
Article:
Cull, Alyssa H., Kent, David G. orcid.org/0000-0001-7871-8811 and Warren, Alan J. (2024) Emerging genetic technologies informing personalized medicine in Shwachman-Diamond syndrome and other inherited BMF disorders. Blood. pp. 931-939. ISSN 1528-0020

https://doi.org/10.1182/blood.2023019986

---

blood

American Society of Hematology
2021 L Street NW, Suite 900,
Washington, DC 20036
Phone: 202-776-0544 | Fax 202-776-0545
editorial@hematology.org

Emerging genetic technologies informing personalized medicine in SDS and other inherited bone marrow failure disorders

Tracking no: BLD-2023-019986-CR2

Alyssa Cull (University of York, United Kingdom) David Kent (University of York, United Kingdom) Alan Warren (University of Cambridge, United Kingdom)

Abstract:
The ribosomopathy Shwachman-Diamond syndrome (SDS) is a rare autosomal recessive inherited bone marrow failure syndrome (IBMFS) caused by mutations in the Shwachman-Bodian-Diamond syndrome (SBDS) gene, that is associated with an increased risk of myeloid malignancy. Tracking how hematopoietic stem cell (HSC) clonal dynamics change over time, assessing whether somatic genetic rescue mechanisms affect these dynamics, and mapping out when leukemic driver mutations are acquired is important to understand which individuals with SDS may go on to develop leukemia. In this review, we will discuss how new technologies that allow researchers to map mutations at the level of single HSC clones are generating important insights into genetic rescue mechanisms and their relative risk for driving evolution to leukemia, and how these data can inform the future development of personalized medicine approaches in SDS and other IBMFSs.

Conflict of interest: COI declared - see note

COI notes: A.J.W. is a consultant for SDS Therapeutics. The remaining authors declare no competing interests.

Preprint server: No;

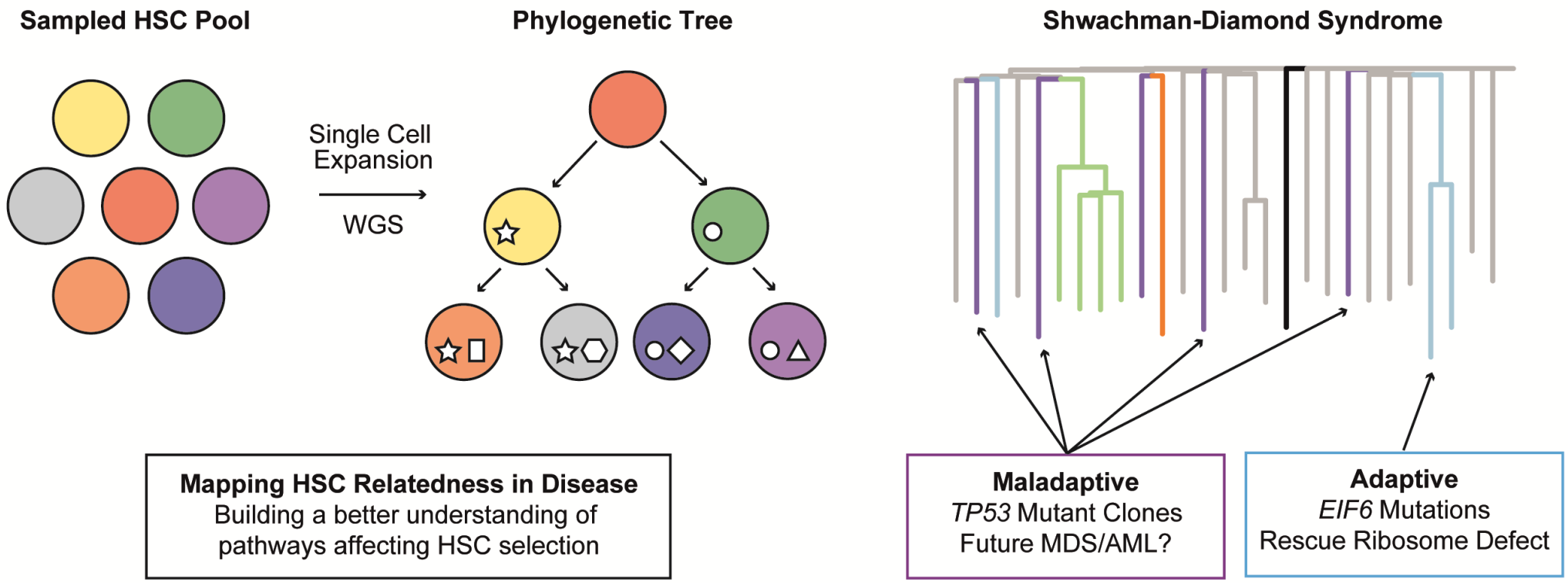
Author contributions and disclosures: A.H.C., D.G.K. and A.J.W. contributed equally to the text of this paper and A.H.C. created the figures.

