

# Evidence that hematopoiesis may be a stochastic process *in vivo*

JANIS L. ABKOWITZ<sup>1</sup>, SANDRA N. CATLIN<sup>2</sup> & PETER GUTTORP<sup>2</sup>

<sup>1</sup>Division of Hematology, Department of Medicine, Box 357710,  
University of Washington, Seattle, Washington 98195, USA

<sup>2</sup>Department of Statistics, University of Washington, Seattle, Washington 98195, USA  
Correspondence should be addressed to J.L.A.

To study the behavior of hematopoietic stem cells *in vivo*, hematopoiesis was simulated by assuming that all stem cell decisions (that is, replication, apoptosis, initiation of a differentiation/maturation program) were determined by chance. Predicted outcomes from simulated experiments were compared with data obtained in autologous marrow transplantation studies of glucose 6-phosphate dehydrogenase (G6PD) heterozygous female Safari cats. With this approach, we prove that stochastic differentiation can result in the wide spectrum of discrete outcomes observed *in vivo*, and that clonal dominance can occur by chance. As the analyses also suggest that the frequency of feline hematopoietic stem cells is only 6 per  $10^7$  nucleated marrow cells, and that stem cells do not replicate on average more frequently than once every three weeks, these large-animal data challenge clinical strategies for marrow transplantation and gene therapy.

Hematopoiesis is the ordered process in which stem cells, through sequential divisions, differentiate into mature blood cells. Much information is known concerning the commitment of progenitor cells to the red cell, granulocyte and platelet lineages. The antigenic characteristics of these cells, their growth factor requirements and their cell-cycle kinetics have been identified in both *in vitro* and *in vivo* analyses. It is at this level of differentiation that the day-to-day regulation of blood cell production occurs.

Although the commitment of a progenitor cell to a specific lineage is likely to be a stochastic (or probabilistic) event<sup>1-4</sup>, little is known about the kinetics of the earliest cells, the hematopoietic stem cells. Stem cells are difficult to isolate, because they do not have a unique physical or antigenic phenotype<sup>5-15</sup>. As such cells are most rigorously defined by their ability to reconstitute and to support long-term hematopoiesis after transplantation<sup>16</sup>, experimental analyses are especially complex.

We have studied hematopoiesis in female Safari cats because of their unique cellular marker. Safari cats are the offspring of matings between Geoffroy cats (a South American wildcat) and domestic cats (which are of Eurasian origin). These species have evolved independently for 12 million years<sup>17</sup>, and have electrophoretically distinct phenotypes of the X chromosome-linked enzyme glucose-6-phosphate dehydrogenase (G6PD). Because of X-chromosome inactivation during embryogenesis, female Safari cats have some somatic cells that contain domestic-type G6PD (d G6PD) and other cells that contain Geoffroy-type G6PD (G G6PD). In individual animals, the ratio of G6PD phenotypes among committed erythroid progenitor cells (BFU-E) is always similar to that among committed granulocytic progenitor cells (CFU-GM)<sup>18-20</sup>. Female Safari cats are generally balanced heterozygotes, with on average, equal numbers of progenitor cells (and differentiated blood cells) of each parental phenotype<sup>20</sup>.

By observing the G6PD phenotype of progenitor cells, we

tracked the contributions of stem cells to the progenitor cell compartment. As shown in the representative study in Fig. 1a (cat 63848), the percentage of progenitors with d G6PD remained relatively constant during six years of repeated observations, and variation about its mean value was not significant with chi-square analysis. Therefore, hematopoiesis appeared polyclonal and stable, and little information could be derived concerning the behavior of individual stem cells<sup>18,19</sup>. In contrast, when six female Safari cats were lethally irradiated and their hematopoiesis was then reconstituted by a small number ( $1-2 \times 10^7$ /kg) of autologous nucleated marrow cells, the G6PD phenotype of progenitors varied extensively (Fig. 1b)<sup>19</sup>. Significant variation extended for 1-4½ years after transplantation when analyzed with successive chi-square analyses<sup>20</sup>. In addition, as seen in Fig. 1b, the pattern of clonal contributions to hematopoiesis in each cat was unique. Wide fluctuations in clonal contributions for 4½ years and subsequent support of hematopoiesis by one or few clones (as for cat 40005, Fig. 1b); wide fluctuations for one year then polyclonal hematopoiesis (for example, cat 40004, Fig. 1b); and more moderate initial fluctuations in clonal contributions (as for cat 40629, Fig. 1b) were observed.

To understand how stem cells might contribute to the progenitor cell compartment, and might result in these divergent outcomes in individual cats, we simulated hematopoiesis assuming that all decisions were determined by chance. Stochastic simulation facilitates the analysis of stem cell frequency and kinetics in this large-animal system, where the requirement for blood cell production is more similar to that of humans.

## Computer simulation

The concept of stochastic differentiation is diagrammed in Fig. 2. Aspirated marrow contains  $R_0$  quiescent hematopoietic stem cells (shown as four lettered cells). Rather than transplant these cells into a cat, the cells were "placed" in a computer where dif-

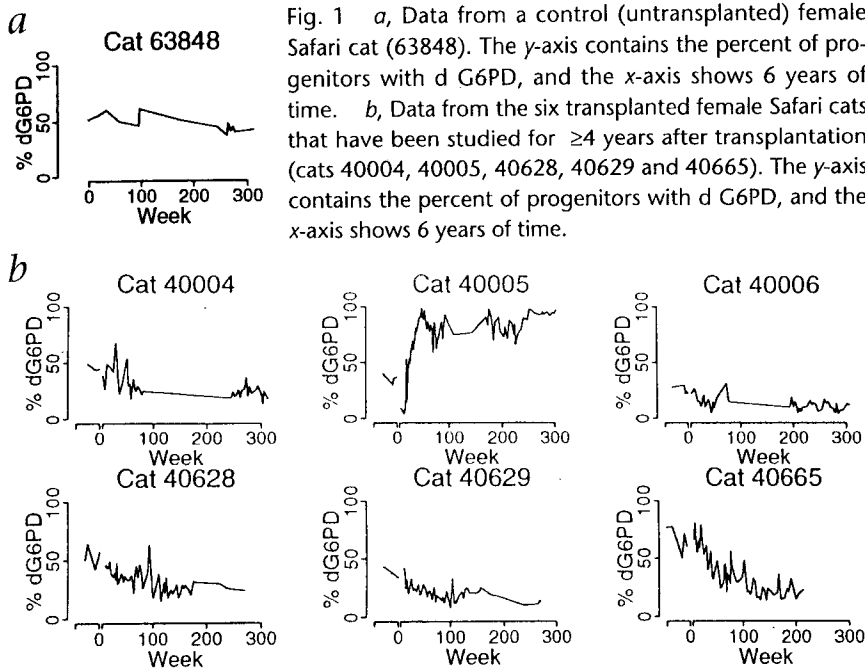


Fig. 1 *a*, Data from a control (untransplanted) female Safari cat (63848). The y-axis contains the percent of progenitors with d G6PD, and the x-axis shows 6 years of time. *b*, Data from the six transplanted female Safari cats that have been studied for  $\geq 4$  years after transplantation (cats 40004, 40005, 40628, 40629 and 40665). The y-axis contains the percent of progenitors with d G6PD, and the x-axis shows 6 years of time.

recorded. We tested different values for  $R_0$ , for  $C_0$  (the number of contributing stem cell clones immediately following transplantation), and for each intensity to determine whether any set of parameter values could lead to the spectrum of outcomes that we observed in female Safari cats.

