

Aging, clonal hematopoiesis and preleukemia: not just bad luck?

Liran I. Shlush^{1,2} · Sasan Zandi¹ · Shalev Itzkovitz² · Andre C. Schuh¹

Received: 6 August 2015 / Revised: 9 September 2015 / Accepted: 18 September 2015 / Published online: 6 October 2015
© The Japanese Society of Hematology 2015

Abstract Chronological human aging is associated with a number of changes in the hematopoietic system, occurring at many levels from stem to mature cells, and the marrow microenvironment as well. This review will focus mainly on the aging of hematopoietic stem and progenitor cells (HSPCs), and on the associated increases in the incidence of hematological malignancies. HSPCs manifest reduced function and acquire molecular changes with chronological aging. Furthermore, while for many years it has been known that the human hematopoietic system becomes increasingly clonal with chronological aging (clonal hematopoiesis), only in the last few years has it become clear that clonal hematopoiesis may result from the accumulation of preleukemic mutations in HSPCs. Such mutations confer a selective advantage that leads to clonal hematopoiesis, and that may occasionally result in the development of leukemia, and define the existence of both preleukemic stem cells, and of ‘preleukemia’ as a clinical entity. While it is well appreciated that clonal hematopoiesis is very common in the elderly, several questions remain unanswered: why and how does clonal hematopoiesis develop? How is clonal hematopoiesis related to the age-related changes observed in the hematopoietic system? And why do only some individuals with clonal hematopoiesis develop leukemia?

Keywords Aging · Clonal hematopoiesis · Preleukemia

Introduction

The incidence of most myeloid hematological malignancies increases with chronological age. It has been suggested that this increasing incidence (also observed in other tissues) is related to the random accumulation of mutations in replicating stem cells [1]. Specifically, it has been proposed that greater the replicative history of a stem cell, the larger will be the corresponding number of mutations accumulated, resulting in higher chance of malignancy. This hypothesis is consistent with a number of recent observations: first, several lines of evidence have shown that acute myeloid leukemia (AML) arises from ancestral preleukemic stem cells (preL-HSCs) [2–5] that already carry a substantial portion (up to 50 % in some cases) of the mutations found later in leukemic blasts [6]. However, despite carrying a spectrum of such preleukemic mutations (pLMs), preL-HSCs are nevertheless functional hematopoietic stem or progenitor cells (HSPCs), as defined by seemingly normal differentiation capacity (Box 1). The accumulation of mutations over time is thus associated with the development of AML. Second, recent exome sequencing studies have demonstrated the progressive appearance of clonal hematopoiesis (Box 1) with aging in a large proportion of healthy individuals (2–20 %) [7–12]. Notably, in the majority of such instances of clonal hematopoiesis, the major clone identified carried mutations which had previously been found in hematological malignancies, and specifically those that comprised the preleukemic mutations carried by preL-HSCs. Taken together, these observations suggest a ‘bad luck’ model, by which replicating HSPCs acquire pLMs over time, thereby gaining a selective advantage that leads to clonal hematopoiesis, and ultimately culminating in overt malignancy in some individuals.

✉ Liran I. Shlush
liranshlush3@gmail.com

¹ Department of Medical Oncology and Hematology, Princess Margaret Cancer Centre, University Health Network (UHN), 610 University Ave, Toronto, ON M5G 2M9, Canada

² Weizmann Institute of Science, Rehovot, Israel

While a relationship among stem cell replication, aging, and the development of cancer seems biologically plausible in light of the stem cell origin of myeloid malignancies as defined above, many open questions remain unanswered: do mutations accumulate randomly in HSPCs? How do pLMs provide HSPCs with a selective advantage, and why does this appear to occur mainly in the elderly? Is this a cell-intrinsic or extrinsic phenomenon? Do clonal hematopoiesis and pLMs contribute to other features of the aging immune/hematopoietic system? Why do only a very small proportion of individuals with clonal hematopoiesis bearing pLMs develop leukemia in their lifetime? In this review we will endeavor to answer some of these questions. Specifically, we will review the current understanding of aging of the human hematopoietic system, including clonal hematopoiesis of aging, and will explore the emerging data regarding pLMs, trying to understand how in some cases they might increase the probability of developing leukemia.

HSPC aging

The relationship between aging and the incidence of cancer was documented more than half a century ago, by Nordling and Armitage [13, 14]. Recent work has identified functional human stem cells as the cell of origin not only in AML [2, 5], but probably also in other hematological malignancies (reviewed by Shlush and Minden [6]). Other recent data have suggested that the lifetime risk of developing cancer correlates with the number of normal stem cell divisions in a particular tissue [1]. Taken together, and considering that it is only a relatively small number of HSPCs which maintains hematopoiesis for the lifetime of an individual, these data suggest that it may not be unexpected for HSPCs biology to vary with aging, and occasionally to give rise to leukemia. Acknowledging that HSPCs are the unit of aging responsible for the excess of myeloid malignancies in older age [15, 16] is the first step in understanding the sequential steps that lead to leukemia. Accordingly, understanding HSPC aging might help us understand why leukemia evolves. We will first review the various features of the aging HSPCs.

Reduced self-renewal capacity

Human studies suggested that although the total number of immunophenotypically defined HSPCs from healthy individuals increases with advanced age, both their self-renewal capacity and quiescence state decline [17, 18]. Because the self-renewal capacity of human preL-HSCs has not been assessed by serial transplantation in xenografts, it is unknown how the reduced self-renewal of bulk HSPCs isolated from healthy elderly individuals relates to

preL-HSC biology, if at all. Direct comparison of the self-renewal capacity of preL-HSCs and of non-mutated HSPCs from the same individuals over a spectrum of different ages is needed to address this question.

