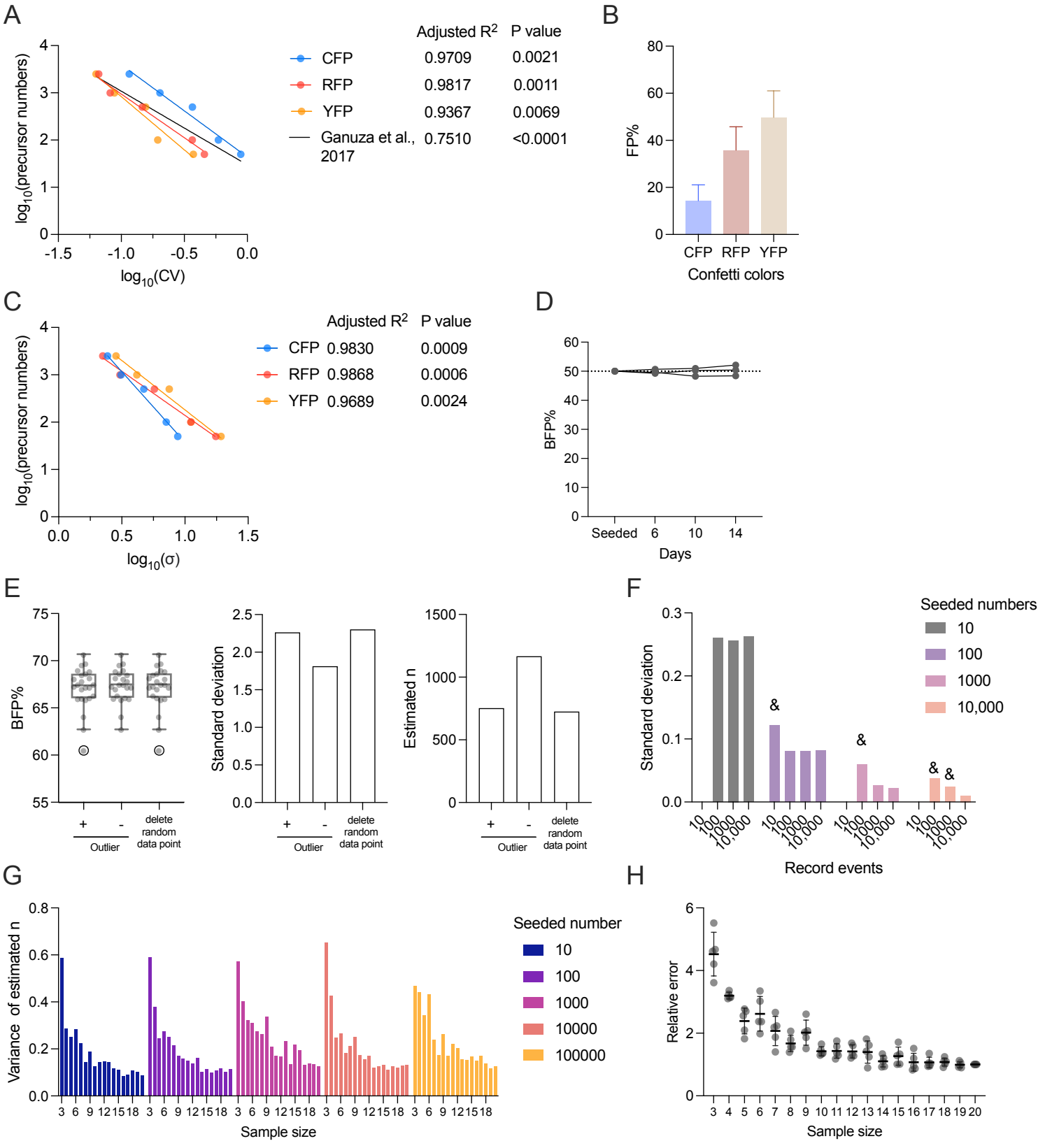
A
Adjusted R2
P value
CFP 0.9709 0.0021
RFP 0.9817 0.0011
YFP 0.9367 0.0069
Ganuza et al., 2017 0.7510 <0.0001
log10(precursor numbers)
log10(CV)
B
FP%
CFP RFP YFP
Confetti colors
C
Adjusted R2
P value
CFP 0.9830 0.0009
RFP 0.9868 0.0006
YFP 0.9689 0.0024
log10(precursor numbers)
log10(σ)
D
BFP%
Seeded 6 10 14
Days
E
BFP%
Standard deviation
Estimated n
+ - delete random data point
Outlier
F
Seeded numbers
10
100
1000
10,000
Standard deviation
Record events
G
Variance of estimated n
Seeded number
10
100
1000
10000
100000
Sample size
H
Relative error
Sample size