In Fig. 3, 20 (of 100) sequential independent simulations of marrow transplantation are shown, beginning with 30 quiescent hematopoietic stem cells. Outcomes remarkably similar to those of cat 40005 (Fig. 3, *b1*, *c5*, and *d5*) and cat 40004 (Fig. 3, *b3* and *b5*) are seen, as well as outcomes resembling studies in other transplanted cats. Variation extends from 1.4 to 6.1 years in these simulations, and then subsides. At times, the subsequent hematopoiesis is dominated by a clone or clones with a single G6PD phenotype, at other times it is polyclonal, as observed in the female Safari cat transplantation studies.

If one uses the same parameter values as given in Fig. 3, but "transplants" 10, and not 30, quiescent hematopoietic stem cells, different

patterns are generated, as shown in Fig. 4. The variation in the G6PD phenotype of progenitor cells immediately following transplantation is extensive, and later in time, BFU-E and CFU-GM express a single G6PD phenotype far more frequently (in 9 out of 20 simulations) than was observed in the feline studies. Some cats (Fig. 4, *b1*, *d1* and *d2*) run out of stem cells. Taken together, these patterns are not similar to those seen in animal experiments.

Simulations of marrow transplantations with 100 hematopoietic stem cells are shown in Fig. 5. Early and late variability are indistinguishable. Comparable outcomes were seen with  $R_0 = 300$  and  $R_0 = 500$  (data not shown). Although the results do not reproduce the experimental data, this might be the predicted outcome if we were to transplant  $\geq 2 \times 10^8$  nucleated marrow cells per kilogram, the quantity of cells generally used in clinical therapies. Therefore, it is of interest that stable polyclonal hematopoiesis has been observed following allogeneic human marrow transplantation<sup>22,23</sup>. These paths are also similar to observations in control (untransplanted) cats (for example, cat 63848, Fig. 1*a*).

Simulation of marrow transplantations with 100 hematopoietic stem cells are shown in Fig. 5. Early and late variability are indistinguishable. Comparable outcomes were seen with  $R_0 = 300$  and  $R_0 = 500$  (data not shown). Although the results do not reproduce the experimental data, this might be the predicted outcome if we were to transplant  $\geq 2 \times 10^8$  nucleated marrow cells per kilogram, the quantity of cells generally used in clinical therapies. Therefore, it is of interest that stable polyclonal hematopoiesis has been observed following allogeneic human marrow transplantation<sup>22,23</sup>. These paths are also similar to observations in control (untransplanted) cats (for example, cat 63848, Fig. 1*a*).

**Criteria for analysis**

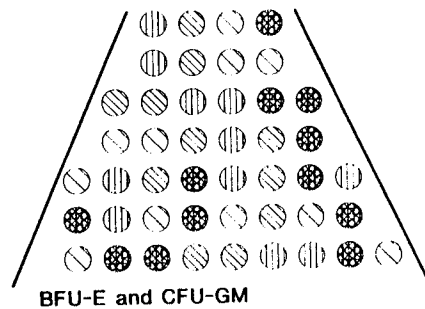
On the basis of the patterns observed in the transplanted female Safari cats, five criteria were derived to permit comparisons with simulated data. These required that a substantial variability in the G6PD phenotype of progenitor cells was present immediately following transplantation (criteria 2 and 3), which subsided (criterion 4) by approximately 1–4 years after transplantation (criterion 1) (see Methods for details). As complete blood cells, frequencies of marrow BFU-E and CFU-GM, as well as progenitor cell-cycle kinetics, returned to baseline values by 10 weeks after transplantation, these values remained normal during the 300 weeks of observations, and there was no evidence of graft failure<sup>20</sup>, the exhaustion of stem cells is not an expected outcome. Criterion 5 was considered to be violated if the sum of paths which "ran out" of stem cells, paths that were maintained entirely by clones expressing d G6PD plus paths that were maintained entirely by clones expressing G G6PD, exceeded 10%. Although  $\geq 90\%$  of progenitors from cats 40005 and 40006 contained a single parental G6PD phenotype after years

tor cell  
Fig. 1a  
5PD re-  
peated  
signifi-  
cantly  
derived  
contrast,  
and their  
r (1–2 ×  
) pheno-  
typic  
when an-  
i, as seen  
topoiesis  
contribu-  
toiesis by  
ations for  
at 40004,  
ial contri-  
the progen-  
divergent  
sis assum-  
Stochastic  
ity and ki-  
ement for  
as.  
med in Fig  
oietic stem  
plant these  
r where dif-

Reserve of stem cells  
 $R = 4$

Contributing stem cell clones  
 $C = 4$

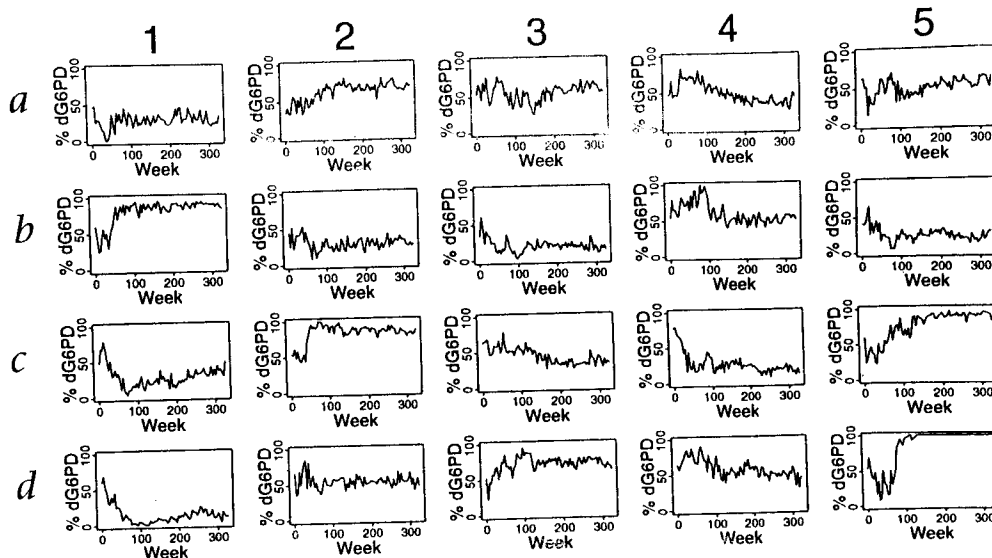
(a) (b) (c) (d)



Fates  
 $p(\text{replication})$   
 $p(\text{differentiation})$   
 $p(\text{apoptosis})$

$p(\text{extinction})$

Fig. 2 Stochastic description of hematopoiesis. Quiescent hematopoietic stem cells form a reserve compartment ( $R$ ). These cells may replicate, initiate a differentiation/maturation program, or undergo an apoptotic death. The probabilities of these events per cell per unit time, called intensities, are  $\lambda_r$ ,  $\mu_r$ , and  $\alpha_r$ , respectively. If a stem cell differentiates, its progeny form a clone, which contributes to hematopoiesis for a random period of time (the intensity for clonal extinction is  $\mu_c$ ). For transplantation studies, BFU-E and CFU-GM are sampled from the contributing compartment ( $C$ ). These compartment descriptions are similar to those proposed by others<sup>21</sup>.



