Constrained optimization of divisional load in hierarchically organized tissues during homeostasis

Peter Ashcroft and Sebastian Bonhoeffer

Institute of Integrative Biology, ETH Zurich, Zurich, Switzerland

It has been hypothesized that the structure of tissues and the hierarchy of differentiation from stem cell to terminally differentiated cell play a significant role in reducing the incidence of cancer in that tissue. One specific mechanism by which this risk can be reduced is by minimizing the number of divisions—and hence the mutational risk—that cells accumulate as they divide to maintain tissue homeostasis. Here, we investigate a mathematical model of cell division in a hierarchical tissue, calculating and minimizing the divisional load while constraining parameters such that homeostasis is maintained. We show that the minimal divisional load is achieved by binary division trees with progenitor cells incapable of self-renewal. Contrary to the protection hypothesis, we find that an increased stem cell turnover can lead to lower divisional load. Furthermore, we find that the optimal tissue structure depends on the time horizon of the duration of homeostasis, with faster stem cell division favoured in short-lived organisms and more progenitor compartments favoured in longer-lived organisms.

1. Introduction

Many tissues in the human body undergo constant regeneration of their constituent cells. This regeneration allows the tissue to maintain effective function and to prevent damage accumulation, for example from exposure to acidic environments in the gut or ultraviolet light on the skin. In organs such as the colon, skin and haematopoietic system, the regeneration is managed through a hierarchical organization of stem and progenitor cells. These cells undergo amplification of numbers and phenotypic differentiation to produce the functional, terminally differentiated cells (TDC) in large enough quantities to support the functioning of the tissue. The haematopoietic system is the archetypal example of such a tissue: in humans approximately $10^5$ haematopoietic stem cells [1] are ultimately responsible for producing approximately $10^{11}$ TDC per day to satisfy respiratory, immune and coagulation demands [2–4].

It has been hypothesized that hierarchical population structures have evolved to protect against the accumulation of mutant cells [5–11]. Empirical observations give weight to these claims: despite blood cells being over 10 times more frequent than any other human cell type [12], haematopoietic cancers make up only a small fraction of cancer incidence statistics [13]. It has been suggested that the protection against cancer is conferred by limiting the cumulative number of cell divisions needed to maintain tissue homeostasis [6], as well as ensuring that the majority of cells are transient such that continued cell turnover washes out mutants [5,8].

Mathematical models in which cells are grouped into homogeneous compartments (so-called compartmental models) have been employed to test different aspects of these cancer prevention hypotheses. With such models it is possible to calculate the number of divisions that cells undergo as they differentiate from stem cells to TDC [6,10,14–17]. The mutation risk is inherently
linked to this number of divisions; every DNA replication event carries a risk of mistakes, such that more divisions means a higher risk of genetic damage. Once an expression is obtained for the number of cell divisions across the tissue, it is then possible to optimize the tissue architecture in order to minimize the number of divisions and the risk of mutation accumulation [6,14].

Using a computational model, Pepper et al. [9] showed that increasing the number of compartments in hierarchically organized tissues limits the somatic evolution (accumulation of multiple beneficial mutations) of those cells. Similarly, Alvarado et al. [14] focussed optimizing tissue structures to limit mutation accumulation. They found that the optimal tissue structure depends on the objective of optimization: minimizing the number of cells with a single mutation in the tissue requires a high degree of self-renewal across the compartments, while the time until two-hit mutants are generated is maximized by the non-self-renewing binary division tree.

2. Methods
2.1. Model structure
We consider a compartmental model for cell proliferation in a hierarchical population structure. For the purpose of our analyses, we primarily consider a reduced model in which only symmetric division events can occur. The full model, which considers symmetric and asymmetric cell divisions, as well as cell death, is described in the electronic supplementary material. This structure is highlighted in figure 1a. We note that we only consider a linear population structure, with population branching here ignored to reduce the number of model parameters. In the full model, we additionally consider a compartment of non-dividing (quiescent) SCs, labelled \( i = 0 \). See appendix I in the electronic supplementary material for details.

2.2. Dynamics
Cells move between compartments through differentiation, which is strictly coupled to cell division in this model. A division event produces two daughter cells, each of which may remain in the same compartment as the mother cell (\( i \)), or differentiate to the next compartment along the hierarchy (\( i + 1 \)). We do not consider the possibility of de-differentiation (i.e. cells cannot move from compartment \( i \) to compartment \( i - 1 \)). De-differentiation would lead to an increase in divisions accumulated during homeostasis compared to a forward-only differentiating tissue.