HSPC myeloid bias

Aging HSPCs also exhibit differentiation bias toward the myeloid lineage that can be explained by clonal selection of a sub-population of HSPCs that are myeloid biased [19, 20]. Likewise, myeloid genes are up regulated and lymphoid lineage-associated genes are mostly down regulated in aging HSPCs [17]. Whether clonal hematopoiesis and the expansion of preL-HSCs contribute to the age related myeloid bias is not known. However, indirect evidence from the initial studies which identified preL-HSCs suggests that such bias does exist, at least in the setting of AML in remission [5]. In these studies, preL-HSCs survived chemotherapy and in some cases expanded through remission. Of note, the allele frequency of mutant *DNMT3a* was higher in remission in myeloid cells than it was in lymphoid cells, suggesting that preL-HSCs can differentiate more efficiently to myeloid cells. Furthermore, in most cases, either at diagnosis or remission, the allele frequency of mutated *DNMT3a* was higher in preL-HSC than it was in B and T cells. In almost half of AML cases, the pLMs could not be identified in T cells, although they were clearly present in preL-HSCs, again suggesting altered differentiation. We acknowledge that there exist alternative possible interpretations of these AML studies, and that the role of aging in these data is unclear. We suggest therefore that direct differentiation assays be performed using preL-HSCs from healthy individuals over a range of ages, to test whether such cells might contribute to the myeloid bias observed in aging HSPCs. In this regard, it is important to note that murine HSPCs demonstrate a similar myeloid bias with no documentation of pLMs or clonal hematopoiesis.

Reduced bone marrow transplantation capacity

The clinical implications of the HSPC aging can be observed in the setting of allogeneic bone marrow transplantation (BMT). HSPCs from older individuals are inferior to those of younger donors when used in the BMT setting. Significantly reduced engraftment and overall survival is observed when older donors are used [21]. In this regard, it is interesting to note that preL-HSCs isolated from AML patients demonstrated a selective advantage over non-mutated HSPCs in xenograft assays [5]. Assuming that the progressive increase in clonal hematopoiesis incidence with aging results from the expansion of preL-HSCs, one might expect the stem cell pool harvested from elderly donors to contain more preL-HSCs than that of young

individuals. If indeed preL-HSCs have a selective advantage under replicative stress, one could expect better outcomes in the allogeneic BMT setting, in contrast to what is actually observed. Several possibilities might account for the poor engraftment of HSPCs from elderly individuals, despite containing more pLMs. First, preL-HSCs from normal elderly individuals (i.e. without AML) may be impaired functionally, and accordingly may engraft poorly in BMT setting. Second, in studies of clonal hematopoiesis, the frequency of preL-HSCs has been as low as 0.2 % [12]. Accordingly, the majority of harvested HSPCs before transplant would be expected to be non-mutated, and thus would not be expected to manifest a selective advantage. Of note, at least one case report of donor-derived leukemia has identified the transplantation capability of preL-HSCs in the setting of allogeneic BMT [22]. In this case, donor-derived preleukemic cells gave rise to recurrent leukemia in the recipient. Genetic analysis of the donor and recipient identified the same somatic *DNMT3a* and *IDH2* mutations in both, indicating that the donor's preL-HSCs were transplanted to the sibling recipient. These preL-HSCs from the healthy donor were subsequently able to grow in the recipient, and ultimately gave rise to overt leukemia in the recipient but not in the donor. These observations indicate that preL-HSCs can engraft in an allogeneic BMT setting. We cannot exclude the additional possibility in this case that donor and recipient BM microenvironments may differ in their permissiveness for preL-HSCs and further leukemia evolution.

HSPC aging and the bone marrow microenvironment

The bone marrow microenvironment also changes with chronological aging. Studies on human BM have demonstrated that BM fat content (yellow marrow) increases with aging. This increased BM fat content is associated with a decrease in the absolute numbers of HSPCs, and with alterations in BM cytokine levels [23, 24]. One can speculate, therefore, that the senescent microenvironment may impose a selective pressure that favors the fittest HSPC clones.

In support of this hypothesis, clonal tracking of retrovirally transduced murine HSCs transplanted into either young or old recipients suggests that the aged microenvironment exerts a selective pressure that favors only some HSPC clones [25]. Consistent with this, the role of the microenvironment has also been implicated in pre-leukemic conditions such as MDS and in the development of secondary leukemia. Specifically, deletion of *Dicer1* or of *RAR γ* in murine osteoprogenitors (part of the bone marrow environment) resulted in dysregulated myeloid lineage differentiation, eventually leading to myelodysplasia and secondary leukemia [26, 27]. Human observations also support this hypothesis. For example, in one study, 38 % of

patients with MDS or AML were shown to have increased β -catenin activity and Notch ligand jagged 1 expression in their osteoblasts, resulting in increased HSPC Notch activation [28]. More studies are needed to elucidate the features of the aging BM microenvironment that might apply a selective pressure to HSPCs, and to define those factors in the BM microenvironment that might accelerate the progression from preleukemia to overt leukemia. Finally, one can speculate that the immune system, another component of the hematopoietic microenvironment, and age-related changes in immune function (reduced diversity of the T cell repertoire, for example) may also play a role in clonal hematopoiesis, preL-HSC growth, and leukemia evolution. Details of such potential interactions remain to be elucidated.

Molecular changes in aging HSPCs

Cellular pathways such as the stress response and inflammation are up regulated in HSPCs from older individuals, whereas the chromatin remodeling, the expression of DNA repair associated genes, and TGF- β signaling are down regulated [29, 30]. Alterations in the transcriptional profiles in old HSPCs are remarkably stable in steady state and after transplantation, suggesting that at least some of the age-related changes in HSPCs are irreversible under physiological conditions [31]. The mechanisms responsible for the various phenotypic changes occurring in HSPCs as they age are not fully understood, but both intrinsic and extrinsic factors are believed to be contributing to this process. Telomere attrition, prolonged exposure to reactive oxygen species (ROS) and subsequent DNA damage, [32, 33], and the accumulation of epigenetic alterations have all been implicated as potential causes of the molecular changes observed in HSPCs during aging. The potential contribution(s) of pLMs and of clonal hematopoiesis to these molecular phenotypes remain to be defined.