4–5, exclusive contributions were not observed in the experimental animals.

### Parameter analysis

Using criteria 1–5 to assess the similarity of simulations to observed data, we derived feasible values for all parameters. For example, the intensity of stem cell replication  $\lambda_r$  determines how often self-renewal occurs, and consequently how fast the stem cell reserve expands. For simulations in Figs 3–5,  $\lambda_r = 1$  per 10 weeks. When  $\lambda_r \geq 1$  per 3 weeks, the reserve compartment grew too quickly and without the early variability required for criteria 2 and 3. If smaller values of  $R_0$ ,  $\lambda_r/\mu_R$  and/or  $\mu_R/\mu_C$  were used to satisfy criteria 2 and 3, we violated criterion 5. If  $\alpha_r$  were increased to satisfy criteria 2 and 3 and, in effect, to mandate asymmetric cell division<sup>24,25</sup>, criterion 5 was also violated.

When  $\lambda_r$  was small ( $\leq 1$  per 30 weeks), too few stem cell replications occurred during the course of the simulation, so that if the early variance was large enough to satisfy criteria 2 and 3,  $R$  remained too small to permit the small late variability required by criteria 1 and 4. This is illustrated in Fig. 6. Stem cell replication must therefore occur on average between every 3 and every 30 weeks to generate outcomes comparable to that observed *in vivo*.

Table 1 shows the bounds in range for other parameters ( $R_0$ ,  $C_0$ ,

$\lambda_r/\mu_R$ ,  $\mu_R/\mu_C$ ,  $\mu_R$ ,  $\mu_C$ ,  $\alpha_r$ ,  $N$ ) as derived by similar manipulations, as well as the values selected for the simulations in Fig. 3.

As shown in Table 1, the number of transplanted quiescent hematopoietic stem cells  $R_0 = 10$ –100. This implies that  $10$ – $100$  stem cells were present in the  $3$ – $8 \times 10^7$  transplanted nucleated marrow cells ( $1$ – $2 \times 10^7$  buffy coat cells per kilogram  $\times$   $3$ – $4$  kg/cat), yielding a frequency of quiescent hematopoietic stem cells between  $1$  per  $10^7$  and  $3$  per  $10^6$  nucleated marrow cells. The selected value of  $R_0 = 30$  (Table 1) can be thought of as an estimate of the average number of stem cells present per  $5 \times 10^7$  nucleated marrow cells (the average number of cells transplanted in the six Safari cats). This yields a frequency estimate of  $6$  per  $10^7$ . This estimate does not account for the loss of hematopoietic stem cells because of seeding efficiency, but it is probably more accurate than estimates of stem cell frequency based on *in vitro* properties alone<sup>26</sup>, or based on monoclonal reconstitution in single animals (for example, cat 40005, Fig. 1b) in which at least one hematopoietic stem cell must have been present among the  $8 \times 10^7$  nucleated marrow cells infused.

In contrast to  $R_0$ ,  $C_0$  is not an important determinant of outcome. Values between  $0$  and  $2000$  yielded equivalent results (Table 1). This is consistent with the observation that reconstitution is a property of hematopoietic stem cells, and not of differentiating stem cell clones<sup>15</sup>.

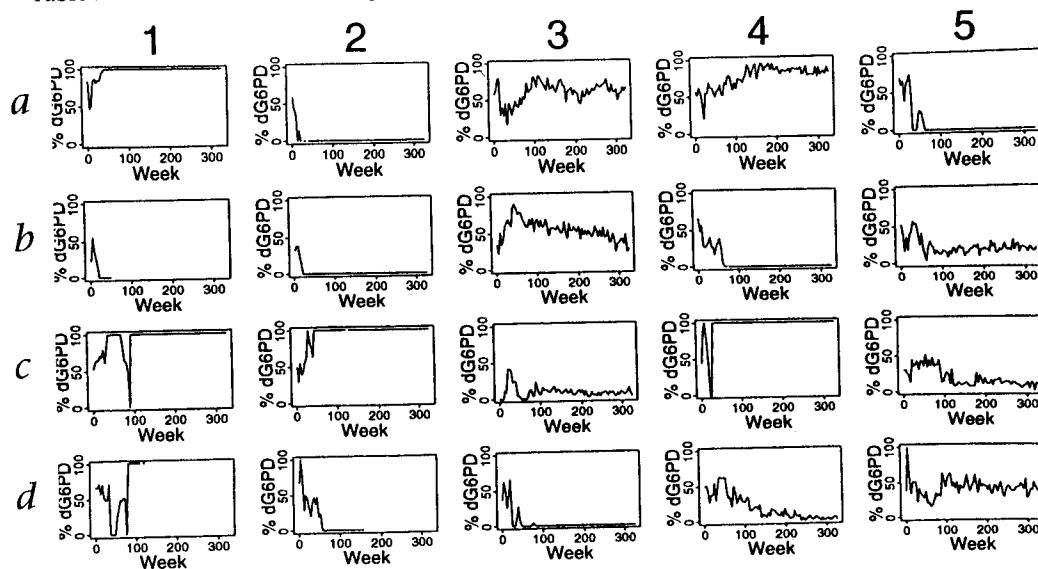


Fig. 4 Simulation of autologous transplantation with  $R_0 = 10$ . The other parameters are as in Fig. 3.

Fig. 3 Simulations of autologous transplantation of 30 quiescent hematopoietic stem cells ( $R_0 = 30$ ). Other parameters for these simulations are  $\lambda_r = 1$  per 10 weeks,  $\mu_R = 1$  per 12.5 weeks,  $\alpha_r = 0$ ,  $\mu_C = 1$  per 6.7 weeks. Equilibrium sizes for  $R$  and  $C$  are 750 and 400, respectively. As stem cells are harvested at a time of steady-state hematopoiesis (before radiation exposure),  $750/400 = R_0/C_0$  and  $C_0 = 16$ . The initial number of transplanted stem cells with d G6PD is chosen randomly from a Bin(30,0.5) distribution. The y-axis contains the percent of progenitors with d G6PD, and the x-axis shows 6 years of time.