To build intuition and guide analytical calculations, we primarily consider the reduced model (figure 1b) in which cells only divide symmetrically. We ignore cell death and asymmetric cell division, as well as SC quiescence compared to the full model that is described in appendix I in the electronic supplementary material. This reduced model is parametrized by \( \beta_i \) which is the division rate of cells in compartment \( i \), and \( a_i \) which is their probability of self-renewal (correspondingly, \( 1 - a_i \) is the probability of differentiation) following a division event. The reduced model is described by the set of ordinary differential equations

\[
\dot{N}_i = 2(1 - a_i \beta_i) N_i - (1 - 2a_i \beta_i) N_{i+1} \quad \text{for} \quad i > 0,
\]

where the first term describes the inflow of two daughter cells from compartment \( i - 1 \) and the second term is the net loss rate (differentiation minus self-renewal) of cells from compartment \( i \). Furthermore, we set \( a_0 = 0 \) and the parameter \( \beta_i \) assumes the role of the death rate of the TDCs.

Under homeostasis the number of cells per compartment does not change (\( N_i = 0 \)). The steady state of equation (2.1),
which we write as \( N'_i \), satisfies the relation
\[
(1 - 2\alpha_i)\beta_i N'_i = 2(1 - \alpha_i)\beta_i N_i - 1, \tag{2.2}
\]
i.e. the net loss from compartment \( i \) is balanced by the influx from differentiating cells in compartment \( i - 1 \). For this steady state to exist, we require \( \alpha_1 = 0.5 \) and \( \alpha_i < 0.5 \) for \( 1 < i < n \). Solving equation (2.2) recursively, we have
\[
N'_i = \frac{\beta_i}{a_i} \left[ \prod_{k=1}^{i-1} \frac{2(1 - \alpha_k)}{1 - 2\alpha_k} \right] N'_i.
\]
(2.3)

In particular, for TDCs (\( i = n \)), we have
\[
N'_n = \frac{\beta_i}{a_i} \left[ \prod_{k=1}^{n-1} \frac{2(1 - \alpha_k)}{1 - 2\alpha_k} \right] N'_i, \tag{2.4}
\]

2.3. Counting divisions

For compartmental models, the expected number of divisions that cells accumulate in a tissue has previously been calculated using two methods: firstly, Derényi & Szöllősi [6] constructed ordinary differential equations (ODE) for the cumulative number of divisions across all cells in a compartment, and from this they extracted the average number of divisions that cells undergo during their lifetime. Alternatively, Böttcher et al. [15] constructed an explicit set of ODEs for sub-populations of cells of type \( i \) that have undergone \( j \) division events, from which the mean number of divisions per cell is easily extracted. These models yield identical results for the mean number of divisions in the reduced and full models, as described in appendices II and III in the electronic supplementary material, respectively.

Briefly, following the method of Derényi & Szöllősi [6], the cumulative number of divisions accumulated across all SCs, \( T' \), in the reduced model satisfies the ODE
\[
T'_1 = \alpha_1 \beta_1 N'_1 \frac{T'_1 + 2}{N'_1} - (1 - \alpha_1) \beta_1 N'_1 \frac{T'_1}{N'_1} = \beta_1 N'_1, \tag{2.5}
\]
where we have used \( \alpha_1 = 0.5 \). Together with the initial condition \( T'(0) = 0 \), equation (2.5) can be solved to give \( T'(t) = \beta_1 N'_1 t \).

Hence, the average number of divisions per SC at time \( t \), \( D(t) \), satisfies
\[
D(t) = \frac{T'(t)}{N'_1} = \beta_1 t, \tag{2.6}
\]
i.e. SC divisions are accumulated linearly with time.

The derivation of the equations for \( T'_i \) [similar to equation (2.5)] for all compartments \( 1 \leq i \leq n \) is described in appendix II in the electronic supplementary material. Ultimately, the expected number of divisions that cells accumulate after a time period \( t \), \( D(t) \), is described by a set of linear ODEs which can be solved analytically. However, the resulting expression becomes increasingly complicated for large \( n \). Instead, we here make the assumption that the SC dynamics are much slower than non-SCs. We can then express the mean number of divisions per TDC at time \( t \), \( D_n(t) \), as the sum of the SC divisions up to that point in time (equation (2.6)) plus the expected number of divisions that are accumulated after SC differentiation, \( D_n \). The latter term can be calculated (see appendix II in the electronic supplementary material) as
\[
D_n = \sum_{j=2}^{n} \frac{1}{1 - 2\alpha_j}. \tag{2.7}
\]

This equation is equivalent to the result of Derényi & Szöllősi [6] for the reduced dynamics. The calculation differs from Derényi & Szöllősi [6] in the optimization procedure in the next steps.