Non-author contributions and disclosures: No;

Agreement to Share Publication-Related Data and Data Sharing Statement:

Clinical trial registration information (if any):

Emerging genetic technologies informing personalized medicine in Shwachman-Diamond Syndrome and other inherited bone marrow failure disorders



New technologies that allow researchers to map mutations at the level of single HSC clones are generating important insights into the pathways central to Shwachman-Diamond Syndrome disease biology.

Figure 1

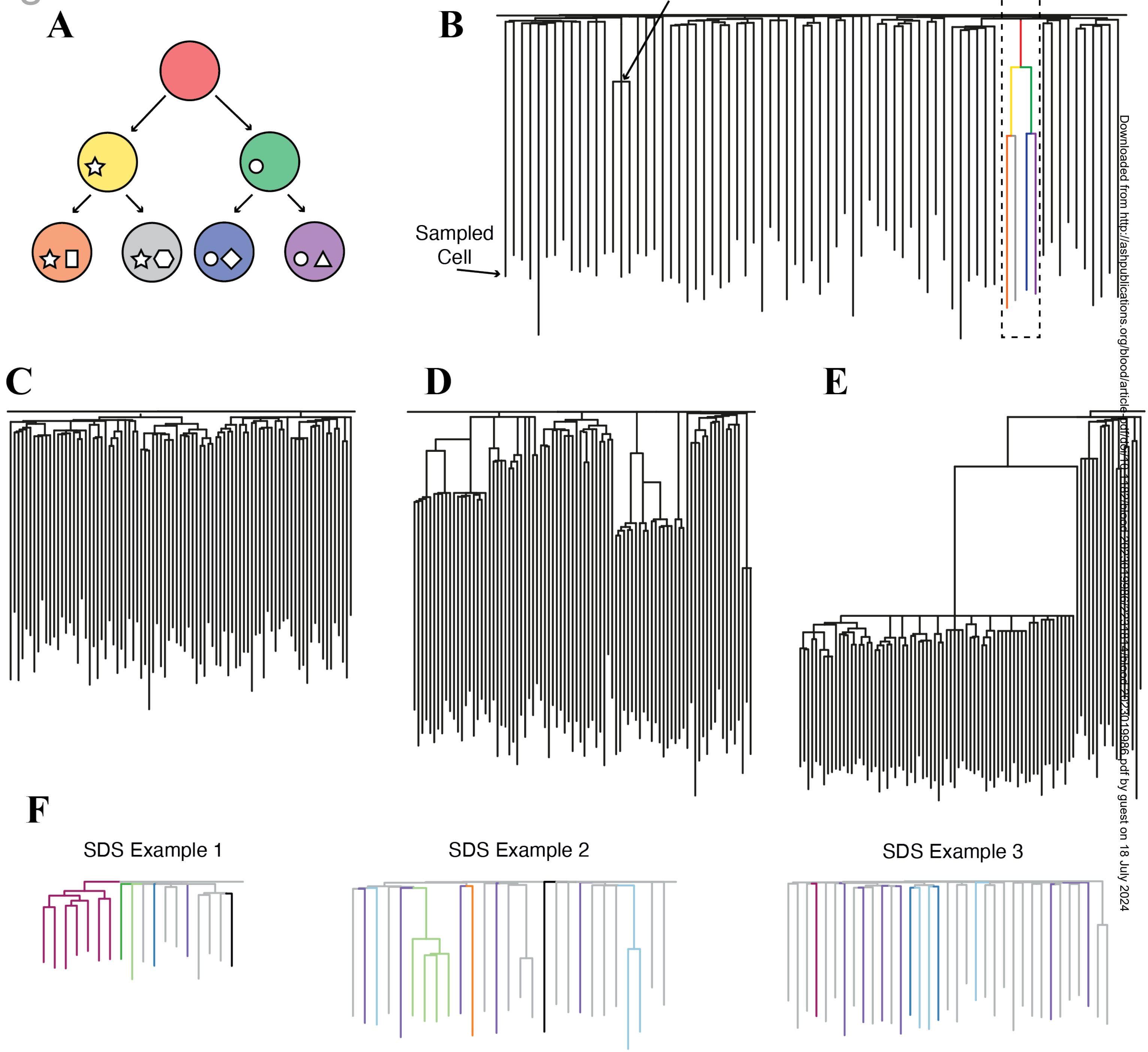
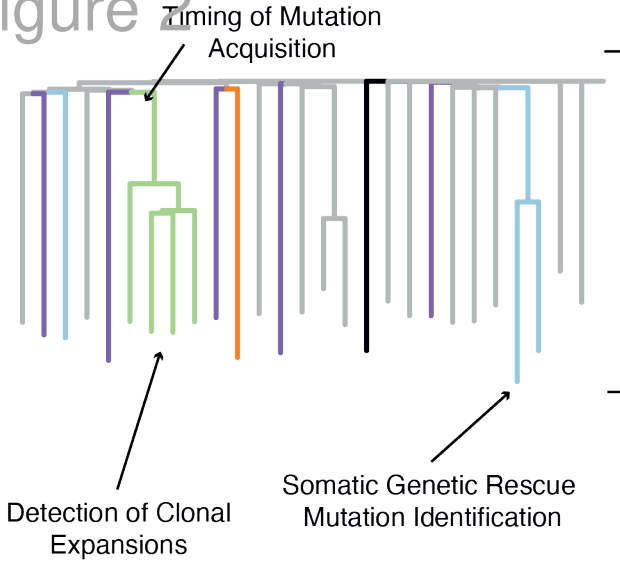