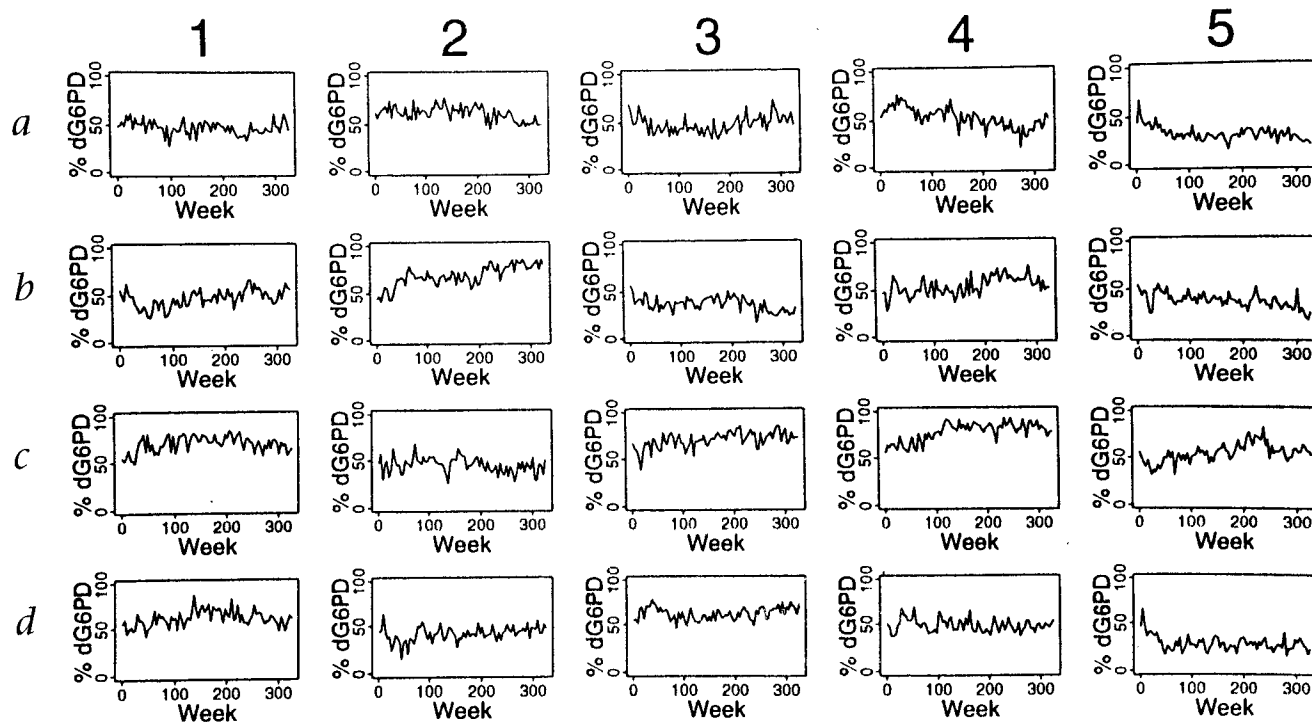


Fig. 5 Simulation of autologous transplantation with  $R_0 = 100$ . The other parameters are as in Fig. 3.

A last analysis relates to the maximum size  $N$  of the quiescent hematopoietic stem cell reserve:  $N$  cannot be less than 50, or simulations are rejected by using criteria 1 and 5, but an upper limit is not required (Table 1). Although  $N = 750$  was selected for computational convenience, there is also a biological relevance to a finite  $N$ . It is likely that an adult animal supports steady-state hematopoiesis by maintaining a reasonably constant number of hematopoietic stem cells. This may reflect the number of geographic niches within the marrow microenvironment in which stem cells can adhere and/or survive<sup>27</sup>. Therefore, after transplantation,  $R$  increases (because of stem cell self-renewal) until a steady state ( $R = N$ ) is achieved, but should not increase indefinitely. Once  $R = N$ , niches open only if there is stem cell

apoptosis or differentiation/maturation. An extension of this assumption that is inherent in our approach is that there are two mechanisms for stem cell death once  $R = N$ . First, there is spontaneous cell death or apoptosis, an intrinsic random event with intensity  $\alpha_R$  (bounds 0 and  $3 \times \mu_R$ , see Table 1). Second, should stem cell replication occur when no additional niche is vacant, one daughter cell will fail to survive.

**Discussion**

After autologous transplantation of small numbers of marrow cells in G6PD heterozygous (female Safari) cats, clonal contributions to hematopoiesis vary<sup>19</sup>. It appears that clones contribute and then stop, while other clones emerge. This disequilibrium

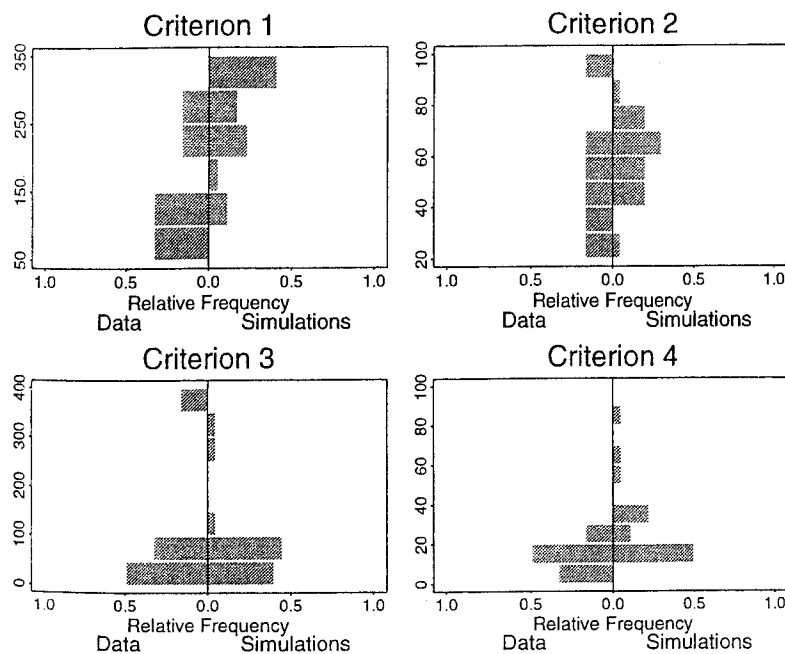


Fig. 6 A comparison of simulated and observed data. Back-to-back histogram of simulated characteristics (right-hand side) and observed characteristics (left-hand side) for criteria 1–4, for a simulation using  $R_0 = 20$ ,  $\lambda_R = 1$  per 30 weeks, and the other parameters as in Fig. 3. In all histograms the x-axis depicts relative frequency, whereas the y-axis corresponds to the respective criterion values. The binomial forward rejection point (criterion 1) has a distribution shifted upwards compared with that of the data. There is no significant difference between the distributions of overall range of variation (criterion 2), or of early variance (criterion 3), whereas the late variance of the simulations (criterion 4) is long-tailed compared with the distribution in the data. Parameters are rejected by using criteria 1 and 4. These parameter values constitute the best combination for  $\lambda_R = 1$  per 30 weeks.

ous cent 30). ula- = 1 r 6.7 nd C y. As ne of efore  $R_0/C_0$  er of G6PD n a y-axis nitors ows 6 ns, as escent 0–100 leated  $\times 3-4$  c stem 7 cells. f as an er  $5 \times$  ; trans- imate loss of ut it is quency onal re- Fig. 1b) ve beed. of out- results onstitu- not of autologous = 10. The ; in Fig. 3.

extends for 1–4.5 years<sup>20</sup>, and thus contrasts the short disequilibrium interval (2–6 months) seen in comparable murine studies<sup>28–36</sup>. These biological observations led us to hypothesize that the stem cell compartment reaches a steady state long after transplantation in cats, because of the relatively large size of the hematopoietic stem cell reserve<sup>20</sup>.