2.4. Optimizing tissue architecture

Given the expressions for the expected number of divisions accumulated per TDC, we can minimize these values—subject to physical constraints—to deduce the optimal tissue structure parameters which minimize the divisional load. The constraint that we apply is that during homeostasis the tissue should maintain a given number of TDCs, such that the tissue remains functional. For the reduced model, we derive the constraint from the equilibrium TDC population size, equation (2.4), which we rearrange as
\[
\sum_{i=2}^{n} \frac{2(1 - \alpha_i)}{1 - 2\alpha_i} \beta_i N'_i = 0. \tag{2.9}
\]

The parameters \( \alpha_i \) (\( 1 < i < n \)), \( \beta_i \), \( N'_i \) and \( n \) must satisfy this relation to ensure that enough TDCs are present in the tissue. We keep the number of TDCs (\( N'_i \)) and the death rate of TDCs (\( \beta_i \)) fixed, as well as setting \( \alpha_i = 0.5 \) and \( a_i = 0 \) as required in the model definition.

We then use the method of Lagrange multipliers [22] to constrain the homeostatic number of TDCs and minimize the tissue’s divisional load. The objective function that is to be optimized is the number of divisions accumulated per TDC, which we can construct in one of two ways: firstly, we can consider only the divisions accumulated after a SC has differentiated, \( D_n \) (equation (2.7)). In this scenario, we drop the time dependence and only focus on optimizing the differentiation structure of the tissue. The Lagrangian function is then constructed as the linear combination of equations (2.7) and (2.9), i.e.
\[
\mathcal{L} = \sum_{i=2}^{n} \frac{1}{1 - 2\alpha_i} - \lambda \left[ \sum_{i=2}^{n} \frac{2(1 - \alpha_i)}{1 - 2\alpha_i} \frac{\beta_i N'_i}{\beta_i N'_1} \right], \tag{2.10}
\]
where the coefficient \( \lambda \) is the Lagrange multiplier.

Secondly, we can include SC divisional history and ask how many divisions have TDCs accumulated after a given time \( t \), \( D_n(t) \), using the approximation in equation (2.8) as the objective function. The Lagrangian function in this case is
\[
\mathcal{L}(t) = \beta_1 t + \sum_{i=2}^{n} \frac{1}{1 - 2\alpha_i} - \lambda \left[ \sum_{i=2}^{n} \frac{2(1 - \alpha_i)}{1 - 2\alpha_i} \frac{\beta_i N'_i}{\beta_i N'_1} \right]. \tag{2.11}
\]

The optimum tissue architecture is then found by computing the stationary points of \( \mathcal{L} \) or \( \mathcal{L}(t) \) considered as a function of \( a_i \) (\( 1 < i < n \)), \( \beta_i \), and the Lagrange multiplier \( \lambda \). Due to the discrete nature of \( n \), we cannot identify its optimum value through stationary point analysis. We can, however, identify it through graphical methods. The number of SCs does not appear in the objective function in the Lagrangian, so this too must be optimized by graphical methods.

The Lagrangian functions for the full model can be constructed analogously. See appendix V in the electronic supplementary material for details.

3. Results

Although our results are algebraic, we illustrate these results using the observed parameters of human erythropoiesis (red
blood cell production). A healthy human has approximately $10^{13}$ red blood cells at any given time [12] and these cells have a lifetime of approximately 120 days [3]. The most recent estimates of human haematopoietic stem cell (HSC) numbers are approximately $10^8$ cells which divide approximately once per year [1].

3.1. Optimized tissue architectures in the reduced model

3.1.1. Additional divisions after stem cell differentiation

To determine the architecture that minimizes the divisional load, one needs to find the stationary points of the Lagrangian function in equation (2.10). These stationary points satisfy $\partial L/\partial a_i = 0$ (for $1 < i < n$) and $\partial L/\partial \lambda = 0$. As the SC division rate $\beta_i$ features only in the constraint and not in the objective function, this parameter cannot be optimized through stationary point analysis. The values of the $a_i$ at the stationary point of $L$ are the parameters which minimize the number of divisions accumulated by the TDCs after leaving the SC compartment, while still maintaining the homeostatic population of TDCs.

From the $\partial L/\partial a_i = 0$ (for $1 < i < n$) condition, we arrive at

$$2(1 - a_i) \over 1 - 2a_i = \lambda \prod_{k=2}^{n-1} \frac{2(1 - a_k)}{1 - 2a_k},$$

which is only satisfied if

$$\frac{2(1 - a_i)}{1 - 2a_i} = \lambda^{1/(3-n)}.$$  \hspace{1cm} (3.2)

This implies that all $a_i$ are equal for $1 < i < n$ in a tissue which minimizes the divisional load after SC differentiation. This equivalence of the $a_i$ across compartments was used by Böttcher et al. [15] and Alvarado et al. [14] as a simplifying assumption without derivation. Furthermore, this result is consistent with the result of Derényi & Szollosi [6], who found that the ratio of differentiation output between two consecutive compartments is constant in optimized tissues.