A key observation from studies of HSPCs over the human lifespan [34], is the accumulation of mutations, both neutral and conferring a selective advantage, with increasing age. How such age-related mutations may relate to the aforementioned phenotypic changes observed with increasing age remains obscure. In the next part of this review we will focus on the genetic changes occurring during HSPC aging.

Clonal hematopoiesis and aging

Clonal hematopoiesis has long been identified as a feature of human HSPC aging [35]. Clonal hematopoiesis is not unique to healthy elderly individuals, and has also been observed in AML patients in remission [36], and in

lymphoma patients after autologous stem cell transplantation [37]. Initially, identified by X chromosome inactivation skewing (XIS) assays [35], clonal hematopoiesis with aging has more recently been identified by the detection of recurrent somatic mutations. For example, elderly individuals with clonal hematopoiesis (defined by XIS) tended to carry mutations in *TET2* (a gene commonly mutated in AML) [8]. Clonal hematopoiesis is now defined by the presence of somatic mutations in the peripheral blood with relatively high allele frequency. In one such study, the pLM *DNMT3a* R882 was detected by exome sequencing of cohorts of individuals with no hematological malignancy [5]. In these cohorts, peripheral blood was used as a germline surrogate with the goal of identifying inherited variants that may contribute genetic risk to complex traits. At the time it was not understood that preleukemic mutations can expand in bone marrow and blood to such a degree that their allele frequency can resemble that of a heterozygous germ-line variant. The variant allele frequency (VAF) cutoff that was used to call germ-line variants in these studies was 0.2 (20%), and at this cutoff ~1/1500 individuals carried the *DNMT3a* R882 mutation. Three additional studies have identified numerous other somatic gene variants in the exomes of peripheral blood cells analyzed for a variety of reasons. With a mutation calling cutoff of 0.02 (2%), the frequency of *DNMT3a* R882 mutations in one study increased to 1:790 [9]. In another study which used a more sensitive targeted sequencing approach, *DNMT3a* R882 mutations were identified in 1:92 healthy individuals [12]. Of note, about half of the clones identified in this study had a VAF of less than 3%. In unpublished results from our group, droplet digital PCR (VAF sensitivity 0.1%) was used to detect *DNMT3a* R882 variants and a carrier rate of 1:63 was observed (Fig. 1). It is becoming increasingly clear from these studies that more sensitive the detection method for pLMs, the higher the incidence of carriers that will be identified (Fig. 1). What is less clear is what should be defined as abnormal, and what is clinically relevant. In a recent review defining a new clinical-pathologic entity called clonal hematopoiesis of indeterminate potential (CHIP), the authors suggest to define clonal hematopoiesis based on the presence of a somatic mutation with a VAF greater than 0.02 (2%) [38]. The authors provide several arguments in support of this criterion. We suggest, however, that this conclusion may be premature, and that further sequencing-based studies are needed to better define clonal hematopoiesis, and more importantly to better understand when clonal hematopoiesis is actually relevant clinically.

Regardless of the VAF cutoff used to define clonal hematopoiesis, all of the studies mentioned above were able to demonstrate an increase in the frequency of clonal hematopoiesis (as defined by somatic mutation analysis)

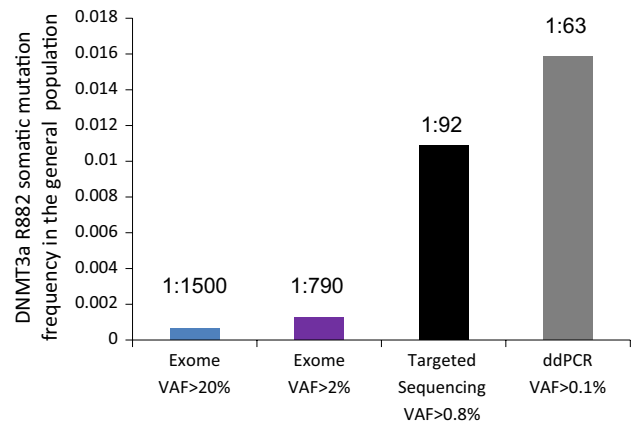


Fig. 1 The incidence of the most common preleukemic mutation (*DNMT3a* R882) increases as more sensitive detection methods are used. Initial reports on *DNMT3a* R882 suggested it was a germline mutation as it was identified in high allele frequencies in the peripheral blood of non-leukemia patients, however, once it was realized this is a preleukemic mutation it became clearer that healthy carriers can exist [5]. In different studies, different variant allele frequencies (VAF) were used to call the *DNMT3a* R882 as a somatic mutation out of bulk blood cells. A cutoff had to be used as no reference tissue was available to validate the somatic nature of the mutation. Interestingly the lower the cutoff, the higher the incidence of the variant that was identified, and it is still not clear if more sensitive methods will be used whether even higher incidence will be identified

with aging. This observation raises several interesting questions: why do specific mutations accumulate more than others? Do mutations accumulate in all stem cells at the same rate? Do pLMs have any effects on health?

Clinically relevant manifestations of clonal hematopoiesis

Clonal hematopoiesis defined by the presence of somatic mutations was associated with an increased risk of developing subsequent hematological malignancies. In one study, subsequent hematologic cancers were more than tenfold more common in the group with clonal hematopoiesis than they were in the group without [10]. Of note, clonal hematopoiesis with known pLMs and with unknown driver mutations carried the same increased risk of hematological malignancy. This observation is difficult to interpret, as structural variants and copy number variation were not assessed in these studies [9, 10]. For example, it has been demonstrated that both *inv(16)* [2] and *t(8;21)*, [4] which are structural variants that would not be detected by exome sequencing can be preleukemic lesions. Such structural variants would not have been identified in these studies, and could conceivably account for the clonal hematopoiesis with unknown drivers.