Figure 2



Emerging genetic technologies informing personalized medicine in SDS and other inherited bone marrow failure disorders

Running Title: Applications of single-cell analyses to IBMFSs

Alyssa H. Cull[1], David G. Kent[1]*, and Alan J. Warren[2,3]*.

[1]York Biomedical Research Institute, Department of Biology, University of York, York, UK; [2]Cambridge Institute for Medical Research, Department of Clinical Biochemistry, University of Cambridge, Cambridge Biomedical Campus, The Keith Peters Building, Hills Road, Cambridge CB2 0XY, UK.

[3]Wellcome MRC Cambridge Stem Cell Institute, University of Cambridge, Cambridge, UK

* These authors contributed equally. Correspondence: david.kent@york.ac.uk and ajw1000@cam.ac.uk

Text word count: 4590

Abstract word count: 135

Figure count: 2

Reference count: 79

Abstract

The ribosomopathy Shwachman-Diamond syndrome (SDS) is a rare autosomal recessive inherited bone marrow failure syndrome (IBMFS) caused by mutations in the Shwachman-Bodian-Diamond syndrome (*SBDS*) gene, that is associated with an increased risk of myeloid malignancy. Tracking how hematopoietic stem cell (HSC) clonal dynamics change over time, assessing whether somatic genetic rescue mechanisms affect these dynamics, and mapping out when leukemic driver mutations are acquired is important to understand which individuals with SDS may go on to develop leukemia. In this review, we will discuss how new technologies that allow researchers to map mutations at the level of single HSC clones are generating important insights into genetic rescue mechanisms and their relative risk for driving evolution to leukemia, and how these data can inform the future development of personalized medicine approaches in SDS and other IBMFSs.