Figure S1

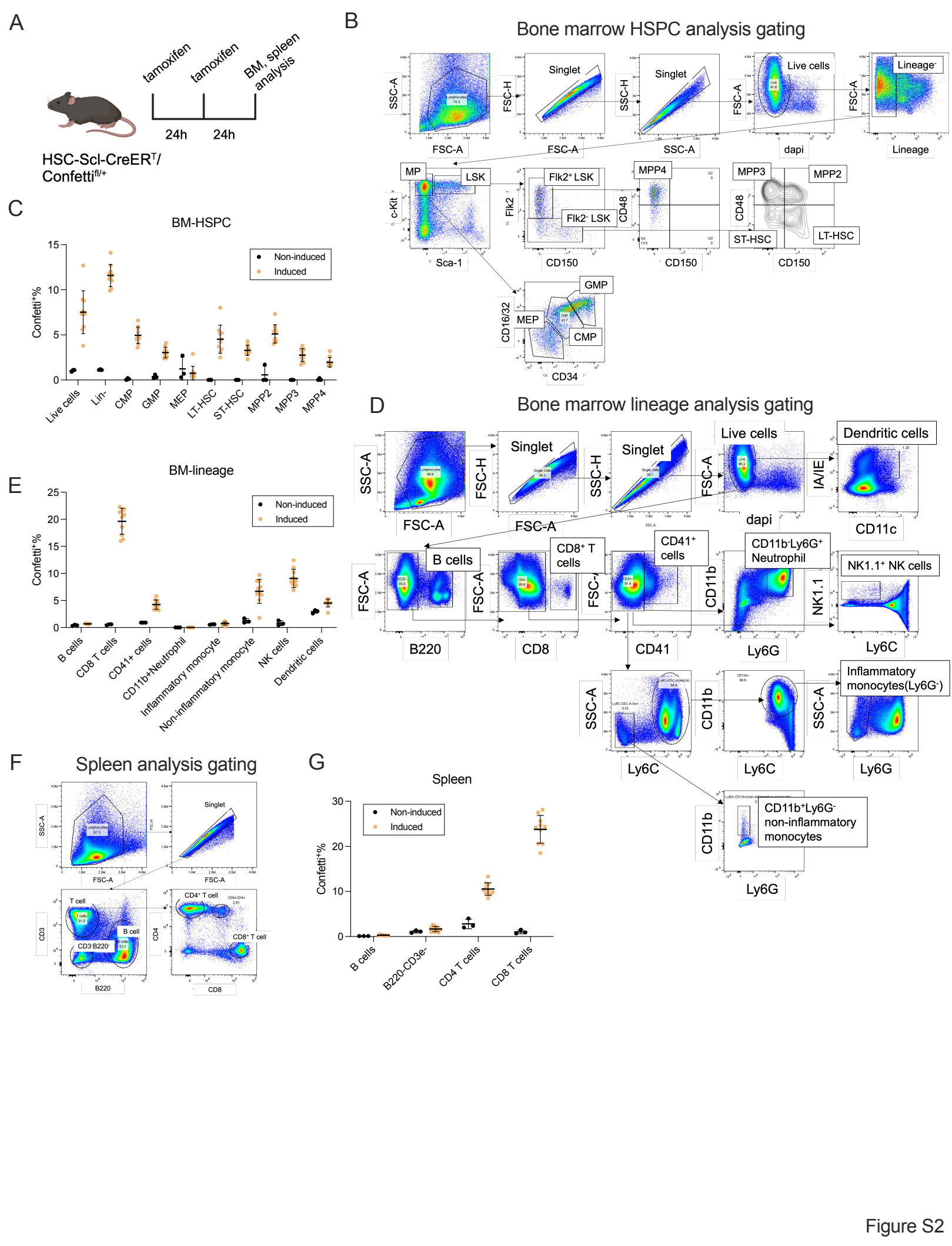
A
tamoxifen
tamoxifen
BM, spleen analysis
24h
24h
HSC-Scl-CreER^{T}/
Confetti$^{fl/+}$
B
Bone marrow HSPC analysis gating
SSC-A
FSC-A
Singlet
FSC-H
SSC-H
Live cells
dapi
Lineage^{-}
Lineage
MP
LSK
c-Kit
Sca-1
Flk2^{+} LSK
Flk2^{-} LSK
Flk2
CD150
MPP4
CD48
MPP3
MPP2
ST-HSC
LT-HSC
GMP
MEP
CMP
CD16/32
CD34
C
BM-HSPC
Confetti^{+}%
Non-induced
Induced
Live cells
Lin-
CMP
GMP
MEP
LT-HSC
ST-HSC
MPP2
MPP3
MPP4
D
Bone marrow lineage analysis gating
Dendritic cells
IA/IE
CD11c
B cells
B220
CD8^{+} T cells
CD8
CD41^{+} cells
CD41
CD11b^{-}Ly6G^{+} Neutrophil
CD11b
Ly6G
NK1.1^{+} NK cells
NK1.1
Ly6C
Inflammatory monocytes(Ly6G^{-})
CD11b^{+}Ly6G^{-} non-inflammatory monocytes
E
BM-lineage
Confetti^{+}%
Non-induced
Induced
B cells
CD8 T cells
CD41+ cells
CD11b+Neutrophil
Inflammatory monocyte
Non-inflammatory monocyte
NK cells
Dendritic cells
F
Spleen analysis gating
T cell
B cell
CD3^{-}B220^{-}
CD3
CD4^{+} T cell
CD8^{+} T cell
CD4
CD8
G
Spleen
Confetti^{+}%
Non-induced
Induced
B cells
B220-CD3e-
CD4 T cells
CD8 T cells