Interestingly, besides a higher incidence of subsequent hematological malignancies, clonal hematopoiesis was also

correlated with other clinical findings and outcomes. Among individuals with pLMs the only significant difference in blood-cell indices was an increase in red-blood-cell distribution width (RDW) [10]. Furthermore, carrying a pLM was associated with increased all-cause mortality with a hazard ratio of 1.4. Death from hematologic neoplasms alone could not account for the observed increase in mortality, as the increased death rate was specifically due to cardiovascular causes. When the combined effect of increased RDW and pLMs on mortality was assessed, individuals with a mutation in conjunction with a RDW of 14.5 % (the upper limit of the normal range) or higher, had a markedly increased risk of death as compared with those with a normal RDW and without pLMs (hazard ratio 3.7); increased RDW was recently found to be related with mortality in general. Increased RDW reflects a dysregulation of erythrocyte homeostasis involving both impaired erythropoiesis and abnormal red blood cell survival, which may be attributed to a variety of underlying metabolic abnormalities that occurs with human aging, including shortening of telomere length, oxidative stress, and inflammation [39]. The correlation between increased RDW and clonal hematopoiesis might be incidental as both are related to human aging. Alternative hypothesis would assume that both increased RDW and clonal hematopoiesis evolved due to the same selective pressures that were introduced by the aging human body.

Despite these anecdotal correlations, it is not clear whether the genetic changes accumulating over time in stem cells is a random process heavily dependent on stem cell replication as was suggested recently [1], or whether it is shaped by other factors. Answering this question is of great importance for understanding how pLMs occur, whether they confer a selective advantage, and how preL-HSCs acquire further mutations in some cases, leading to AML.

Mutation accumulation in HSPCs

When analyzing the accumulation of somatic mutations in stem cells it is important to take into account the proliferative workload of different stem cells. In a recent analysis of the correlation between the total number of stem cell replications and the incidence of malignancy [1], the authors identified a tight correlation between the two factors. It was thus concluded that the majority of cancer occurs due to ‘bad luck’, random mutations arising during DNA replication in normal, noncancerous, stem cells. In this study, it was hypothesized that the lifetime number of stem cell divisions within a tissue accounts for the stochastic/random accumulation of mutations. Accordingly, tissues with greater stem cell replication will carry more mutations, and consequently will have a higher incidence of cancer. While the authors have provided evidence for such a correlation, several questions remain: first, the number of stem cell

replications in a tissue is the product of the number of tissue stem cells and how many times each of them divides [40, 41]. As we will next demonstrate, the accumulation of mutations is influenced not only by the total number of stem cell divisions, but also by the topology of the lineage tree [40–42]. A ‘shallow’ tree topology with many stem cells that divide at a low rate has a significantly lower probability of yielding an oncogenic clone than does a ‘deep’ tree containing a few rapidly dividing stem cells (Fig. 2). Moreover, we will show that the presence of a minority of stem cells that replicates at a faster rate than do the others can increase the chance of cancer in a highly non-linear manner. Since tissue stem cells have been shown to not only undergo asymmetric divisions creating both stem and progenitor daughter cells, but also occasional symmetric divisions, generating two stem cell daughters (and thus replacing other stem cells) [43], these few rapidly dividing stem cells have a higher chance to increase in frequency in the stem cell pool, leading to clonal hematopoiesis (Fig. 2).

The effect of proliferative heterogeneity on mutation accumulation

To evaluate the effect of proliferative heterogeneity on tumor formation, we analyze a simple mathematical model of L stem cells, each dividing strictly asymmetrically for D divisions over a time period T . In this scenario the total number of accumulated stem cell divisions is $N = L \cdot D$. We also assume that a cell must acquire at least k mutations to give rise to a tumor clone. In the first scenario, we assume L stem cells that are equivalent in their proliferative dynamics. An oncogenic mutation occurs at a probability p per division, which we assume is independent of the previous oncogenic mutations that have already been acquired (thus we ignore mutations that may increase genomic instability [44]). The probability that by time T there have been k mutations in a stem cell (we assume k mutations are sufficient to create an oncogenic clone) is the binomial probability:

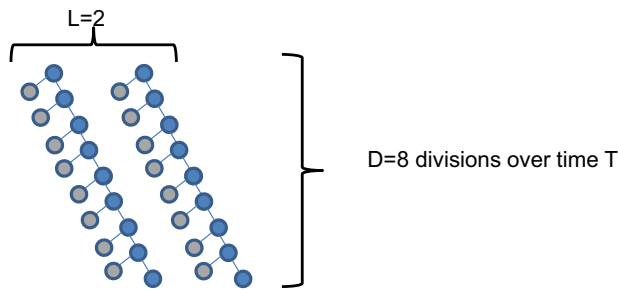
$$\begin{aligned} P(k|D) &= P(k \text{ muts. over } D \text{ divs.}) \\ &= \binom{D}{k} \cdot p^k \cdot (1-p)^{D-k} \sim \binom{D}{k} \cdot p^k \sim D^k \cdot p^k \end{aligned} \quad (1)$$

where we assumed p is small so that $1-p$ is negligible and that $D \gg 1$.