The value of $\lambda$, and ultimately the value of the $a_i$, are obtained from the $\partial L/\partial \lambda = 0$ condition, which is equal to the population size constraint (equation (2.9)). By substituting equation (3.2) into equation (2.9) and then solving for $\lambda$, we find the value of the Lagrange multiplier

$$\lambda = \left( \frac{\beta_i N_i^c}{\beta_1 N_1^c} \right)^{(3-n)/(n-2)}.$$ \hspace{1cm} (3.3)

By rearranging equation (3.2), we can now define the values of $a_i$ which minimize the number of divisions per TDC during homeostasis and for a given number of compartments. These values will depend explicitly on the number of compartment size $n$, hence we write this value as $\tilde{a}(n)$, which satisfies

$$\tilde{a}(n) = \frac{\lambda^{1/(3-n)} - 2}{2\lambda^{1/(3-n)} - 2} = \left( \frac{\beta_i N_i^c}{\beta_1 N_1^c} \right)^{1/(n-2)} - 2.$$ \hspace{1cm} (3.4)

We here want to differentiate between the subscript notation, which is just the index of the compartment, and the bracket notation, denoting that the optimized parameter value $\tilde{a}(n)$ depends on the number of compartments in the hierarchy.

From equation (3.4), we see that an increase in the number of compartments $n$ must be compensated by a decrease in the progenitor self-renewal probability $\tilde{a}(n)$ to maintain the homeostatic level of TDCs. However, $\tilde{a}(n)$ is a probability and is bounded from below by zero. Once $n$ is increased sufficiently such that $(\beta_i N_i^c/\beta_1 N_1^c)^{1/n-2} < 2$, we have to artificially set $\tilde{a}(n) = 0$ and there will be a new (increased) steady state TDC population size $N_i$. From equation (2.4), the number of TDCs will satisfy

$$N_i = \frac{\beta_i N_i^c}{\beta_1} \left( \frac{2(1 - \tilde{a}(n))}{1 - 2\tilde{a}(n)} \right)^{-n-2},$$ \hspace{1cm} (3.5)

while the number of divisions per TDC after SC differentiation (equation (2.7)) satisfies

$$\tilde{D}_i = (n - 2) \frac{1}{1 - 2\tilde{a}(n)} + 1.$$ \hspace{1cm} (3.6)

The minimum number of divisions occurs when $\tilde{a}(n)$ first reaches zero, i.e. when $(\beta_i N_i^c/\beta_1 N_1^c)^{1/(n-2)} = 2$. At this point, the number of compartments is given by

$$n = \log_2 \left( \frac{\beta_i N_i^c}{\beta_1 N_1^c} \right) + 2.$$ \hspace{1cm} (3.7)

and the divisional load per TDC is

$$\tilde{D}_i = \log_2 \left( \frac{\beta_i N_i^c}{\beta_1 N_1^c} \right) + 1 = n - 1.$$ \hspace{1cm} (3.8)

Therefore, perhaps unsurprisingly, the binary division tree $(\tilde{a}(n) = 0)$ with minimal number of compartments $(n = \bar{n})$ is the tissue structure which minimizes divisional load while maintaining the TDC population. As shown by Böttcher et al. [15], both the mean and variance of the cumulative number of cell divisions is minimal for this non-self-renewing structure. If the number of compartments $n$ is smaller than $\bar{n}$, then there must be some self-renewal in the progenitor compartments, which leads to an increase in the cumulative number of divisions. If $n > \bar{n}$, then there are more divisions than required to maintain the TDC population and the tissue produces more TDCs than necessary.

The values of $\tilde{a}(n)$ (probability of self-renewal that minimizes divisional load per TDC), $N_i$ (the corresponding number of TDCs) and $\tilde{D}_i$ (the divisional load per TDC) are illustrated in figure 2a–c as a function of total SC turnover $(\beta_i N_i^c)$ and the number of compartments in the tissue $(n)$. The line described by equation (3.7) (dashed diagonal) does indeed follow the minimum divisional load; any change in either SC turnover or the number of compartments leads to an increase in divisional load. However, by moving along this line (equation (3.7)), in particular to higher SC turnover and lower number of compartments, the divisional load can be reduced, as shown in figure 2f. Therefore, the tissues which minimize divisional load after SC differentiation would have very high SC turnover and few differentiation steps.