Figure S2

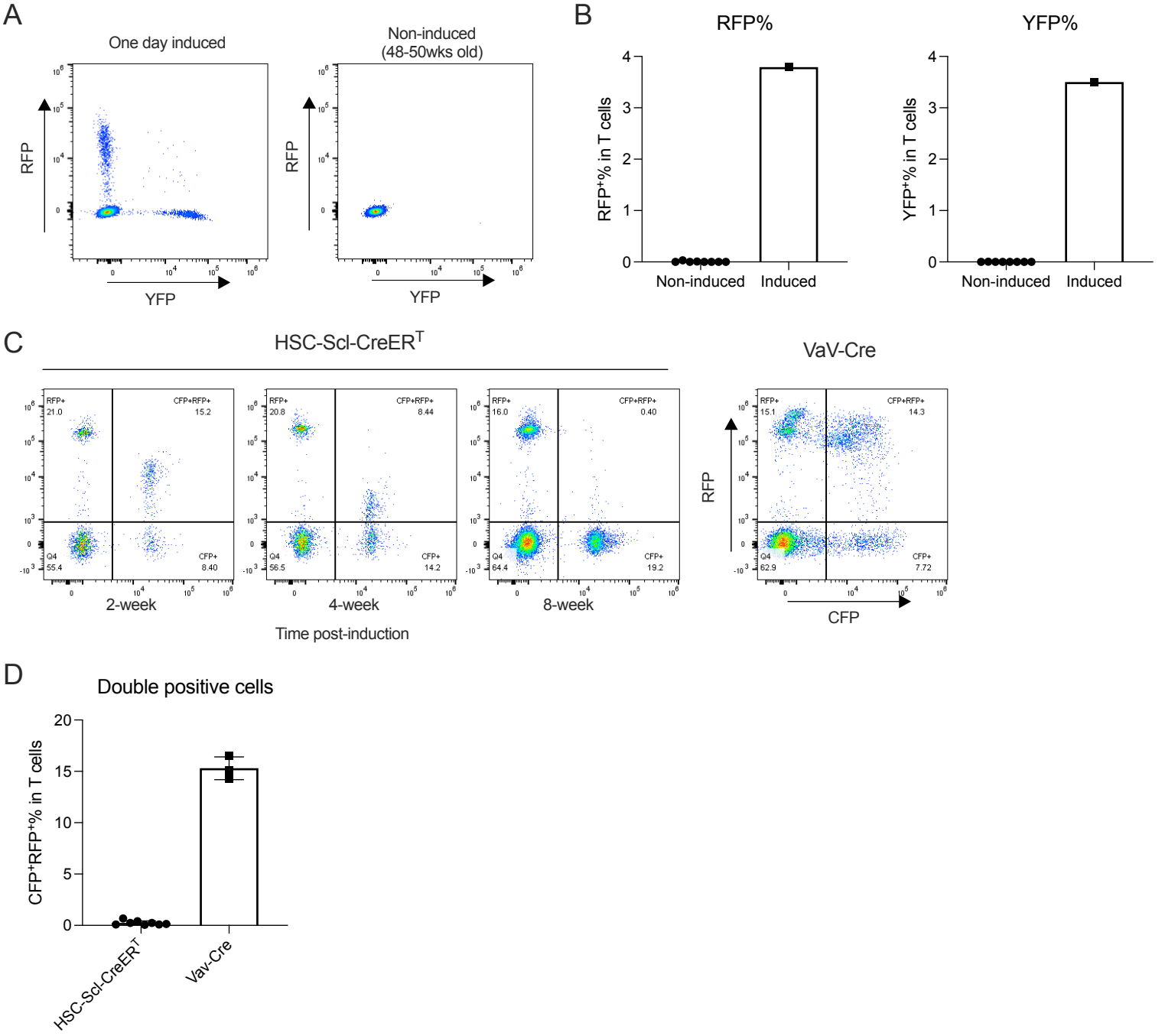
A
One day induced
Non-induced
(48-50wks old)
RFP
YFP
B
RFP%
YFP%
RFP+% in T cells
YFP+% in T cells
Non-induced
Induced
C
HSC-Scl-CreERT
VaV-Cre
RFP+
21.0
CFP+RFP+
15.2
Q4
55.4
CFP+
8.40
RFP+
20.8
CFP+RFP+
8.44
Q4
56.5
CFP+
14.2
RFP+
16.0
CFP+RFP+
0.40
Q4
64.4
CFP+
19.2
RFP+
15.1
CFP+RFP+
14.3
Q4
62.9
CFP+
7.72
2-week
4-week
8-week
Time post-induction
RFP
CFP
D
Double positive cells
CFP+RFP+% in T cells
HSC-Scl-CreERT
Vav-Cre