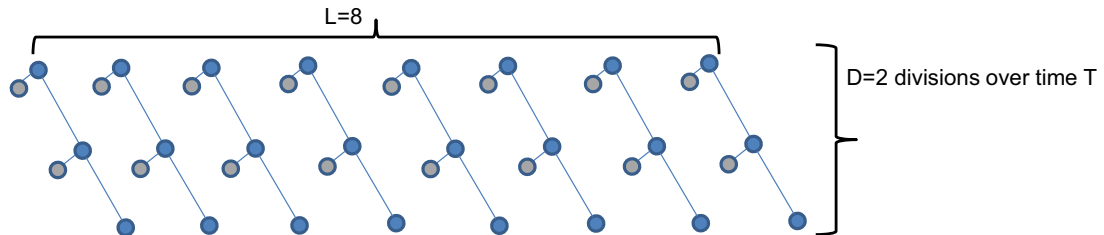
Hence the expected number of stem cells at time T that have undergone exactly k mutations, \bar{n} , is:

$$\begin{aligned} \bar{n} &= L \binom{D}{k} \cdot p^k \cdot (1-p)^{D-k} = L \cdot P(k|D) \sim k! \cdot L \cdot D^k \cdot p^k \\ &= k! \cdot \frac{N}{D} \cdot D^k \cdot p^k = k! \cdot N \cdot p^k \cdot D^{k-1} \end{aligned} \quad (2)$$

(A) Deep lineages – L (number of clones)=2 stem cells dividing $D=8$ times each yielding $N=2 \times 8=16$ SC divisions



(B) Shallow lineages – 8 stem cells dividing $D=2$ times each yielding $N=8 \times 2=16$ SC divisions



(C) Occasional symmetric divisions lead to clonal hematopoiesis (red clone)

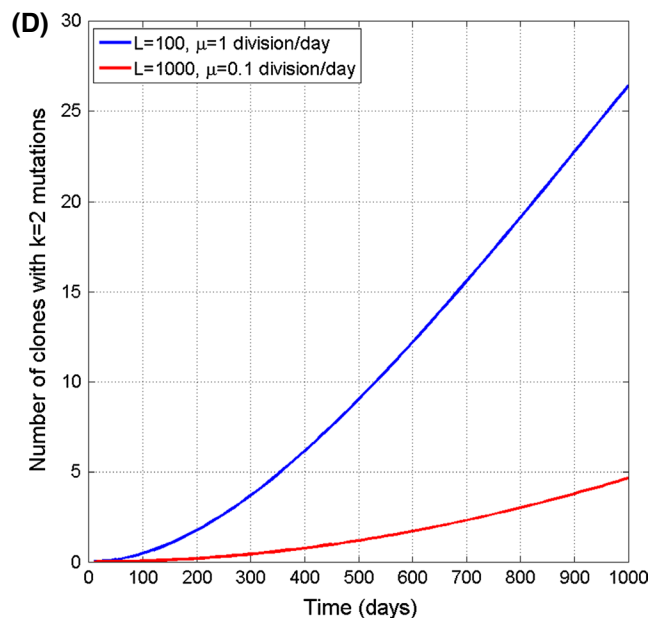
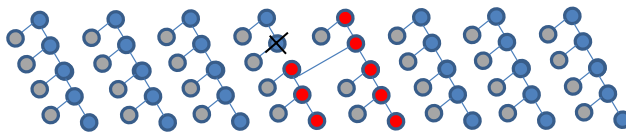


Fig. 2 The effect of lineage structure and stem cell kinetics on mutation accumulation in stem cells. **a** In shallow stem cell lineages less stem cells [L number of replicating stem cells (SC)] divide so that each stem cells replicate more (D number of SC divisions over time T). **b** In deep stem cell lineages more stem cell divide as that each stem cell divide less. **c** regardless of the lineage topology SCs undergoing asymmetric divisions will increase clone size and lead to stem cell expansion (clonal hematopoiesis). **d** The number of clones with

$k = 2$ somatic oncogenic mutations under two lineage topologies—one with low number of rapidly dividing stem cells (*blue* division rate $\mu = 1/\text{day}$), and another with higher number of stem cells that divide less rarely (*red* $\mu = 0.1/\text{day}$). While both topologies have the same number of total stem cell divisions at each time point the accumulation of clones with two or more mutations is significantly higher in the ‘deep’ topology. Suggesting that less the stem cells contributing to hematopoiesis the higher the chance for mutation accumulation

where we have used the fact that the total number of division is $N = L \cdot D$ and approximated the binomial coefficient by $k! \cdot D^{k-1}$.

Equation (2) indicates that the number of mutated tissue stem cells indeed increases linearly with the total number of stem cell divisions, N [1], but importantly, it increases non-linearly with the average lineage depth D . Thus, preserving the total number of stem cell divisions but ‘spreading’ them over larger number of tissue stem cells L each undergoing fewer divisions D dramatically decreases the probability of having a stem cell that has acquired the necessary number of k mutations. We will demonstrate this effect in numerical examples below.

It is interesting to note here the special case of $k = 1$, if only a single hit is needed to create an oncogenic clone Eq. (2) indicates that $\bar{n} = N \cdot p^1 \cdot D^{1-1} = N \cdot p$. Thus $k = 1$ is the only case where the probability of cancer is independent of lineage depth. In all other cases the average lineage depth has at least as important an effect. However, to our knowledge very little acute leukemia is the result of a single hit. Maybe the mixed lineage leukemia MLL is a good example of such a case, while all other acute leukemia require more than one mutation.