Introduction

Inherited bone marrow failure syndromes (IBMFSs) are heterogeneous diseases that typically present with cytopenias and multi-organ dysfunction. The most common IBMFSs include Fanconi anemia (FA), Diamond-Blackfan anemia (DBA), dyskeratosis congenita (DC), GATA2 deficiency, SAMD9/9L-related syndromes, and Shwachman-Diamond syndrome (SDS). In addition to bone marrow (BM) failure and ineffective hematopoiesis, the genetic drivers of these diseases often lead to abnormalities in other tissues and patients are at a much-increased risk of developing a number of cancers early in life, including leukemia.[1]

The ribosomopathy SDS is a fascinating autosomal recessive BM failure disorder caused by defective ribosome assembly.[2] In 90% of cases, SDS is caused by mutations in a single gene, Shwachman-Bodian-Diamond syndrome (*SBDS*).[3,4] The ribosome assembly defect in SDS results in a range of highly variable clinical manifestations including neutropenia, poor growth, exocrine pancreatic insufficiency, skeletal abnormalities, cognitive impairment and other organ dysfunction.[2] Individuals with SDS are at higher risk of developing poor prognosis MDS/AML within the first few decades of life.[5–7,8] While most individuals are diagnosed as children, patients may still remain undiagnosed into adulthood. Understanding why some SDS patients progress to MDS/AML whereas others do not is key to developing preventative strategies and personalized therapeutic interventions.

Given that it is a monogenic disorder driving a fundamental biological defect in all cells, SDS represents an excellent model disease to explore somatic genetic rescue mechanisms. For many years, a key question in the field of SDS research remains: how does a germline ribosome defect associated with a growth disadvantage lead to the development of leukemia? The risk of transformation for individuals with SDS is high. In the French SDS registry, approximately 30% of patients developed a

myeloid malignancy by the age of 30.[6] The cumulative 20-year incidence was 9.8% in the Italian SDS registry,[7] while in the North American SDS registry, 17% developed a myeloid cancer by 18 years of age.[8,9] Outcomes for patients who progress to myeloid neoplasms are very poor. In a recent retrospective study, the three-year overall survival for patients who developed MDS or AML was estimated to be 51% and 11%, respectively.[8] Particularly in the case of SDS-related MDS, these survival rates are worse than those observed in the non-SDS population.[8] Understanding the genetic changes that drive transformation to these aggressive blood cancers becomes of paramount importance and while acquired biallelic *TP53* mutations are common in myeloid cancers in SDS, this may not be the only route.[8,10]

In this review, we detail how recently developed technologies have expanded our ability to map the mutations involved in genetic rescue and leukemic transformation at the level of single hematopoietic stem cell (HSC) clones. Understanding the range and type of genetic rescue mechanisms and their relative risk for driving disease evolution of SDS to leukemia will help inform future drug design and clinical decision making for IBMFSs more generally.

Molecular basis of SDS

Great strides have been made in understanding the molecular basis of SDS.[2,11–13] Most individuals with SDS carry germline mutations in the highly conserved *SBDS* gene encoding the SBDS protein that licenses entry of the large ribosomal subunit into the actively translating pool of ribosomes.[2] Approximately 200 proteins and 76 small nucleolar RNA molecules coordinate the assembly of 80 ribosomal proteins and four ribosomal RNAs into the small (40S) and large (60S) ribosomal subunits. The initial steps in ribosome biogenesis occur in the nucleolus, where the precursors of the 40S and 60S subunits are produced via endonucleolytic cleavage of the 90S pre-ribosome.[14–18] Upon export into the cytoplasm, the nascent 40S and 60S ribosomal subunits complete the final steps in maturation. One of the last cytoplasmic maturation events is eviction of the anti-association factor eIF6 from the 60S ribosomal subunit by the concerted action of the GTPase elongation factor-like 1 (EFL1) and its cofactor SBDS.[19–22] Removal of eIF6 from the intersubunit face of the nascent 60S subunit is essential to allow 80S ribosome assembly as eIF6 physically blocks joining of the nascent 60S to the 40S ribosomal subunit. SBDS and EFL1 also act as general eIF6 release factors by promoting the recycling of post-termination eIF6-bound 60S subunits back into active translation.[22]