Figure S3

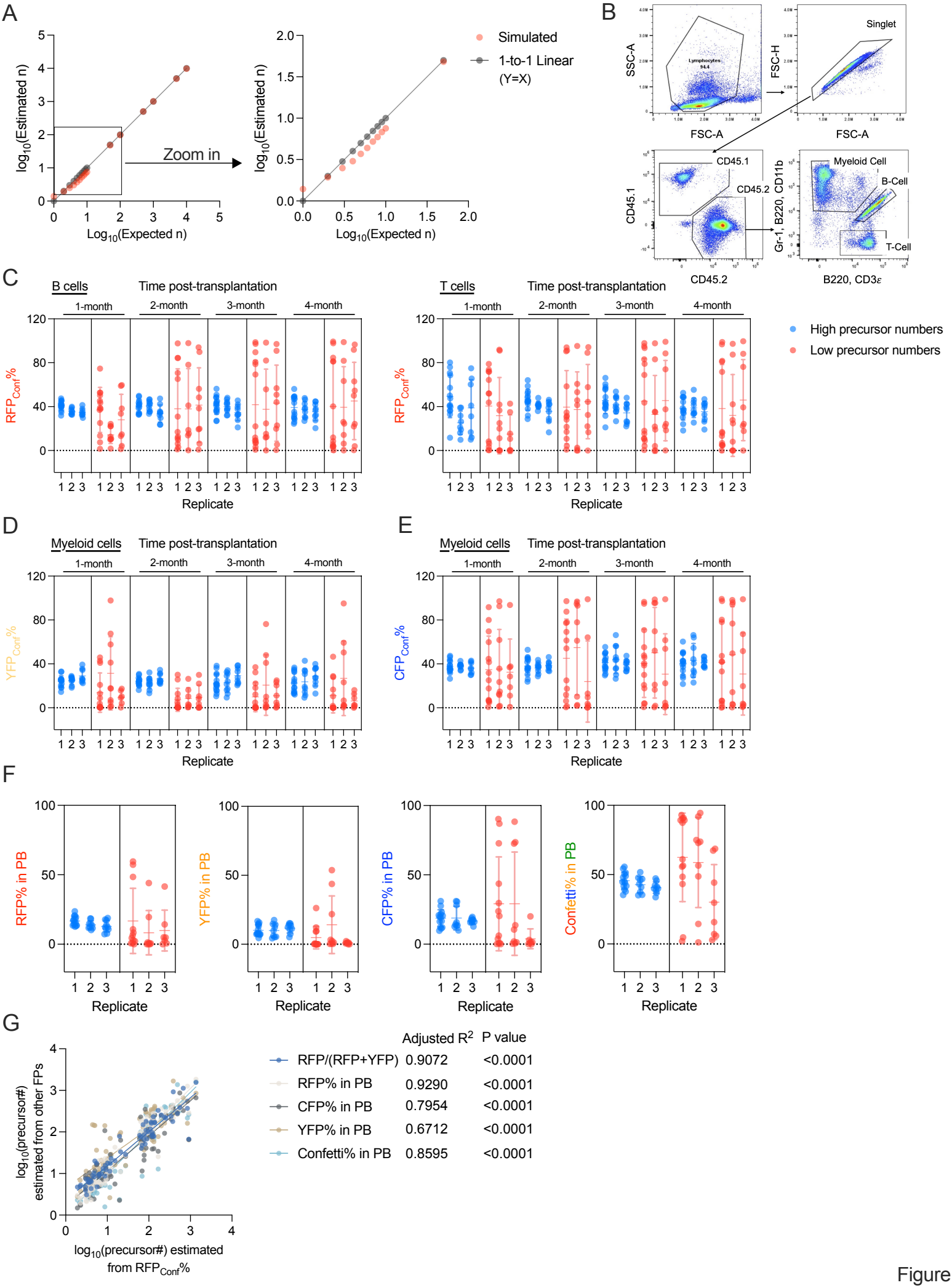
A
Simulated
1-to-1 Linear (Y=X)
Zoom in
log10(Estimated n)
Log10(Expected n)
B
Singlet
Lymphocytes
SSC-A
FSC-A
FSC-H
CD45.1
CD45.2
Myeloid Cell
B-Cell
T-Cell
Gr-1, B220, CD11b
B220, CD3ε
C
B cells
T cells
Time post-transplantation
1-month
2-month
3-month
4-month
RFPConf%
Replicate
High precursor numbers
Low precursor numbers
D
Myeloid cells
YFPConf%
E
CFPConf%
F
RFP% in PB
YFP% in PB
CFP% in PB
Confetti% in PB
G
log10(precursor#) estimated from other FPs
log10(precursor#) estimated from RFPConf%
Adjusted R2 P value
RFP/(RFP+YFP) 0.9072 <0.0001
RFP% in PB 0.9290 <0.0001
CFP% in PB 0.7954 <0.0001
YFP% in PB 0.6712 <0.0001
Confetti% in PB 0.8595 <0.0001