Hematopoietic stem cells cannot be isolated and observed. Therefore, further insights cannot be derived directly from experimental data. Stochastic simulation, however, allows one to study the behavior of cells in the unobserved stem cell compartment. This approach is feasible because all hematopoietic stem cells and their progeny are labeled (with either d or G G6PD), and absolute determinations of the G6PD phenotypes of progenitor cells (percentage with d G6PD s.e.m.) are obtained<sup>18,19</sup>.

By comparison of simulated outcomes to experimental results from the transplantation studies of female Safari cats, we estimate that the frequency of quiescent hematopoietic stem cells is approximately 6 per 10<sup>7</sup> nucleated marrow cells. Comparable studies of mice, however, estimate a higher concentration of hematopoietic stem cells (approximately 50–300 per 10<sup>7</sup> nucleated marrow cells<sup>28,37</sup>). This raises the possibility that the concentration of hematopoietic stem cells decreases with increasing animal size or longevity.

In the simulations of Fig. 3, hematopoietic stem cells replicated an average of once every 10 weeks. If self-renewal occurred more frequently than once every three weeks, we could not simulate the observed outcomes. As proviral integration requires stem cell replication<sup>38</sup>, these results could help to explain the difficulty of gene transfer into the hematopoietic stem cells of large animals<sup>39–42</sup>.

Taken together, these simulations provide insights into the behavior of hematopoietic stem cells. This approach should also provide sufficient flexibility to test hypotheses about stem cell kinetics and to predict outcomes of experimental studies, such as gene transfer experiments.

When we initially analyzed our transplantation data, we used a more traditional mathematical approach<sup>19,43,44</sup>. Such calculations required, however, simplifying assumptions. For example, we assumed that from 10 weeks after transplantation (a time when BFU-E and CFU-GM numbers and cell-cycle kinetics were normal<sup>19</sup>) onward, a steady state existed so that every time a stem cell clone stopped contributing to hematopoiesis, a new one was activated. Although some useful information was obtained, this model could not predict or explain the stable polyclonal hematopoiesis or the clonal dominance of hematopoiesis that developed in experimental animals one to four years after transplantation. Other investigators have used similar methods, such as feedback loops<sup>45,46</sup> or branching processes<sup>24,47</sup> to model hematologic data.

Stochastic simulation is a conceptually simple process. Multiple events occur independently and randomly. Information about steady-state kinetics can be derived from data, and need not be assumed *a priori*. The mathematical analysis is difficult only because of the need to score efficiently all replication, commitment/differentiation, and apoptotic decisions of hundreds to thousands of cells.

The stochastic process underlying the simulations also has a biological relevance. Each stem cell, located in a geographical niche within the marrow space, experiences unique inputs from its environment. Cytokines are produced by monocytes, T cells, fibroblasts, and endothelial cells; they are expressed on the cell surfaces of these microenvironmental cells; and they are at-

Table 1 Results of parameter analyses

Parameter	Minimum value	Maximum value	Selected value
$\lambda_r$	1 per 30 weeks	1 per 3 weeks	1 per 10 weeks
$\lambda_r/\mu_r$	1.1	2.0	1.25
$\mu_r/\mu_c$	0.1	2.0	0.533
$\mu_r$	1 per 60 weeks	1 per 3.3 weeks	1 per 12.5 weeks
$\mu_c$	1 per 120 weeks	1 per 0.33 weeks	1 per 6.7 weeks
$R_0$	10	100	30
$C_0$	0	2000	16
$N$	50	$\infty$	750
$\alpha_r$	0	$3 \times \mu_r$	0

Maximum and minimum bounds are presented, as well as the values used in the simulations of Fig. 3, which excellently approximate all Safari cat data.  $\lambda_r/\mu_r$  determines how quickly the reserve compartment grows.  $\mu_r/\mu_c$  determines how fast C increases relative to R. The number of transplanted quiescent hematopoietic stem cells,  $R_0$ , could be as low as 2 and satisfy criteria 1–4. However, all simulations using  $R_0 < 10$  violated criterion 5. With  $R_0 > 50$ , there was no path similar to that seen in cat 40005 (Fig. 1b), but outcomes equivalent to those of other animals were observed. With  $R_0 > 100$ , criteria 1 and 3 were clearly violated, and thus  $R_0 = 100$  is recorded as the bounds. The maximum size of the hematopoietic stem cell reserve is termed  $N$ .  $N$  cannot be less than 50 or criteria 1 and 5 are violated. There does not appear to be an upper limit on  $N$ . When  $\lambda_r$  is below the upper bound, simulations using  $N = 750$  were not noticeably different from those using a higher  $N$ . For computational convenience, we limited the number of stem cells  $N$  available in the simulated reserve, and selected  $N = 750$ . Including a positive  $\alpha_r$  in the simulation increased the variability of paths, by changing the behavior of the reserve compartment just as if we were to increase the differentiation intensity  $\mu_r$  by that amount. Hence, the growth in the reserve compartment became slower, as described by the ratio  $\lambda_r/(\alpha_r + \mu_r)$ . For a given growth rate, the contributing compartment filled up more slowly as  $\alpha_r$  is increased, causing more variability earlier in the simulation. When stimulating paths with  $\alpha_r + \mu_r$  held constant at the previously selected value for  $\mu_r$ , namely 1 per 12.5 weeks, there was no impact on the observed paths even using  $\alpha_r$  as high as  $3 \times \mu_r$ .

tached to matrix protein<sup>48</sup>. Redundant, synergistic, and contradictory factors are present<sup>48,49</sup>. Cell–cell interactions and the genetic programming inherent to individual cells also have an impact on self-renewal and differentiation decisions. A naive stem cell acts in a stochastic fashion (that is, responds to the sum of the divergent influences that it uniquely experiences). In this way, hematopoiesis as a process has similarities to the evolution of protein structure, as well as cosmic ray theory, some of the areas in which stochastic simulation has provided insights to complex scientific processes<sup>50–52</sup>.

This analysis, like any mathematical analysis, does not prove that the differentiation *in vivo* is stochastic. Rather, it proves that stochastic differentiation is sufficient to explain all patterns of clonal contributions to hematopoiesis seen in experimental animals. It also demonstrates that clonal dominance (such as was observed in cat 40005, Fig. 1b) can occur by chance, and does not require a clone with a specialized phenotype or a genetic advantage.

## Methods

**The Markov assumption.** A stochastic process is, for each instance  $t$

Table 2 Application of assessment criteria

Cat.	Years	Length	Crit. 1	Crit. 2	Crit. 3	Crit. 4
40004	6.0	44	273	55	59	21
40005	5.8	78	230	95	356	7
40006	5.9	52	73	27	29	13
40628	5.2	59	92	49	15	11
40629	5.2	54	103	35	23	8
40665	4.1	59	103	67	94	14

Assessment criteria applied to the data for the transplanted female Safari cats in Fig. 1b. The column "years" is the number of years of observation post-transplantation. Length refers to the number of observations 10 or more weeks following transplantation. The four rightmost columns are the values of criteria 1-4, namely the binomial forward rejection point, the overall range of variation, the early variance, and the late variance, respectively. The last four columns constitute the distributions to which the distributions of criterion values in the simulations were compared.

of time, a random variable (or vector), which takes values in some state space. The value of the process at a given time is its state at that time. The key assumption we make when describing the dynamics of hematopoiesis is the Markov assumption. This states that the future is influenced by the past only through the present. The stochastic process thus has only limited memory.