Located on chromosome 7q11, the *SBDS* gene is most commonly mutated within exon 2 (c.258>+2T>C and c.183_184TA>CT) due to a gene conversion event arising from recombination between *SBDS* and an adjacent pseudogene *SBDSP1*, markedly reducing functional SBDS protein

expression.[3,4] This reduction in SBDS expression impairs the tight coupling of EFL1 GTP hydrolysis on the 60S subunit to eIF6 release, promoting eIF6 retention, impaired ribosome assembly and reduced translation.[21] In addition, mutations in genes such as *EFL1,*[23–25] *DNAJC21*[26,27] and *SRP54*[28–30] cause an SDS-like phenotype, consistent with their role in ribosome assembly (*EFL1, DNAJC21*) or the recruitment of signal peptide-containing proteins to the endoplasmic reticulum for secretion or membrane insertion (*SRP54*). All of these mutations lead to the striking defect in protein synthesis observed in cells from individuals with SDS.[2,22,25,31,32] Translational dysregulation due to loss of SBDS function may contribute to the development of neutropenia in SDS.[33–36]

Somatic genetic rescue of defective ribosome assembly in SDS

Somatic genetic rescue occurs in inherited disorders when an acquired somatic mutation or gross chromosomal change partially or completely offsets the deleterious effects of a germline mutation.[12] This can take the form of removal or correction of the mutation itself, or modification of other target genes such that the pathways affected by the germline mutation are altered. While these changes benefit the survival and function of individual cells, they may in fact have a neutral or detrimental impact on further disease development. First reported in a patient with severe combined immunodeficiency,[37] somatic genetic rescue mechanisms have been described in a number of IBMFSs including FA, SDS, SAMD9/9L syndromes, DC and DBA.[12] In SAMD9/9L syndromes, gain-of-function mutations in the *SAMD9* and *SAMD9L* genes lead to a general decrease in cellular proliferative capacity.[38] In this context, one of the most common rescue mechanisms involves the loss of chromosome 7 which reduces expression of the mutant allele in hematopoietic cells.[38–43] While this may be beneficial in the short-term, these changes have also been linked to an increased likelihood of developing MDS and AML.[44,45]

The p53 protein has been implicated as a key player in driving the clinical features of several ribosomopathies including SDS, Treacher Collins syndrome, DC, DBA, North American Indian Childhood Cirrhosis and 5q- syndrome[36,46–51]. Impaired ribosome assembly stabilizes p53 via the nucleolar surveillance pathway[52]. Thus, somatic mutations that reduce p53 activation could increase cellular fitness either by directly inactivating *TP53* or by rescuing the underlying ribosome maturation defect. In keeping with this hypothesis, significantly more *TP53* loss-of-function mutations were detected in patients with biallelic *SBDS* mutations compared to those with one or no *SBDS* mutations or other forms of neutropenia.[53] Mutations in eIF6 suppress the fitness defect of yeast cells lacking the SBDS homologue.[19] It was therefore striking to find that in SDS, hematopoietic cells' common compensatory mechanisms include diverse mosaic somatic genetic events such as missense mutations in *EIF6*, reciprocal chromosomal translocation or interstitial deletions of chromosome 20 (del(20)q)

that encompass the *EIF6* gene[11,54], either reducing eIF6 expression or disrupting its interaction with the 60S subunit to confer a selective advantage over non-modified cells[31,32]. Somatic genetic rescue mutations have also been described in *EFL1* mutated patients.[55] If acquired in HSCs, the contribution of an individual stem cell to hematopoiesis alters over time if these somatic genetic rescue mutations and chromosomal changes enhance clonal fitness. Understanding the dynamic behavior of clones carrying specific somatic mutations is therefore vital to determine which patients are at greater risk of developing leukemia and how these blood cancers may differ from leukemia that occurs later in adulthood.