Figure S4

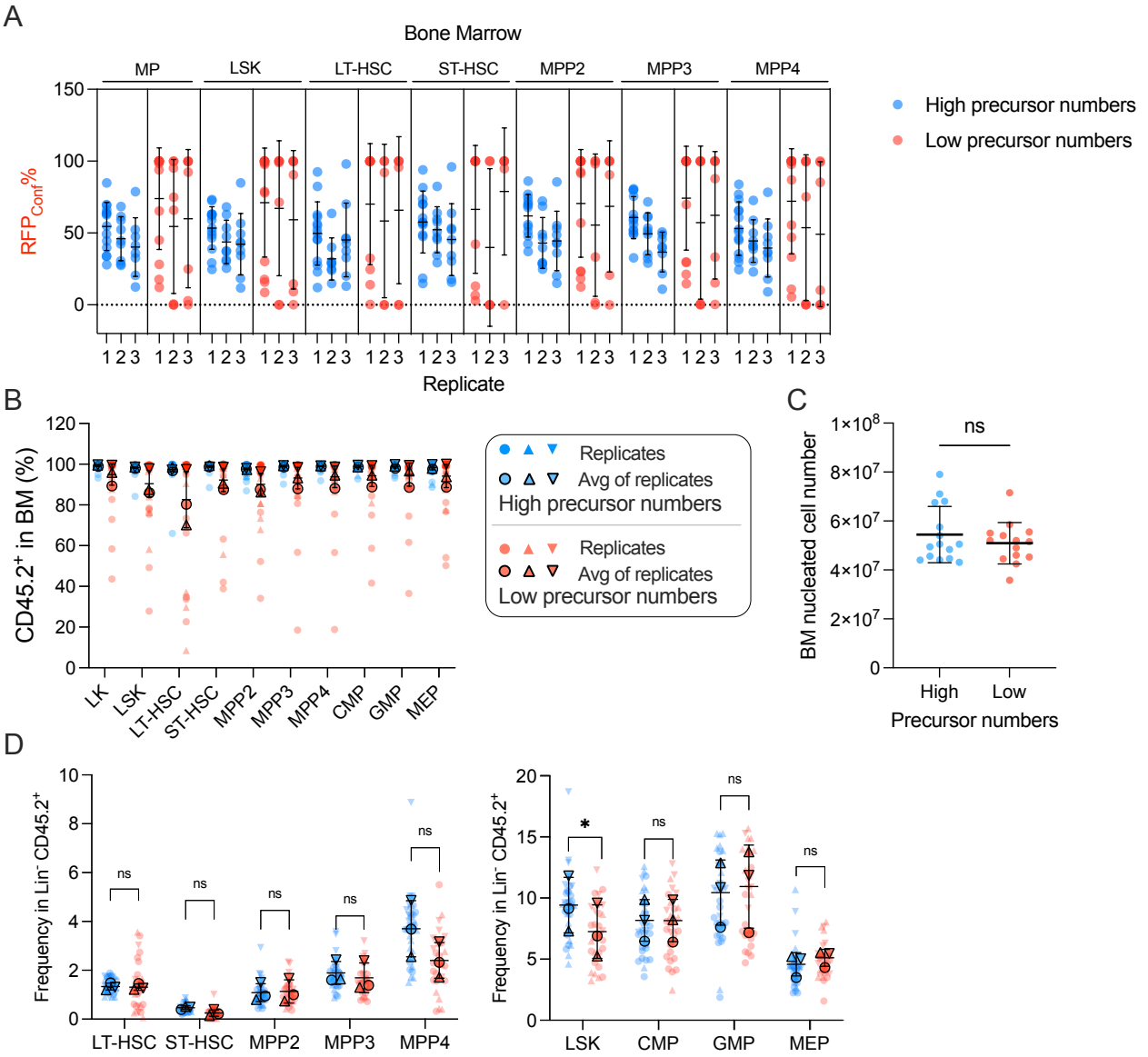
A
Bone Marrow
MP
LSK
LT-HSC
ST-HSC
MPP2
MPP3
MPP4
RFP_Conf %
Replicate
High precursor numbers
Low precursor numbers
B
CD45.2+ in BM (%)
LK
LSK
LT-HSC
ST-HSC
MPP2
MPP3
MPP4
CMP
GMP
MEP
Replicates
Avg of replicates
High precursor numbers
Replicates
Avg of replicates
Low precursor numbers
C
BM nucleated cell number
ns
High
Low
Precursor numbers
D
Frequency in Lin- CD45.2+
ns
LT-HSC
ST-HSC
MPP2
MPP3
MPP4
Frequency in Lin- CD45.2+
*
ns
LSK
CMP
GMP
MEP

Figure S5

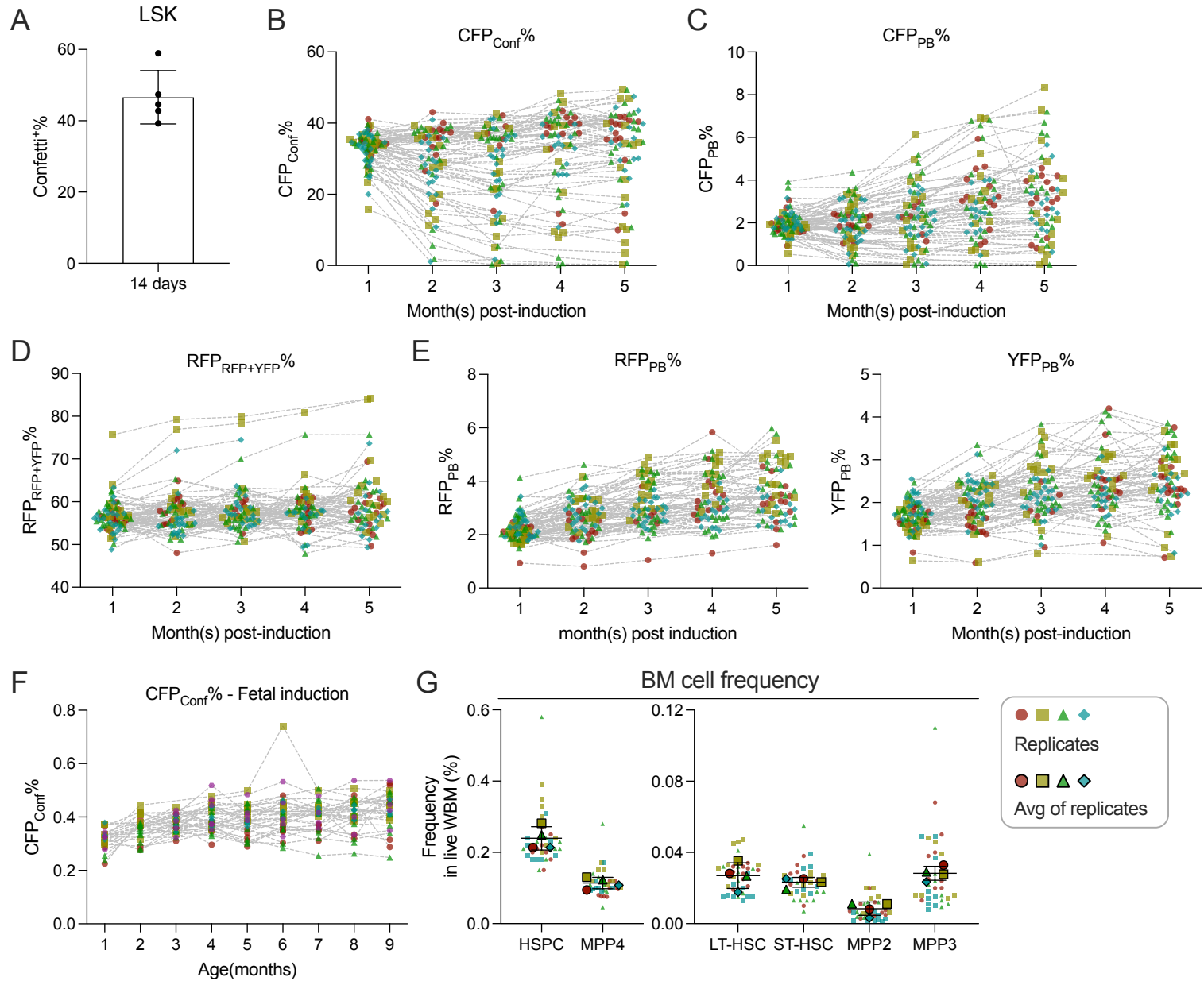
A
LSK
Confetti+%
14 days
B
CFPConf%
Month(s) post-induction
C
CFPPB%
Month(s) post-induction
D
RFPRFP+YFP%
Month(s) post-induction
E
RFPPB%
month(s) post induction
YFPPB%
Month(s) post-induction
F
CFPConf% - Fetal induction
Age(months)
G
BM cell frequency
Frequency in live WBM (%)
HSPC
MPP4
LT-HSC
ST-HSC
MPP2
MPP3
Replicates
Avg of replicates