If the SC turnover is fixed for another reason (e.g. due to spatial and nutritional niche competition), then the minimum divisional load occurs when there are enough compartments such that there is no progenitor self-renewal $(\tilde{a}(n) = 0)$, but not too many such that TDCs are not overproduced $(N_i = N_i^c)$ (figure 2f).
3.1.2. Cumulative divisions after lifetime $t$

The objective function in the Lagrangian $\mathcal{L}(t)$ (equation (2.11)), which now counts SC divisions too, is explicitly dependent on the SC division rate parameter $\beta_i$, as well as on the organism’s lifetime. We can now determine the parameters and tissue structure which minimize the divisional load after a time period $\tau$, which we call the timescale of optimization. The optimum parameter values must satisfy $\partial \mathcal{L}(\tau)/\partial \alpha_i = 0$ ($1 \leq i \leq n$), $\partial \mathcal{L}(\tau)/\partial \lambda = 0$, as well as $\partial \mathcal{L}(\tau)/\partial \beta_i = 0$. Note that $\tau$ is the timescale of optimization: it is entirely possible that the organism lives for a shorter ($t < \tau$) or longer ($t > \tau$) time and we therefore have to consider two timescales.

The partial derivative of $\mathcal{L}(\tau)$ (equation (2.11)) with respect to $\alpha_i$ are unchanged from the previous section, such that the condition equation (3.2) must be satisfied. From the condition $\partial \mathcal{L}(\tau)/\partial \beta_i = 0$, we arrive at

$$\beta_i = \sqrt{\frac{\lambda \beta_i N_i}{\tau N_1}}.$$  
(3.9)

Note that the SC division rate $\beta_i$ is now decoupled from the SC population size $N_1$ for this time-dependent problem, and therefore we can no longer consider SC turnover $\beta_i N_1$ as an independently variable quantity as was done in figure 2.

These values of $\alpha_i$ (equation (3.2)) and $\beta_i$ (equation (3.9)) can be substituted into the population size constraint equation (2.9), i.e. the $\partial \mathcal{L}(\tau)/\partial \lambda = 0$ condition, and we can solve for the value of the Lagrange multiplier $\lambda$:

$$\lambda = \left(\frac{\beta_i N_i^3}{\tau N_1}\right)^{3-n/(n-1)}.$$  
(3.10)

The values of $\alpha_i$ and $\beta_i$, which minimize the number of divisions per TDC after time $\tau$ are thus given by

$$\alpha_i(\tau) = \hat{\alpha}(\tau) = \frac{(\tau \beta_i N_i^3 / N_1)^{1/(n-1)} - 2}{2(\tau \beta_i N_i^3 / N_1)^{1/(n-1)} - 2}$$  
and

$$\beta_i(\tau) = \frac{1}{\tau} \left(\frac{\tau \beta_i N_i^3}{N_1}\right)^{1/(n-1)}.$$  
(3.11)

To satisfy $\hat{\alpha}(\tau) \geq 0$, we require $(\tau \beta_i N_i^3 / N_1)^{1/(n-1)} \geq 2$. This condition imposes $\beta_i(\tau) \geq 2/\tau$. Again, here we want to clarify the difference between the subscript notation, which is just the index of the compartment, and the bracket notation, denoting that the optimized parameter values $\hat{\alpha}(\tau)$ and $\beta_i(\tau)$ depend on the number of compartments in the hierarchy.
From equation (2.4), the number of TDCs will satisfy

$$N_n = \frac{\hat{\beta}_1(n) N_1}{\hat{\beta}_n} \left( \frac{2(1 - \alpha(n))}{1 - 2\alpha(n)} \right)^{n-2},$$

while the number of divisions per TDC after time $t$ (equation (2.8)) satisfies

$$D_n(t) = \hat{\beta}_1(n) t + (n - 2) \frac{1}{1 - 2\alpha(n)} + 1. \quad (3.13)$$

The minimum number of divisions occurs when $\alpha(n) = 0$ and $\beta_1(n) = 2/\tau$, or when $(\tau \beta_1 N_1^0 / N_1)^{1/\tau} = 2$. At this point, the number of compartments is given by

$$\bar{n} = \log_2 \left( \frac{\tau \beta_1 N_1^0}{N_1} \right) + 1, \quad (3.14)$$

and the divisional load per TDC is

$$\bar{D}_n = 2 \frac{t}{\tau} + \bar{n} - 1. \quad (3.15)$$

Firstly, we consider the optimized parameters and divisional load when the lifetime is equal to the optimization timescale ($t = \tau = 70$ years). For this case, the values of $\alpha(n)$, $\beta_1(n)$, $N_1^0$, and $\bar{D}_n$ are illustrated in figure 3a–d as a function of SC number ($N_1^0$) and the number of compartments in the tissue ($n$). The line described by equation (3.14) (dashed diagonal) again follows the minimum divisional load; any change in either SC population size or the number of compartments leads to an increase in divisional load. However, by moving along this line (equation (3.14), in particular to higher SC population size and lower number of compartments, the divisional load per TDC can be reduced, as shown in figure 3f. Therefore, the tissues which minimize TDC divisional load have many SCs and few differentiation steps. The optimum SC division rate $\beta_1(n)$ is substantially lower than the observed once per year [1], with SCs dividing twice per lifetime in the optimized tissue.