New approaches to addressing unanswered questions

Despite a general understanding of which somatic rescue mutations are preferentially acquired in *SBDS*-mutated hematopoietic cells, little is known about the nature and dynamics of those mutations and whether multiple mutations occur in the same or different clones.[10,32,56] Recent efforts using new technologies have been undertaken to tie mutations to specific clones and to assess clonal dynamics. While next generation sequencing (NGS) technologies and targeted sequencing panels have revolutionized diagnostic medicine within the field of hematology, the high error rate of traditional NGS limits its use in the identification of mutant subclones, especially at lower allele burdens.[57] While digital droplet PCR is a sensitive and cost-effective way of monitoring low variant allele fraction (VAF) clones, it requires prior knowledge of mutations.[58] In the last decade, duplex sequencing approaches have enabled the identification of very low frequency somatic mutations from heterogeneous cell populations by using specialized tags and sequencing both strands of DNA to greatly reduce the error rate.[57,59] This technology is able to reliably detect a single somatic mutation in $>10^7$ sequenced bases.[60] An even more recent approach (NanoSeq) further reduced error rates down to <5 errors per billion bases sequenced.[61] Once generated, whole genome sequencing (WGS) data can then be used to interrogate features such as intratumoral heterogeneity.[62] In addition to detection of single nucleotide variants (SNVs), technologies such as optical genome mapping provide detailed information on structural variants (SVs).[63] While these advances have allowed for single-molecule detection of SNVs, high resolution mapping of SVs, and potential longitudinal monitoring of very low VAF mutant clones, none of these bulk sequencing technologies can be used to identify new clones with extremely low VAF mutations or to comprehensively map the origins of, or relationships between, different HSC clones.

One of the few ways in which unperturbed human HSC relatedness can be studied involves the use of naturally occurring somatic mutations as a genetic barcode.[64,65] In brief, this technique tracks the spontaneously acquired SNVs that occur in an organism through normal cellular processes that lead to DNA damage (e.g., cell division, mutagen exposure, etc). Human HSCs acquire these mutations

linearly over time, at a rate of approximately 14-18 SNVs and 0.65-0.77 insertions/deletions (indels) each year.[65–67] The majority of these changes are neutral with no effect on clonal fitness or HSC function and the patterns of shared and unique SNVs can be used to create phylogenetic trees (Figure 1A-B). These diagrams visually depict HSC clonal relationships, with branch points (termed 'coalescences') representing historic stem cell self-renewal divisions where one HSC gave rise to two daughter HSCs, each of whose progeny can be distinguished through mutational profiling of shared and unique mutations (Figure 1B). This approach has recently been used to study both normal and diseased states.[56,64,67–71] In healthy adults under the age of 70, the HSC pool is diverse and highly polyclonal, with many HSCs contributing to hematopoiesis (Figure 1C).[67] In elderly individuals over the age of 70, clonal expansions become much more prevalent with a relatively limited pool of stem cells supporting blood production (Figure 1D).[67] Tree structures become even more skewed towards dominant clones in the context of blood cancers (Figure 1E).[68] In addition to detailing HSC relatedness, clonal information can also be used to estimate overall mutation burden and stem cell numbers, identify specific mutational processes that might drive SNV acquisition (mutational signatures), and detect and estimate the timing of specific and sometimes unexpected driver mutations in expanded clones (Figure 2).[71] Since the above described phylogenetic approach relies on creating numerous WGS libraries per patient to map and track large numbers of competing clones, it can become prohibitively expensive. Additionally, limitations to this approach exist. For example, the assay relies on the expansion of single hematopoietic stem and progenitor cells (HSPCs), either in liquid culture or in Methocult medium, to provide enough genetic material for sequencing. The efficiency of colony growth in liquid culture from FACS-sorted HSPCs may be ~70%.[67] However, depending on the cell source (sorted versus enriched), this number can be lower since more differentiated cells will not produce colonies. This may also be disease- or patient-specific, with some IBMFS samples growing poorly *in vitro* and it is possible that some rare mutant clones will be missed as a result. As such, alternative ways to map clonal dynamics have been pursued, including the use of mitochondrial DNA mutations rather than genome-wide SNVs as molecular barcodes.[72] Another option is to use single-cell long-read WGS,[73] but a range of technical challenges exist that currently limit its use in establishing lineage relationships.[74] Therefore, while targeted sequencing panels, especially those utilizing high sensitivity duplex sequencing, will become a powerful clinical tool for tracking mutations (and their respective clones) in patients, if clonal origins and/or lineage relationships are required, then whole genome approaches remain the only robust method.