Figure S6

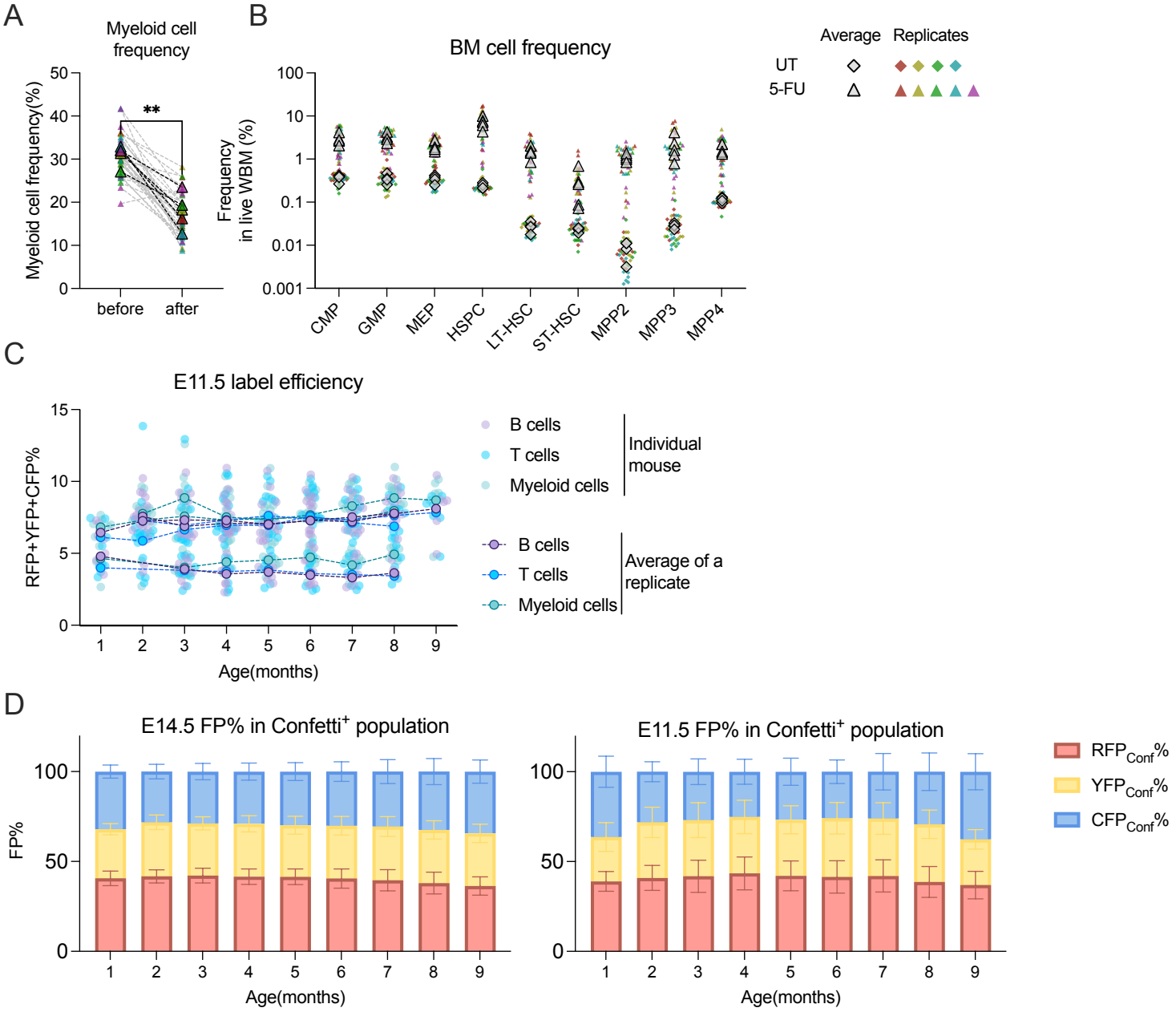
A
Myeloid cell frequency
Myeloid cell frequency(%)
50
40
30
20
10
0
**
before
after
B
BM cell frequency
Frequency in live WBM (%)
100
10
1
0.1
0.01
0.001
CMP
GMP
MEP
HSPC
LT-HSC
ST-HSC
MPP2
MPP3
MPP4
Average
Replicates
UT
5-FU
C
E11.5 label efficiency
RFP+YFP+CFP%
15
10
5
0
1
2
3
4
5
6
7
8
9
Age(months)
B cells
T cells
Myeloid cells
Individual mouse
B cells
T cells
Myeloid cells
Average of a replicate
D
E14.5 FP% in Confetti+ population
FP%
100
50
0
Age(months)
E11.5 FP% in Confetti+ population
RFPConf%
YFPConf%
CFPConf%