While Eq. (2) provides us with some intuition for the scaling behavior of the number of mutated stem cells vs. the total number of stem cells divisions and the average lineage depth, it is only an approximation, since in fact we need k mutations or more to create an oncogenic clone, rather than exactly k mutations. The probability to get k mutations or more in a lineage of depth D assuming probability p for a mutation over 1 division is:

$$P(> k|D) = P(\text{more than } k \text{ muts. over } D \text{ divs.}) \\ = \sum_{i=k}^D \binom{D}{i} \cdot p^i \cdot (1-p)^{D-i} \quad (3)$$

Using Eq. (3) we obtain the expected number of stem cells with k oncogenic mutation or more by time T :

$$\bar{n} = L \cdot \sum_{i=k}^D \binom{D}{i} \cdot p^i \cdot (1-p)^{D-i} \quad (4)$$

As a numeric example, let us consider two tissue designs. Design A has $L = 100$ stem cells, each dividing $D = 1000$ times over a given period T . If at least $k = 2$ oncogenic mutations are needed for a cancer clone, and the probability of an oncogenic mutation per division is $p = 0.001$, Eq. (4) yields: $\bar{n}_A = 26.4$ mutated stem cell by time T . In design B there are $L = 1000$ stem cells, each dividing $D = 100$ times over time period T . In this case, the number of mutated stem cells would only be $\bar{n}_B = 4.6$. Thus, having ten times more stem cells each dividing ten times slower decreased the chance of cancer by about

sixfold. If $k = 3$ oncogenic mutations were needed, the decrease would be almost tenfold, and if $k = 6$, we would get a 50,000-fold decrease in the chance of having an oncogenic clone in design B. Importantly, both designs have the same number of total stem cell divisions: $N = 100,000$.

The nonlinear dependence on the lineage depth D obtained from Eqs. (2) and (4) also gives rise to a high sensitivity to a few clones that divide significantly faster than do the others. As an example, consider another tissue design where $L = 100$ stem cells have undergone $N = 100,000$ divisions by time T , but this time, ten stem cells divide faster than do the others and ‘carry’ half the load of 50,000 divisions (thus each of these ten ‘fast’ stem cells divides $D_1 = 5000$ times over time T), whereas the remaining 90 ‘slow’ stem cells divide $D_2 = 555$ times over this time period ($10 \times 5000 + 90 \times 555 = 100,000$ divisions). In this case, if $k = 6$ or more oncogenic mutations are needed to initiate cancer, the ‘fast’ stem cells will generate on average 3.8 oncogenic stem cell clones whereas the 90 ‘slow’ ones will only generate on average 0.002 oncogenic clones. If all 100 stem cells were to divide at these rates, only 0.06 oncogenic clones would have been generated. Thus, a slight imbalance in the proliferative workload of tissue stem cells can lead to a dramatic effect on the chance of tumors. Now if we include occasional symmetric divisions leading to neutral drift dynamics [43], the faster a stem cell divides the higher is its probability to expand and give rise to clonal hematopoiesis.

The recent observations of clonal hematopoiesis with aging suggests that such an imbalance in proliferation dynamics does indeed occur, as blood counts are generally preserved with aging, while at least in some individuals, specific clones contribute more to the total pull, suggesting that at least at some level of the hematopoietic hierarchy, HSPCs with pLMs replicate more than do wild type HSCPs. The concept of a decreased effective population size replicating more in response to physiological needs might be a major contributor to mutation accumulation and aging related malignancy in the hematopoietic system.

Preleukemic mutations and progression to leukemia

Clearly, the above model suffers from oversimplification, due to the limitations of our current knowledge. It is not clear whether the initiating preleukemic mutations do, or do not, have an effect on the probability of the next events to occur. What is also becoming clearer is that the nature of the mutations being acquired at early stages, and correlated with clonal hematopoiesis, has a great deal of deterministic effect on the probability to evolve to overt leukemia. For example, *DNMT3a* mutations, which are one of the most common pLMs, are also one of the most common leukemic mutations, and are highly correlated with the appearance

of *NPM1c* and *FLT3/ITD* mutations at the leukemic phase. On the other hand, *JAK2* mutations, which are also common variants in the setting of clonal hematopoiesis, only account for low percentages of AMLs and MDSs. Other mutations such as *PPM1D* have not been described so far in myeloid malignancies, but are abundant in individuals with clonal hematopoiesis [11]. As all of these variants (*DNMT3a*, *JAK2*, *PPM1D*) are recurrently mutated in clonal hematopoiesis, and their frequency increases with age, it is accordingly reasonable to assume that they have been selected due to the specific phenotypes they provide to the HSPCs in the aging marrow environment. However, their unequal distribution among the various hematological malignancies suggests that HSPCs carrying each of these variants have a different probability to progress to leukemia. This may be related to changes in stem cell dynamics or in mutation rates specific to the different initiating events. For example, *DNMT3a* (part of the epigenetic machinery) mutations might induce higher global mutation rates, or an increased regional mutation rate, that will increase the probability to develop leukemia. The role of the epigenetic machinery in maintaining genomic stability has been described and reviewed in the past [45]. Another possible explanation could be that both *DNMT3a*- and *JAK2*-mutated HSPCs have the same probability to acquire the next events required for overt leukemia, but *DNMT3a*-mutated HSPCs have gained a clonal advantage under specific environmental conditions that also provide a selective advantage to leukemic mutations. On the other hand, *JAK2* clonal hematopoiesis might have developed under conditions that are less permissive for leukemia. Another non-random phenomenon observed in individuals was clonal hematopoiesis was recently described by McKerrel et al. who demonstrated that pLMs in spliceosome machinery genes accumulate in individuals with higher chronological age, as compared to mutations in *DNMT3a* and *JAK2* [12]. This observation suggests that although mutations might accumulate randomly as stem cells replicate, they might endow a selective advantage only at a specific age or maybe in a specific microenvironment.

The above taken together, it is clear that almost all elderly individuals develop clonal hematopoiesis, and in many cases this is associated with pLMs. A major remaining question is whether the probability of developing leukemia from a preleukemic clone is random or deterministic. No clear answer can be given at this point, and future studies comparing individuals who develop leukemia to other carriers of pLMs should be conducted to address this question.

Preleukemic mutations repertoire

Initial studies on pLMs demonstrated that a large number of pLMs in AML occur in epigenetic modifying genes

[2]. The same mutations were also identified in the clonal hematopoiesis studies [9–12] among individuals with no overt leukemia. Why genes from the epigenetic machinery are specifically mutated in human clonal hematopoiesis is not clear. Other mutations observed in individuals with clonal hematopoiesis were related to lymphoid malignancies (MYD88 in Waldenstrom's Macroglobulinemia and MGUS, and STAT3 in T-ALL) [9], suggesting that clonal hematopoiesis might be a premalignant state in these malignancies as well.