The probabilistic description of the stochastic process has three parts: one describing the dynamics of the stem cell reserve, one describing the behavior of differentiating stem cell clones, and one describing the data collection.

**Reserve compartment.** Let  $R_t = (R_t^d, R_t^c)$  be the random vector of quiescent stem cells with d G6PD and G G6PD phenotypes, respectively, available to the Safari cat at time  $t$ , this time being measured from the time of transplantation. The conditional probability of increase (by replication) or decrease (by differentiation or apoptosis) of the process within a small time increment  $h$  is assumed linear in  $h$ . More precisely, given a current state of  $R_t = r = (r^d, r^c)$ , we have

$$P(R_{t+h}^d = r^d + 1 | R_t = r) \approx \begin{cases} \lambda_r^d r^d h & \text{if } r^d + r^c < N \\ 0 & \text{if } r^d + r^c \geq N \end{cases}$$

$$P(R_{t+h}^d = r^d - 1 | R_t = r) \approx (\mu_r^d + \alpha_r^d) r^d h$$

$$P(R_{t+h}^c = r^c + 1 | R_t = r) \approx \begin{cases} \lambda_r^c r^c h & \text{if } r^d + r^c < N \\ 0 & \text{if } r^d + r^c \geq N \end{cases}$$

$$P(R_{t+h}^c = r^c - 1 | R_t = r) \approx (\mu_r^c + \alpha_r^c) r^c h$$

The parameters  $\lambda_r^d$  and  $\lambda_r^c$  are the replication intensities for stem cells containing d G6PD and G G6PD, respectively, whereas the parameters  $\mu_r^d$  and  $\mu_r^c$  are the differentiation intensities, and the parameters  $\alpha_r^d$  and  $\alpha_r^c$  are the apoptosis intensities. The probability that any of the components of  $R_t$  change by more than one unit in a vanishingly small time interval  $h$  is negligible, as is the probability that both components change in the same interval.

As a consequence of the Markov assumption, the time to the next replication or, say, a stem cell with d G6PD, is an exponential random variable with mean  $1/\lambda_r^d$ . The time to differentiation or to apoptosis is likewise exponentially distributed random variables.

**Contributing compartment.** As shown in Fig. 2, after stem cells ini-

tiate their differentiation program, further cell division leads to committed progenitor cells, and then to mature blood cells. The contributing compartment consists of these "active" stem cell clones. When a quiescent stem cell differentiates, it enters into this compartment. The current state at time  $t$  is written  $C_t = (C_t^d, C_t^c)$  and again conditional probabilities of one-step change during a small time increment  $h$  are linear in  $h$ . We have

$$P(C_{t+h}^d = c^d + 1 | C_t = c, R_t = r) \approx \mu_r^d r^d h$$

$$P(C_{t+h}^d = c^d - 1 | C_t = c, R_t = r) \approx \mu_c^d c^d h$$

$$P(C_{t+h}^c = c^c + 1 | C_t = c, R_t = r) \approx \mu_r^c r^c h$$

$$P(C_{t+h}^c = c^c - 1 | C_t = c, R_t = r) \approx \mu_c^c c^c h$$

A stem cell clone leaves the contributing compartment when its numbers are depleted. The parameters  $\mu_c^d$  and  $\mu_c^c$  are the depletion intensities. As before, the time until depletion of a stem cell clone will be an exponential random variable with the mean the inverse depletion intensity.

As it appears that G6PD phenotype is a neutral label, in that the behavior of cells containing d G6PD is equivalent to that of those containing G G6PD (ref. 18), we make the simplifying assumption that intensities for cells containing d G6PD and G G6PD are equal.

**Observations.** Provided that all clones in the contributing compartment contribute equally to hematopoiesis<sup>19,44</sup>, the distribution of G6PD phenotypes among sampled cells (BFU-E and CFU-GM) should be proportional to this distribution among active stem cell clones. Hence we assume that our observations are binomial samples (because of the large number of cells being sampled from, we can ignore the finiteness of the population) from the contributing compartment. Formally, if we observe  $n_i$  colonies at time  $t$ ,  $Y_i$  of which contain d G6PD, we have

$$Y_i \sim \text{Bin}(n_i, C_t^d / (C_t^d + C_t^c)).$$

The resulting process can be termed a hidden Markov process, where the hidden process is the four-dimensional process  $(R_t^d, R_t^c, C_t^d, C_t^c)$  which is partially revealed only through binomial observation as described above.

**Equilibrium behavior.** When the reserve compartment is entirely reconstructed, its numbers will remain relatively stable. Thus the total intensity with which clones enter the contributing compartment is approximately  $Np\mu_r$  for stem cells with d G6PD, where  $p$  is the proportion of cells containing d G6PD in the reserve compartment. Because there is a large number of cells in the reserve once this compartment is full, this proportion will change only very slowly. Similarly, the intensity of entering stem cells with G G6PD is  $N(1-p)\mu_r$ . This allows us to look at the contributing compartment in equilibrium as what is called an immigration-death process. The equilibrium distribution of such a process is a Poisson distribution, so that the number  $C^d$  of clones of cells with d G6PD is distributed  $\text{Po}(Np\mu_r/\mu_c)$  in equilibrium, whereas the number  $C^c$  of clones of cells with G G6PD is distributed  $\text{Po}(N(1-p)\mu_r/\mu_c)$ . We can redefine the parameters using  $\phi = N\mu_r/\mu_c$  and  $p$  as unknown parameters. Although  $N$ ,  $\mu_r$  and  $\mu_c$  are all unknown, they cannot be separated out from equilibrium considerations.

Given the total number  $C^d + C^c$  of clones in the contributing compartment (which, on average, is  $\phi$ ), the number  $C^d$  is distributed  $\text{Bin}(C^d + C^c, p)$ . Hence each observation  $Y_i$  is binomially distributed with success probability  $C^d / (C^d + C^c) \approx p$  when the process is in equilibrium.

**Simulation.** By the Markov assumption, the time to the next apoptosis, differentiation, or replication in a reserve compartment of size  $R$  is the minimum of six independent exponential random variables with three for each G6PD type. We use pseudo-random exponential random numbers to simulate the reserve compartment. If  $R = N$  and the next event is a replication, it is ignored. Once a stem cell enters the contributing compartment to head a clone of differentiating cells, we simulate another independent exponential random variable with mean  $1/\mu_c$ , which indicates how long this clone will persist. The binomial sampling is done every 4 weeks using samples of size 70 (the average size of the samples from the experimental cats). In this way simulated outcomes incorporate the variability that results from biological sampling. To initiate simulations,  $R_0$  hematopoietic stem cells were randomly selected (harvested) from a population containing equal numbers of stem cells with d G6PD and with G G6PD. Using ratios of 0.23–0.67 (as observed in baseline determinations in the experimental animals) did not alter conclusions. Except for simulations intending to study the effect of  $C_0$  on the resulting paths, we took  $C_0$  to be  $R_0\mu_r/\mu_c$ , because at the time of marrow harvest (that is, before radiation exposure) the animals had normal steady-state hematopoiesis. For each analysis, 20–100 independent simulations were performed using the S-Plus statistical system, Ver. 3.1-2 (Statistical Sciences Inc., Seattle, Washington).