Figure S7

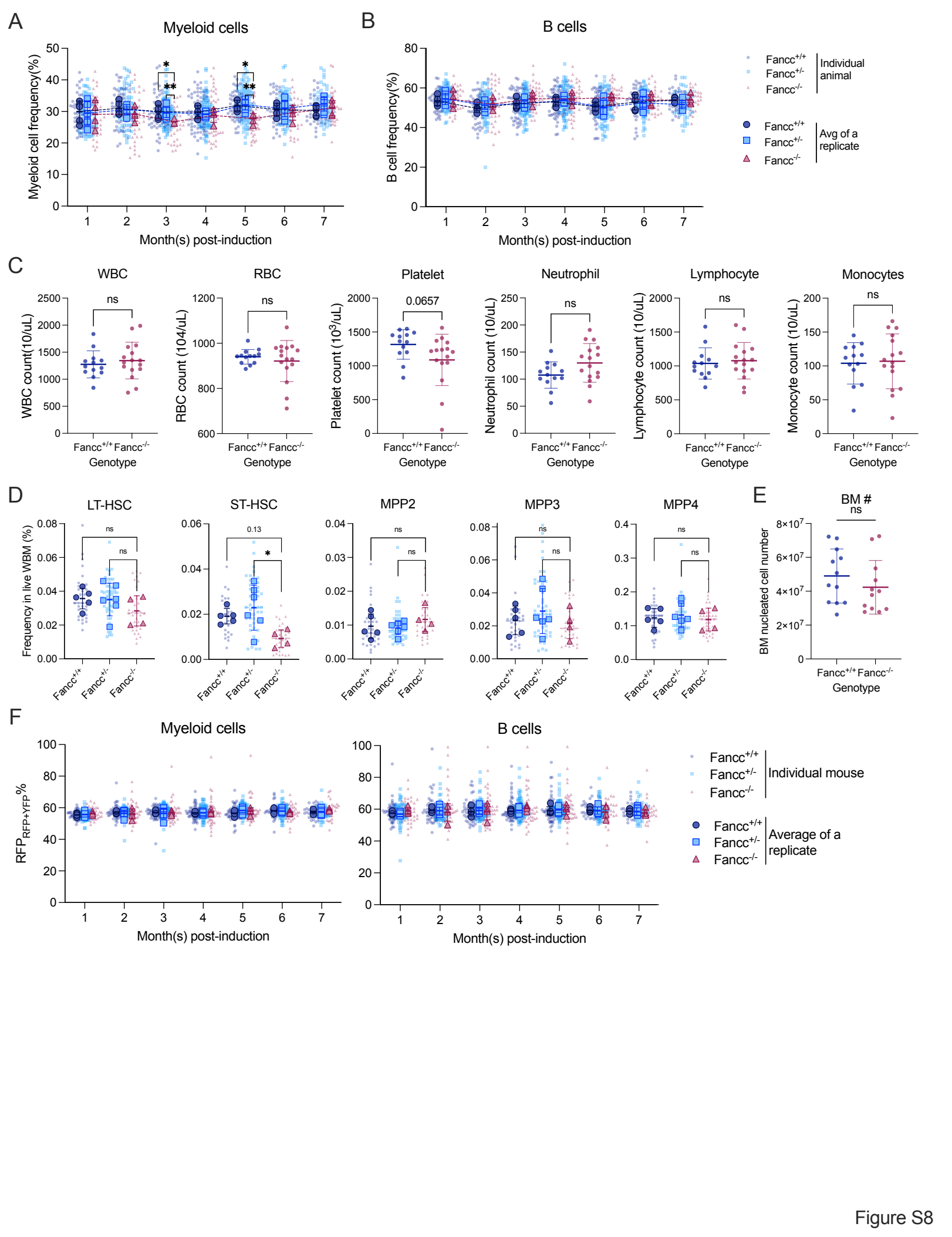
A
Myeloid cells
Myeloid cell frequency(%)
Month(s) post-induction
B
B cells
B cell frequency(%)
Month(s) post-induction
Fancc+/+
Fancc+/-
Fancc-/-
Individual animal
Avg of a replicate
C
WBC
RBC
Platelet
Neutrophil
Lymphocyte
Monocytes
WBC count(10/uL)
RBC count (104/uL)
Platelet count (10^3/uL)
Neutrophil count (10/uL)
Lymphocyte count (10/uL)
Monocyte count (10/uL)
ns
0.0657
Genotype
D
LT-HSC
ST-HSC
MPP2
MPP3
MPP4
Frequency in live WBM (%)
0.13
E
BM #
BM nucleated cell number
F
Myeloid cells
B cells
RFP+YFP %
Month(s) post-induction
Individual mouse
Average of a replicate