Mapping relationships between HSCs in SDS

Using the powerful method of SNV profiling and HSC phylogeny building in the context of SDS, Machado *et al.* recently mapped HSC relatedness and assessed somatic genetic rescue mechanisms in

samples taken from ten individuals with biallelic germline loss-of-function mutations in *SBDS*.[56] WGS was performed on DNA isolated from 323 individually expanded single hematopoietic stem and progenitor cells (HSPCs) obtained from either the peripheral blood (PB) or bone marrow (BM) of patients. For seven of the ten patients (age range 4-33 years old), substantial clonal expansions were observed, as evidenced by clustering of the sampled HSPCs into large clades (Figure 1F). Phylogenetic tree structures like this are typically only observed much later in life[67] which suggests that for patients with SDS, HSC diversity and clonal architecture is profoundly altered within the first few decades of life. Whether reduced polyclonality is observed in the HSC pools of patients with other IBMFSs remains to be seen, especially since this feature was not uniformly observed in all patients studied by Machado *et al.*

The mutational information generated by WGS also identifies specific somatic mutations acquired by individual ancestral HSCs. Machado *et al.* found that *TP53* was the most commonly mutated gene in individual HSCs which accorded with data from larger cohorts of SDS patients reported by Kennedy *et al.*[32] and Xie *et al*.,[53] where 45% and 48% of SDS patients harbored a *TP53* mutation, respectively. The Kennedy study also reported that ~60% of patients carried *EIF6* mutations, a finding reinforced by both the Machado et al. and Tan *et al*. studies.[31] Beyond these most common mutations, Machado *et al*. also detected SNVs in genes reported as recurrently mutated in SDS patients (e.g. *PRPF8*), while also for the first time in SDS identifying mutations in ribosomal protein genes *RPL5* (encoding uL18) and *RPL22* (encoding eL22). Although still to be established, these mutations may decrease p53 activation by reducing signaling through the nucleolar stress pathway, potentially diminishing the drive to select *TP53* mutations, thereby lowering the risk of disease progression. Together, this approach allows researchers to comprehensively map the landscape of potential mutations that rescue the SBDS-associated ribosome deficiency, but do not compromise cell viability or lead to frequent leukemic transformation. Since every single cell in SDS is under evolutionary pressure to rescue the SBDS-dependent ribosome deficiency, the exploration of possible genetic rescue mechanisms across multiple patients is extensive. Mutations that are under positive selection in SDS lead to improved cellular fitness and the persistence of clones bearing those mutations.

Whether or not individual clones need single or multiple somatic genetic rescue events and whether mutations involved in clonal hematopoiesis (CH) co-occur with rescue mutations is less clear. Across both the Kennedy and Machado studies, CH itself was detected in up to 72% of SDS patients.[32,56] Most of these expansions were associated with somatic genetic rescue mutations rather than typical age-related CH mutations in genes such as *DNMT3A*, *TET2* and *ASXL1*.[32,56] By sequencing serial blood samples taken from 49 SDS patients with detectable CH, Kennedy *et al*. demonstrated that many of these driver mutations were consistently seen across multiple samples and the VAFs of these mutations remained low and stable over time.[32] While this is a powerful approach to follow clonal