#### Criteria for comparison of simulated and observed data.

**Criterion 1.** The time after transplantation when/if the variation in the percent of progenitors with d G6PD subsides (termed the forward binomial variation rejection point). For this determination we computed a standard chi-square statistic of proportions equality based first on the complete data set, and subsequently by leaving out the first observation, the first two observations, and so on. The week listed is the last week for which the  $P$  value for the data set consisting of that week through the final observation was less than 0.05.

**Criterion 2.** The overall range of variation. This was expressed as the difference between the maximum percentage of progenitors with d G6PD observed and the minimum percentage.

**Criterion 3.** The relative amount of variation in the percent of progenitors with d G6PD immediately following transplantation. This was measured by Pearson's binomial goodness-of-fit statistic of the first 15 observations beginning week 10.

**Criterion 4.** The relative amount of variation in the percentage of progenitors with d G6PD late after transplantation. This measurement assessed the last 15 observations.

**Criterion 5.** The percentage of simulations in which hematopoiesis was exclusively supported by cells with a single G6PD phenotype, or where the stem cell reserve was exhausted.

Variability was assessed beginning at week 10 after transplantation to assure that contribution of stem cells, rather than more differentiated cells also present in the marrow inoculum, were assayed. The first four criteria applied to data from the six transplanted female Safari cats are given in Table 2. The rejection of a set of parameters is always a strong statement indicating a substantial inadequacy of the stochastic process in simulating observed data. There is considerable interaction between the parameters, and it was often necessary to change several parameters simultaneously when investigating allowable ranges. We assessed criteria 1–4 by comparing the observed distribution (from six cats) of each criterion to the distribution of the same criterion in the simulated paths. Criterion 5 was considered to be violated if greater than 10% of the paths were maintained only by clones expressing d G6PD, were maintained only by clones expressing G

G6PD, or died out because of an exhaustion of stem cells within the reserve compartment.

#### Acknowledgments

This work was supported by R01 HL 46598 from the National Institutes of Health. J.L.A. is a recipient of a Faculty Research Award from the American Cancer Society. The authors thank Allan Dimanahan for help in preparation of this manuscript.

RECEIVED 19 MAY; ACCEPTED 20 DECEMBER 1995

- Till, J.E., McCulloch, E.A. & Siminovitch, L. A stochastic model of stem cell proliferation, based on the growth of spleen colony-forming cells. *Proc. Natl. Acad. Sci. USA* 51, 29–36 (1964).
- Ogawa, M., Porter, P.N. & Nakahata, T. Renewal and commitment to differentiation of hematopoietic stem cells (an interpretive review). *Blood* 61, 823–829 (1983).
- Gordon, M.Y. & Amos, T.A.S. Stochastic effects in hematopoiesis. *Stem Cells* 12, 175–179 (1994).
- Fairbairn, L.J., Cowling, G.J., Reipert, B.M. & Dexter, T.M. Suppression of apoptosis allows differentiation and development of a multipotent hemopoietic cell line in the absence of added growth factors. *Cell* 74, 823–832 (1993).
- Jordan, C.T., McKearn, J.P. & Lemischka, I.R. Cellular and developmental properties of fetal hematopoietic stem cells. *Cell* 61, 952–963 (1990).
- Jones, R.J., Wagner, J.E., Celano, P., Zicha, M.S. & Sharkis, S.J. Separation of pluripotent haematopoietic stem cells from spleen colony-forming cells. *Nature* 347, 188–189 (1990).
- Terstappen, L.W.M.M., Huang, S., Safford, M., Lansdorp, P.M. & Loken, M.R. Sequential generations of hematopoietic colonies derived from single nonlineage-committed CD34<sup>+</sup> CD34<sup>-</sup> progenitor cells. *Blood* 77, 1218–1227 (1991).
- Keifer, F., Wagner, E.F. & Keller, G. Fractionation of mouse bone marrow by adherence separates primitive hematopoietic stem cells from *in vitro* colony-forming cells. *Blood* 78, 2577–2582 (1991).
- Ogata, H. *et al.* Separation of hematopoietic stem cells into two populations and their characterization. *Blood* 80, 91–95 (1992).
- Uchida, N. & Weissman, I.L. Searching for hematopoietic stem cells: Evidence that Thy-1.1<sup>+</sup> Lin<sup>-</sup> Sca-1<sup>+</sup> cells are the only stem cells in C7SBL/Ka-Thy-1.1 bone marrow. *J. Exp. Med.* 175, 175–184 (1992).
- Baum, C.M., Weissman, I.L., Tsukamoto, A.S., Buckle, A.-M. & Peault, B. Isolation of a candidate human hematopoietic stem cell population. *Proc. Natl. Acad. Sci. USA* 89, 2804–2808 (1992).
- Spain, L.M. & Mulligan, R.C. Purification and characterization of retrovirally transduced hematopoietic stem cells. *Proc. Natl. Acad. Sci. USA* 89, 3790–3794 (1992).
- Onishi, M. *et al.* CD4<sup>int</sup> hematopoietic progenitor cells in murine bone marrow. *Blood* 81, 3217–3225 (1993).
- Wolf, N.S., Kona, A., Priestley, G.V. & Bartalmez, S.H. *In vivo* and *in vitro* characterization of long-term repopulating primitive hematopoietic stem cells isolated by sequential Hoechst 33342-rhodamine 123 FACS selections. *Exp. Hematol.* 21, 514–22 (1993).
- Uchida, N., Aguila, H.L., Fleming, W. H., Jerabek, L. & Weissman, I.L. Rapid and sustained hematopoietic recovery in lethally irradiated mice transplanted with purified Thy-1.1<sup>+</sup> Lin<sup>-</sup> Sca-1<sup>+</sup> hematopoietic stem cells. *Blood* 83, 3358–3379 (1994).
- Orlic, D. & Bodine, D.M. What defines a pluripotent hematopoietic stem cell (PHSC): Will the real PHSC please stand up! *Blood* 84, 3991–3994 (1994).
- O'Brien, S.J. Molecular genetics in the domestic cat and its relatives. *Trends Genet.* 2, 137–142 (1986).
- Abkowitz, J.L. *et al.* Feline G-6-PD cellular mosaicism: Application to the study of retrovirus-induced pure red cell aplasia. *J. Clin. Invest.* 75, 133–140 (1985).
- Abkowitz, J.L., Linenberger, M.L., Newton, M.F., Ott, R.L. & Guttorp, P. Evidence for the maintenance of hematopoiesis in a large animal by the sequential activation of stem cell clones. *Proc. Natl. Acad. Sci. USA* 87, 9062–9066 (1990).
- Abkowitz, J.L. *et al.* The behavior of hematopoietic stem cells in a large animal. *Proc. Natl. Acad. Sci. USA* 92, 2031–2035 (1995).
- Uchida, N., Fleming, W.H., Alpern, E.J. & Weissman, I.L. Heterogeneity of hematopoietic stem cells. *Curr. Opin. Immunol.* 5, 177–184 (1993).
- Nash, R., Storb, R. & Neiman, P. Polyclonal reconstitution of human marrow after allogeneic bone marrow transplantation. *Blood* 72, 2031–2037 (1988).
- Turhan, A.A., Humphries, R.K., Phillips, G.L., Eaves, A.C. & Eaves, C.J. Clonal hematopoiesis demonstrated by X-linked DNA polymorphisms after allogeneic bone marrow transplantation. *N. Engl. J. Med.* 320, 1655–1661 (1989).
- Vogel, H., Niewisch, H. & Mاتيoli, G. Stochastic development of stem cells. *J. Theor. Biol.* 22, 249–270 (1969).
- Posakony, J.W. Nature versus nurture: Asymmetric cell divisions in *Drosophila* bristle development. *Cell* 76, 415–418 (1994).
- Berardi, A.C., Wang, A., Levine, J.D., Lopez, P. & Scadden, D.T. Functional isolation and characterization of human hematopoietic stem cells. *Science* 267, 104–107 (1995).
- Stewart, F.M., Crittenden, R.B., Lowry, P.A., Person-White, S. & Quesenberry, P.