Figure S8

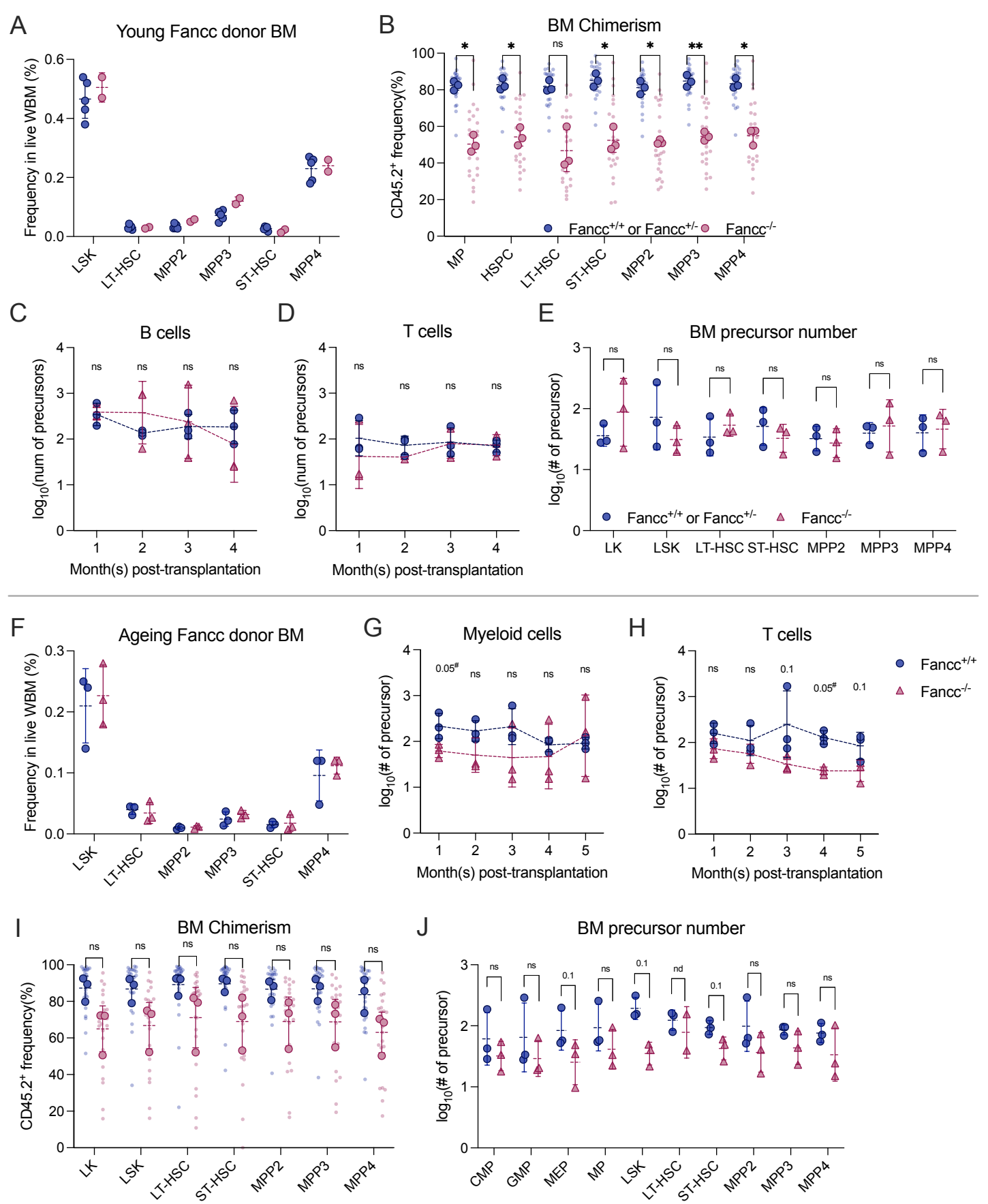
A
Young Fancc donor BM
Frequency in live WBM (%)
LSK
LT-HSC
MPP2
MPP3
ST-HSC
MPP4
B
BM Chimerism
CD45.2+ frequency(%)
MP
HSPC
LT-HSC
ST-HSC
MPP2
MPP3
MPP4
Fancc+/+ or Fancc+/-
Fancc-/-
C
B cells
log10(num of precursors)
Month(s) post-transplantation
D
T cells
log10(num of precursors)
Month(s) post-transplantation
E
BM precursor number
log10(# of precursor)
LK
LSK
LT-HSC
ST-HSC
MPP2
MPP3
MPP4
F
Ageing Fancc donor BM
Frequency in live WBM (%)
G
Myeloid cells
log10(# of precursor)
Month(s) post-transplantation
H
T cells
log10(# of precursor)
Month(s) post-transplantation
I
BM Chimerism
CD45.2+ frequency(%)
J
BM precursor number
log10(# of precursor)
CMP
GMP
MEP
MP
LSK
LT-HSC
ST-HSC
MPP2
MPP3
MPP4

Figure S9