For a fixed SC population size, the SC division rate is a decreasing function of the number of compartments, such that shorter hierarchies require more input from the SC compartment (figure 3f). The minimum divisional load occurs when there are enough compartments such that there is no progenitor self-renewal ($\alpha(n) = 0$), but not too many such that TDCs are not overproduced ($N_n = N_1^0$), and when SCs are dividing as slowly as possible (figure 3e–h).

The optimum progenitor self-renewal rate is $\hat{\alpha}(\bar{n}) = 0$. However, the optimum SC division rate $\hat{\beta}_1(\bar{n})$, as well as the optimum number of compartments in the hierarchy, $\bar{n}$, depend on the timescale of optimization $\tau$. As this timescale increases, the optimum SC activity decreases ($\hat{\beta}_1(\bar{n}) = 2/\tau$; figure 4a) and the number of amplification steps required to maintain homeostasis increases ($\bar{n} = \log_2 (\tau \beta_1 N_1^0 / N_1) + 1$; figure 4b). These properties are independent of the lifetime of the individual. However, the lifetime $t$ does contribute to the accumulated divisional load (equation (3.15)). When we fix $t = \tau$, we have $D_n(t) = \bar{n}(t) + 1$, which increases logarithmically with $t$ (blue line in figure 4c). If the lifetime $t$ exceeds the optimization timescale $\tau$ (for example when an
individual lives beyond reproductive senescence), then TDCs accumulate additional divisions at a rate of $2/\tau$ per year (the SC division rate). For short optimization timescales, this means a rapid accumulation of additional divisions (orange lines in figure 4c).

This rapid accumulation of divisions can be seen more clearly by plotting $D_0(t)$ as a function of lifetime $t$ for different optimization timescales $\tau$ (figure 5). We also see that choosing an optimization timescale $\tau$ that exceeds the lifetime $t$ results in more accumulated divisions than necessary in early life, due to the increased number of differentiation steps.

3.2. Optimized architectures in the full model

For the full model (see appendix I in the electronic supplementary material for parameter descriptions), including quiescence and asymmetric division, we can repeat the above analyses using the Lagrangian function

$$L = \sum_{j=2}^{n-1} \left( \frac{2b_j + c_j}{b_j - a_j} - 1 \right) + 1 - \lambda \left[ \prod_{k=2}^{n-1} \frac{b_k + c_k}{b_k - a_k} \right] - \frac{d_iN_i^N}{(2b_1 + c_1)(\gamma/(\ell + \gamma))N_{SC}^N},$$

(3.16)

for optimizing the divisional load after SC differentiation, and

$$L(t) = \left( 2b_1 + c_1 \right) \frac{\gamma}{\ell + \gamma} + L_{\tau},$$

(3.17)

for optimizing divisional load after a lifetime $t$. Note that cell death would have to be compensated by additional cell divisions and will result in an increase in the average number of divisions of the surviving cells. We have therefore immediately set the death rate $d_i = 0$ for $1 \leq i < n$. We have also used the condition that SC self-renewal and SC differentiation occur at the same rate, $a_1 = b_1$, which is required for homeostasis.

The optimization process is described in full in appendix V in the electronic supplementary material. From this analysis, we find that the ratio

$$\phi_i = \frac{2b_i + c_i}{b_i - a_i},$$

(3.18)

is conserved across all progenitor compartments in optimized tissues, i.e. $\phi_i = \phi(n)$ for $1 < i < n$. Because of the conditions $b_i - a_i > 0$, as well as $a_i \geq 0$, $b_i \geq 0$, and $c_i \geq 0$, this ratio is bounded by $2 \leq \phi(n) < \infty$. When considering only divisions accumulated after SC differentiation (i.e. without time dependence), this ratio takes the value

$$\phi(n) = \left( \frac{d_iN_i^N}{(2b_1 + c_1)(\gamma/(\ell + \gamma))N_{SC}^N} \right)^{1/(n-2)},$$

(3.19)
and the divisional load per TDC is
\[ D_n = (n - 2)\frac{\phi(n)}{N_{\text{SC}}} - 1 + 1 \]  
(3.20)

This divisional load is minimized when \( \phi(n) = 2 \), which can only be achieved if \( a_i = c_i = 0 \), i.e. progenitor cells \((1 < i < n)\) do not self-renew or divide asymmetrically, they only differentiate. This pattern was also described by Derényi & Szölösi [6]. By increasing the SC differentiation flux \((2b_1 + c_1)(\tau f(\tau + 1))\) and constraining the number of compartments while maintaining the TDC population, hence reducing the divisional load (as described for the reduced model).