- ithin  
of  
ican  
ell pro-  
l. Acad.  
ferenti-  
23-829  
cells 12,  
of apoptic cell  
al prop-  
ation of  
s. Nature  
en, M.R.  
nonlin-  
991).  
urrow by  
colony-  
ulations  
Evidence  
-1.1 bone  
'eault, B.  
'roc. Natl.  
rovirally  
790-3794  
one mar-  
vitro char-  
1 cells iso-  
ions. Exp.  
I.L. Rapid  
nsplanted  
Blood 83,  
c stem cell  
94).  
ves. Trends  
o the study  
(1985).  
Guttorp, P.  
l by the se-  
9062-9066  
rge animal.  
ogeneity of  
an marrow  
(1988).  
C.J. Clonal  
r allogeneic  
9).  
stem cells. J.  
n *Drosophila*  
nctional iso-  
Science 267,  
esenberry, P.
- J. Long-term engraftment of normal and post-5-fluorouracil murine marrow into normal nonmyeloablated mice. *Blood* 81, 2566-2571 (1993).
28. Harrison, D.E., Astle, C.M. & Lerner, C. Number and continuous proliferative patterns of transplanted primitive immunohematopoietic stem cells. *Proc. Natl. Acad. Sci. USA* 85, 822-826 (1988).
29. Dick, J.E., Magli, M.C., Huszar, D., Phillips, R.A. & Bernstein, A. Introduction of a selectable gene into primitive stem cells capable of long-term reconstitution of the hematopoietic system of *W/W<sup>v</sup>* mice. *Cell* 42, 71-79 (1985).
30. Lemischka, I.R., Raulet, D.H. & Mulligan, R.C. Developmental potential and dynamic behavior of hematopoietic stem cells. *Cell* 45, 917-927 (1986).
31. Capel, B., Hawley, R., Covarrubias, L., Hawley, T. & Mintz, B. Clonal contributions of small numbers of retrovirally marked hematopoietic stem cells engrafted in unirradiated neonatal *W/W<sup>v</sup>* mice. *Proc. Natl. Acad. Sci. USA* 86, 4564-4568 (1989).
32. Snodgrass, R. & Keller, G. Clonal fluctuation within the haematopoietic system of mice reconstituted with retrovirus-infected stem cells. *EMBO J.* 6, 3955-3960 (1987).
33. Jordan, C.T. & Lemischka, I.R. Clonal and systemic analysis of long-term hematopoiesis in the mouse. *Genes Dev.* 4, 220-232 (1990).
34. Keller, G. & Snodgrass, R. Life span of multipotential hematopoietic stem cells *in vivo*. *J. Exp. Med.* 171, 1407-1418 (1990).
35. Capel, B., Hawley, R.G. & Mintz, B. Long- and short-lived murine hematopoietic stem cell clones individually identified with retroviral integration markers. *Blood* 75, 2267-2270 (1990).
36. Harrison, D.E. & Zhong, R.-K. The same exhaustible multilineage precursor produces both myeloid and lymphoid cells as early as 34 weeks after marrow transplantation. *Proc. Natl. Acad. Sci. USA* 89, 10134-10138 (1992).
37. Micklem, H.S., Lennon, J.E., Ansell, J.D. & Gray, R.A. Numbers and dispersion of repopulating hematopoietic cell clones in radiation chimeras as functions of injected cell dose. *Exp. Hematol.* 15, 251-257 (1987).
38. Coffin, J.M. Retroviridae and their replication. (eds Fields, B.N. & Knipe, D.M.) *Virology*. 1437-1500 (Raven Press, New York, 1990).
39. Karlsson, S. Treatment of genetic defects in hematopoietic cell function by gene transfer. *Blood* 78, 2481-2492 (1991).
40. Lothrop, C.D., Jr., et al. Expression of a foreign gene in cats reconstituted with retroviral vector infected autologous bone marrow. *Blood* 78, 2372-2375 (1991).
41. Bodine, N.M. et al. Long-term *in vivo* expression of a murine adenosine deaminase gene in Rhesus monkey hematopoietic cells of multiple lineages after retroviral mediated gene transfer into CD34<sup>+</sup> bone marrow cells. *Blood* 82, 1875-1890 (1993).
42. Kiem, H.-P. et al. Retrovirus-mediated gene transduction into canine peripheral blood repopulating cells. *Blood* 83, 1467-1473 (1994).
43. Guttorp, P., Newton, M.A. & Abkowitz, J.L. A stochastic model for haematopoiesis in cats. *IMA J. Math. Appl. Med. Biol.* 7, 125-143 (1990).
44. Newton, M., Guttorp, P., Catlin, S.N., Assuncao, R. & Abkowitz, J.L. Stochastic modeling of early hematopoiesis. *J. Am. Stat. Assoc.* (in the press).
45. Wichmann, H.E., Loeffler, M. & Schmitz, S. A concept of hematopoietic regulation and its biomathematical realization. *Blood Cells* 14, 411-429 (1988).
46. Blackett, N.M. Haematopoietic spleen colony growth: A versatile, parsimonious, predictive model. *Cell Tissue Kinet.* 20, 393-402 (1987).
47. Kurnit, D.M., Matthyse, S., Papayannopoulou, P. & Stamatoyannopoulos, G. Stochastic branching model for hematopoietic progenitor cell differentiation. *J. Cell. Physiol.* 123, 55-63 (1985).
48. Hogge, D.E. et al. Cytokines acting early in human hematopoiesis. *Baillière's Clin. Haematol.* 7, 49-63 (1994).
49. Metcalf, D. Hematopoietic regulators: Redundancy or subtlety? *Blood* 82, 3515-3523 (1993).
50. Heilinga, H.W. & Richards, F.M. Optimal sequence selection in proteins of known structure by simulated evolution. *Proc. Natl. Acad. Sci. USA* 91, 5803-5807 (1994).
51. Guttorp, P. *Stochastic Modeling of Scientific Data*. (Chapman and Hall, London, 1995).
52. Arley, N. *On the Theory of Stochastic Processes and their Application to the Theory of Cosmic Radiation*. (Wiley, New York, 1943).