Most of the studies dealing with clonal hematopoiesis and pLMs in the modern sequencing era have focused on available exome sequencing or specifically designed targeted sequencing. The limitations of these approaches when used at relatively low coverage is the underestimation of structural variants, and CNVs, and an over-simplification of the complexity of the genetic landscape of the entire genome as HSPCs age. Consistent with this notion, studies focusing on CNV have reported that clonal chromosomal changes also increase in frequency with aging (chromosomal clonal mosaicism can be observed in 2–3 % of elderly individuals). Notably, chromosomal changes in this study cohort demonstrated changes in chromosomal regions that are highly mutated in hematological malignancies (20q–, 13q–, 11q–, 17p–, 12+ and 8+) [46].

How can we identify individuals at risk for progression?

Transformation to leukemia can occur after a long preleukemic phase (years), but can probably also occur with a relative shorter latent phase in the presence of a strong initiating events [the *t(15;17)* translocation in APL, or translocations involving *MLL* in ALL or AML, for example]. The probability of leukemic transformation depends on the degree of impairment of differentiation in a cell that can still self-renew. By extension, the probability of leukemic transformation in a population of preleukemic cells depends on the degree of impairment of differentiation of the various sub-clones comprising the preleukemic pool. We hypothesize that the accumulation of specific genetic variants in HSPCs will add to the impairment of differentiation, and will conversely increase the likelihood of self-renewal. Accordingly, to determine an individual's risk for progression to AML, we suggest that assays that combine the assessment of the functional output of preleukemic clones, combined with comprehensive genetic analysis, be developed. By this approach, clones with impaired differentiation capacity, but with a selective advantage can be identified in competitive functional assays. If such clone will also carry recurrent preleukemic mutations, it would be predicted to have a higher chance of evolving to AML. One can speculate, however, that while the identification of pLM-bearing individuals at particular risk for leukemic

transformation is a laudable goal, the most effective means to prevent leukemia might be the prevention of clonal hematopoiesis in the first place.

Concluding remarks

Clonal hematopoiesis either with pLMs or with an unknown driver seems to be a very common feature of the aging hematopoietic system, and probably in the aging of other highly replicating tissues as well [47]. Although preL-HSCs likely have a selective advantage that leads to clonal hematopoiesis, the aging hematopoietic system is far from being monoclonal. Whether any of the steps in the evolution from normal wild type HSPC to overt leukemia are random or deterministic is yet to be answered. What is becoming clearer, is that the risk of hematologic malignancy increases in only some individuals with clonal hematopoiesis. While some such individuals are known to share common features such as a high RDW, we are currently unable to identify patients at greatest risk. More research is needed to solve the enigma of why only some individuals evolve to overt leukemia, and whether clonal hematopoiesis has any relevance in other aspects of the aging of the hematopoietic and immune systems.

Box 1

Clonal hematopoiesis A decrease in the genetic diversity of cells from the hematopoietic system, due to the acquisition of mutations.

Preleukemic mutations (pLMs) Mutations that can be recurrently detected in leukemic cells but also in functionally differentiating cells.

Preleukemia A subtype of clonal hematopoiesis due to preleukemic mutations that eventually will evolve overt leukemia.

Preleukemic Stem cells preL-HSCs– HSPCs carrying preleukemic mutations which can still maintain differentiation capacity.

Leukemia Clonal expansion of hematopoietic cells with aberrant differentiation and reduced functionality.

Acknowledgments We are thankful to Georges Natsoulis for helpful discussions.

References

- Tomasetti C, Vogelstein B. Cancer etiology. Variation in cancer risk among tissues can be explained by the number of stem cell divisions. *Science*. 2015;347(6217):78–81.
- Corces-Zimmerman MR, et al. Preleukemic mutations in human acute myeloid leukemia affect epigenetic regulators and persist in remission. *Proc Natl Acad Sci USA*. 2014;111(7):2548–53.
- Jan M, et al. Clonal evolution of preleukemic hematopoietic stem cells precedes human acute myeloid leukemia. *Sci Transl Med*. 2012;4(149):149ra118.
- Miyamoto T, Weissman IL, Akashi K. AML1/ETO-expressing nonleukemic stem cells in acute myelogenous leukemia with 8:21 chromosomal translocation. *Proc Natl Acad Sci USA*. 2000;97(13):7521–6.
- Shlush LI, et al. Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia. *Nature*. 2014;506(7488):328–33.
- Shlush LI, Minden MD. Preleukemia: the normal side of cancer. *Curr Opin Hematol*. 2015;22(2):77–84.
- El Kassas N, et al. X-chromosome inactivation in healthy females: incidence of excessive lyonization with age and comparison of assays involving DNA methylation and transcript polymorphisms. *Clin Chem*. 1998;44(1):61–7.
- Busque L, et al. Recurrent somatic TET2 mutations in normal elderly individuals with clonal hematopoiesis. *Nat Genet*. 2012;44(11):1179–81.
- Genovese G, et al. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N Engl J Med*. 2014;371(26):2477–87.
- Jaiswal S, et al. Age-related clonal hematopoiesis associated with adverse outcomes. *N Engl J Med*. 2014;371(26):2488–98.
- Xie M, et al. Age-related mutations associated with clonal hematopoietic expansion and malignancies. *Nat Med*. 2014;20(12):1472–8.
- McKerrell T, et al. Leukemia-associated somatic mutations drive distinct patterns of age-related clonal hemopoiesis. *Cell Rep*. 2015;10(8):1239–45.
- Nordling CO. A new theory on cancer-inducing mechanism. *Br J Cancer*. 1953;7(1):68–72.
- Armitage P, Doll R. The age distribution of cancer and a multi-stage theory of carcinogenesis. *Br J Cancer*. 1954;8(1):1–12.
- Linton PJ, Dorshkind K. Age-related changes in lymphocyte development and function. *Nat Immunol*. 2004;5(2):133–9.
- Lichtman MA, Rowe JM. The relationship of patient age to the pathobiology of the clonal myeloid diseases. *Semin Oncol*. 2004;31(2):185–97.
- Pang WW, et al. Human bone marrow hematopoietic stem cells are increased in frequency and myeloid-biased with age. *Proc Natl Acad Sci USA*. 2011;108(50):20012–7.
- Yahata T, et al. Accumulation of oxidative DNA damage restricts the self-renewal capacity of human hematopoietic stem cells. *Blood*. 2011;118(11):2941–50.
- Cho RH, Sieburg HB, Muller-Sieburg CE. A new mechanism for the aging of hematopoietic stem cells: aging changes the clonal composition of the stem cell compartment but not individual stem cells. *Blood*. 2008;111(12):5553–61.
- Beerman I, et al. Functionally distinct hematopoietic stem cells modulate hematopoietic lineage potential during aging by a mechanism of clonal expansion. *Proc Natl Acad Sci USA*. 2010;107(12):5465–70.
- Kollman C, et al. Donor characteristics as risk factors in recipients after transplantation of bone marrow from unrelated donors: the effect of donor age. *Blood*. 2001;98(7):2043–51.
- Yasuda T, et al. Leukemic evolution of donor-derived cells harboring *IDH2* and *DNMT3A* mutations after allogeneic stem cell transplantation. *Leukemia*. 2014;28:426–8.
- Griffith JF, et al. Bone marrow fat content in the elderly: a reversal of sex difference seen in younger subjects. *J Magn Reson Imaging*. 2012;36(1):225–30.