When we optimize the tissue architecture based on the total TDC divisional load including SC divisions, we have to include an optimization timescale \( \tau \). From our analysis (appendix V in the electronic supplementary material), we find the following relations for the ratio \( \phi(n) \) and the total SC division rate
\[ \phi(n) = \left( \frac{\tau d_{n} N_{n}^i}{N_{\text{SC}}} \right)^{(1/(n-1))} \quad \text{and} \quad \left(2b_1 + c_1\right) \frac{\tau f(\tau + 1)}{\tau} \left( \frac{\tau d_{n} N_{n}^i}{N_{\text{SC}}} \right)^{(1/(n-1))}, \]
and the divisional load after a lifetime \( t \) satisfies
\[ D_n(t) = (2b_1 + c_1) \frac{\tau f(\tau + 1)}{\tau} \left( n - 2 \right) \left( \frac{\tau d_{n} N_{n}^i}{N_{\text{SC}}} \right)^{(1/(n-1))} + 1. \]  
(3.22)

From the relation for the total SC division rate, we observe that there is a trade-off between the division rate of active SCs \((2b_1 + c_1)\) and the fraction of actively dividing SCs \( f(\tau + 1) \); if active SCs are dividing faster, then the number of actively dividing cells can be smaller. The total SC division rate is bounded by \( 2/\tau \leq (2b_1 + c_1)(\tau f(\tau + 1)) \leq \infty \), thus unconstrained by our optimization procedure. Briefly, if we assume that progenitor cell compartments do not affect the divisional load and are therefore not considered in our optimization procedure, then we would expect that in longer-lived animals having a lower rate of somatic mutation accumulation [23]. Furthermore, we find that underestimating the timescale of optimization relative to lifetime leads to a large increase in divisional load at ages beyond the optimization timescale.

Our full model of cell dynamics also allows asymmetric division, cell death and SC quiescence. Repeating the above analysis, we find that the ratio of differentiation output to net loss in a progenitor compartment is conserved across progenitor compartments in optimized tissues. This is a more general principle than derived by Derényi & Szölösi [6] who found that the ratio of differentiation output between two consecutive progenitor compartments is constant in optimized tissues. Furthermore, the minimum divisional load is achieved when the differentiation output to net loss ratio is equal to two, i.e. when progenitor cells do not self-renew and differentiate only symmetrically. This is something that could be compared to empirical data to check for tissue optimization, if such data existed. Finally, from our model and analysis, we find that there is a trade-off between the SC division rate and the fraction of quiescent SCs: faster dividing SCs would require a larger quiescent fraction to maintain the minimum divisional load (for a given number of progenitor compartments).

Interestingly, although optimal tissues have a constant self-renewal across progenitor cells, the division rates of the progenitor cells do not affect the divisional load and are thus unconstrained by our optimization procedure. Briefly, the progenitor compartment size \( N_{1} \) is inversely proportional to the division rate \( \tau \) (equation (2.3)), so the total differentiation flux out of that compartment \([1 - a_i] / \beta \) is independent of \( \beta \). Therefore, the relationship between compartment sizes in these optimized tissues are unconstrained as well: i.e. it is not necessary to have a monotonically increasing compartment size for the tissue to minimize divisional load. In optimized tissue structures (binary division trees), we do have a doubling of the flux of cells between consecutive progenitor compartments. By assuming a constant division rate across progenitor compartments, we arrive at the doubling of subsequent progenitor compartment sizes in these optimized tissues.

Our conclusions are in agreement with those of Derényi & Szölösi [6], despite different optimization procedures: we both find that hierarchically organized tissues can be tuned to minimize divisional loads, and therefore limit the occurrence of somatic mutations during homeostasis. Beyond...
Derényi & Szöllősi [6], we show that the constant ratio of differentiation outputs between subsequent compartments can be generalized further: we find that the ratio of differentiation output to net loss per compartment is conserved across all compartments, which in the absence of asymmetric divisions means that self-renewal probability is constant across all compartments. Furthermore, we explore the impacts of varying the optimization timescale, lifetime, SC numbers and SC quiescence.

Although the number of divisions itself is an important measure for damage accumulation per cell, it is not the full story in terms of cancer prevention. Intuitively, it is not clear if the total number of divisions per TDC, or only the number of divisions in the SC compartment, is important for cancer initiation and incidence. Tomasetti & Vogelstein have shown that across tissues the dependence of cancer incidence on the number of SC divisions is sub-linear, such that an increase in SC divisions does not lead to a proportional increase in cancer risk [24,25]. Therefore, other factors are at play, which could be extrinsic (e.g. smoking), or related to an increase in cancer risk. Tomasetti & Vogelstein have shown that not only the mean, but also the variance, increases with increasing progenitor self-renewal, leading to fewer cells with very high divisional load. This again suggests that the non-self-renewing tissue would be the optimum for minimizing the probability of a cell evolving an oncogenic phenotype through mutation accumulation.