24. Tuljapurkar SR, et al. Changes in human bone marrow fat content associated with changes in hematopoietic stem cell numbers and cytokine levels with aging. *J Anat.* 2011;219(5):574–81.
25. Vas V, et al. Aging of the microenvironment influences clonality in hematopoiesis. *PLoS One.* 2012;7(8):e42080.
26. Raaijmakers MH, et al. Bone progenitor dysfunction induces myelodysplasia and secondary leukaemia. *Nature.* 2010;464(7290):852–7.
27. Walkley CR, et al. A microenvironment-induced myeloproliferative syndrome caused by retinoic acid receptor gamma deficiency. *Cell.* 2007;129(6):1097–110.
28. Kode A, et al. Leukaemogenesis induced by an activating beta-catenin mutation in osteoblasts. *Nature.* 2014;506(7487):240–4.
29. Chambers SM, et al. Aging hematopoietic stem cells decline in function and exhibit epigenetic dysregulation. *PLoS Biol.* 2007;5(8):e201.
30. Sun D, et al. Epigenomic profiling of young and aged HSCs reveals concerted changes during aging that reinforce self-renewal. *Cell Stem Cell.* 2014;14(5):673–88.
31. Wahlestedt M, et al. An epigenetic component of hematopoietic stem cell aging amenable to reprogramming into a young state. *Blood.* 2013;121(21):4257–64.
32. Norddahl GL, et al. Accumulating mitochondrial DNA mutations drive premature hematopoietic aging phenotypes distinct from physiological stem cell aging. *Cell Stem Cell.* 2011;8(5):499–510.
33. Ahlqvist KJ, et al. Somatic progenitor cell vulnerability to mitochondrial DNA mutagenesis underlies progeroid phenotypes in Polg mutator mice. *Cell Metab.* 2012;15(1):100–9.
34. Welch JS, et al. The origin and evolution of mutations in acute myeloid leukemia. *Cell.* 2012;150(2):264–78.
35. Busque L, et al. Nonrandom X-inactivation patterns in normal females: lyonization ratios vary with age. *Blood.* 1996;88(1):59–65.
36. Fialkow PJ, Janssen JW, Bartram CR. Clonal remissions in acute nonlymphocytic leukemia: evidence for a multistep pathogenesis of the malignancy. *Blood.* 1991;77(7):1415–7.
37. Mach-Pascual S, et al. Predictive value of clonality assays in patients with non-Hodgkin's lymphoma undergoing autologous bone marrow transplant: a single institution study. *Blood.* 1998;91(12):4496–503.
38. Steensma DP, et al. Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes. *Blood.* 2015;126(1):9–16.
39. Salvagno GL, et al. Red blood cell distribution width: a simple parameter with multiple clinical applications. *Crit Rev Clin Lab Sci.* 2015;52(2):86–105.
40. Wasserstrom A, et al. Estimating cell depth from somatic mutations. *PLoS Comput Biol.* 2008;4(4):e1000058.
41. Reizel Y, et al. Colon stem cell and crypt dynamics exposed by cell lineage reconstruction. *PLoS Genet.* 2011;7(7):e1002192.
42. Shlush LI, et al. Cell lineage analysis of acute leukemia relapse uncovers the role of replication-rate heterogeneity and microsatellite instability. *Blood.* 2012;120(3):603–12.
43. Klein AM, Simons BD. Universal patterns of stem cell fate in cycling adult tissues. *Development.* 2011;138(15):3103–11.
44. Nowak MA, Michor F, Iwasa Y. Genetic instability and clonal expansion. *J Theor Biol.* 2006;241(1):26–32.
45. Jones PA, Gonzalgo ML. Altered DNA methylation and genome instability: a new pathway to cancer? *Proc Natl Acad Sci USA.* 1997;94(6):2103–5.
46. Laurie CC, et al. Detectable clonal mosaicism from birth to old age and its relationship to cancer. *Nat Genet.* 2012;44(6):642–50.
47. Martincorena I, et al. Tumor evolution. High burden and pervasive positive selection of somatic mutations in normal human skin. *Science.* 2015;348(6237):880–6.