Our analysis relies on assumptions about the system being in steady state and parameter values being constant (for example, constant HSC activation and deactivation throughout life). This, however, is not representative of early development when SCs undergo significant expansion and tissue compartments are populated. An expansion phase would require an initially increased SC self-renewal rate to allow the SCs to reach the homeostatic level and to populate the progenitor and TDC compartments.

With this study, we have shown that limiting progenitor self-renewal in hierarchical tissues results in the lowest divisional load during homeostatic tissue maintenance. This optimized tissue also satisfies some other somatic evolution constraints put forward by multiple authors. Going forward, we would like to bring all of these aspects together to achieve an overall understanding of the selection pressures that act on and shape tissue organization.

References

1. Lee-Six H et al. 2018 Population dynamics of normal human blood inferred from somatic mutations. Nature 561, 473–478. (doi:10.1038/s41586-018-0497-0)
2. Kaushansky K, Lichtman MA, Prchal JT, Levi M, Press OW, Burns LJ, Caligiuri MA. 2016 Williams hematology, 9th edn. New York, NY: McGraw-Hill.
3. Nombela-Arieta C, Manz MG. 2017 Quantification and three-dimensional microanatomical organization of the bone marrow. Blood Adv., 1, 407–416. (doi:10.1182/bloodadvances.2016003194)

Data accessibility. Code is publicly available at https://github.com/ashcrofttp/optimising-divisions and archived at https://doi.org/10.5281/zenodo.5554123.

Authors’ contributions. P.A. and S.B. conceptualized the study, validated the results and wrote the manuscript. P.A. performed formal analysis, wrote and curated the code and visualized the results. S.B.: conceptualization, data curation, formal analysis, investigation, methodology, resources, software, validation, visualization, writing—original draft, writing—review and editing; S.B.: conceptualization, funding acquisition, project administration, supervision, validation, writing—original draft, writing—review and editing.

Competing interests. We declare we have no competing interests.

Funding. This work was supported by the Swiss National Science Foundation (grant no. 310030B_176401; awarded to S.B.).
4. Vaziri H, Dragowska W, Allsopp RC, Thomas TE, Harley CB, Lansdorp PM. 1994 Evidence for a mitotic clock in human hematopoietic stem cells: loss of telomeric DNA with age. Proc. Natl Acad. Sci. USA 91, 9857–9860. (doi:10.1073/pnas.91.21.9857)

5. Cairns J. 1975 Mutation selection and the natural history of cancer. Nature 255, 197–200. (doi:10.1038/255197a0)

6. Derényi I, Szöllősi GJ. 2017 Hierarchical tissue organization as a general mechanism to limit the accumulation of somatic mutations. Nat. Commun. 8, 14545. (doi:10.1038/ncomms14545)

7. Grajzel D, Derényi I, Szöllősi GJ. 2020 A compartment size-dependent selective threshold limits mutation accumulation in hierarchical tissues. Proc. Natl Acad. Sci. USA 117, 1606–1611. (doi:10.1073/pnas.1913104117)

8. Nowak MA, Michor F, Iwasa Y. 2003 The linear process of somatic evolution. Proc. Natl Acad. Sci. USA 100, 14 966–14 969. (doi:10.1073/pnas.1913104117)

9. Pepper JW, Sprouffske K, Maley CC. 2007 Animal cell differentiation patterns suppress somatic evolution. PLoS Comput. Biol. 3, e250. (doi:10.1371/journal.pcbi.0030250)

10. Rodríguez-Brenes IA, Wodarz D, Komarova NL. 2013 Minimizing the risk of cancer: tissue architecture and cellular replication limits. J. R. Soc. Interface 10, 20130410. (doi:10.1098/rsif.2013.0410)

11. Werner B, Dingli D, Traulsen A. 2013 A deterministic model for the occurrence and dynamics of multiple mutations in hierarchically organized tissues. J. R. Soc. Interface 10, 20130349. (doi:10.1098/rsif.2013.0349)

12. Sender R, Fuchs S, Milo R. 2016 Revised estimates for the number of human and bacteria cells in the body. PLoS Biol. 14, e1002533. (doi:10.1371/journal.pbio.1002533)

13. Cancer Research UK. 2015 Leukaemia (All Subtypes Combined) Incidence Statistics. Cancer Research UK. See http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/leukaemia/incidence.

14. Adams PD, Jasper H, Rudolph KL. 2015 Aging-induced stem cell mutations as drivers for disease and cancer. Cell Stem Cell 16, 601–612. (doi:10.1016/j.stem.2015.05.002)

15. Tomasetti C, Vogelstein B. 2015 Variation in cancer risk among tissues can be explained by the number of stem cell divisions. Science 355, 1330–1334. (doi:10.1126/science.aaf9011)

16. Tomasetti C, Vogelstein B. 2015 Variation in cancer risk among tissues can be explained by the number of stem cell divisions. Science 347, 78–81. (doi:10.1126/science.1260825)