evolution, it is limited by the availability of longitudinal samples, especially those taken prior to SDS diagnosis. By using somatic mutations as natural barcodes, the tree building approach skirts this pitfall because it concomitantly reports the historical origins of the mutation (and therefore the growth rates over the lifetime of the individual) as the sequencing information collected from HSPCs reflects the mutational state of ancestral HSCs. Regarding co-occurrence of CH mutations and gene rescue events, the Machado *et al.* study provides evidence in single clones that points to mutual exclusivity in most cases (including SDS-related or CH-related mutations such as those affecting *DNMT3A*, *ASXL1*, *TET2* and *RUNX1*). The two exceptions to this were (1) a clone with a combination of *TET2* and *TP53* mutations and (2) a patient with biallelic *TP53* mutations who went on to develop MDS with acquisition of multiple copy number alterations and mutations. This, in combination with the observation that the individual genome-wide mutation burden was not increased in *TP53* mono-allelic mutant SDS patients, suggests that heterozygous somatic genetic rescue events were sufficient to restore cellular fitness and compensate for the SBDS deficiency. This conclusion is also consistent with the study by Kennedy *et al.* where they first used bulk whole exome sequencing to identify patients with multiple *EIF6* and *TP53* mutations and then employed a single cell DNA genotyping approach to show that the vast majority of these mutations were acquired alone in distinct clones.[32] In both studies, the one important exception was biallelic *TP53* mutations that were associated with genomic instability and progression to myeloid malignancy.[32,56]

Timing of mutation acquisition

One of the most unique aspects of the phylogenetic tree building approach is its ability to estimate when in an individual's lifetime specific driver mutations were acquired and when specific clonal expansions began. Previous approaches to timing SNVs have involved serial sampling and sequencing to look for earliest detection of low variant allele fraction (VAF) mutations. There are several limitations to this approach. If standard bulk sequencing is used, mutations towards the lower end of detection (e.g. mutations at low VAF) are likely to be missed. Therefore, a single cell approach using serial samples as performed by Kennedy *et al.* would be the best option to yield high quality information.[32] However, this approach is constrained by the availability of serial samples. When using the phylogeny approach, a single sample time point captures information about sequential SNV acquisition in ancestral HSCs over time. Williams *et al.* were able to demonstrate that *JAK2*^{V617F} mutations in myeloproliferative neoplasms were acquired early in life (between 33 weeks of gestation and 11 years old) despite the absence of disease symptoms for many decades[68] and Machado *et al.* were similarly able to estimate that clonal expansions with acquired driver mutations began as early as *in utero* up to 12 years of age in SDS patients. Moreover, in individuals with SDS where there is strong selective pressure to rescue the ribosomal defect, most mutations could be traced back to the

first decade of life and, interestingly, there were no observed patterns in terms of when mutations in specific genes were acquired (including several CH-related mutations). While the sampling method itself is biased toward clones that have had a longer time to expand, it is worth noting that studies in normal aging have shown a much more protracted period of driver mutation acquisition compared to SDS patients.[67]

Implications for patient monitoring

Given the high risk of progression to myeloid cancers for patients with SDS and the dismal outcomes of patients with pre-existing *TP53* mutations even after allogeneic bone marrow transplant[75], the retrospective study of patients who have developed MDS/AML is key to developing new monitoring and treatment strategies. In both the Kennedy *et al.* (whole exome sequencing) and Machado *et al.* (WGS) studies, several driver mutations were identified in SDS patients that can now be investigated for their impact on disease progression. Together, these studies identified recurrent mutations in *TP53*, *EIF6*, *CSNK1A1*, *PRPF8*, *RPL5* and *RPL22* and as well as recurrent chromosomal abnormalities in chromosomes 3, 5, 7 and 20.[32,56] While mutations in genes such as *EIF6*, *CSNK1A1*, and *PRPF8* were not found to be associated with blood cancer development, homozygous *TP53* mutations were observed in all those SDS patients included in the study who had progressed to MDS or AML compared to those who had not. Notably, patients in the Kennedy study who had not yet progressed to cancer still had detectable heterozygous *TP53* mutations and this accords with no increased mutation burden in *TP53* heterozygous clones reported by Machado et al. Together, these data indicate that presence of a heterozygous *TP53* mutation alone does not predict that an individual will go on to develop MDS/AML and suggests that this parameter by itself is of limited clinical value in determining appropriateness for transplantation.

Also of interest, one of the patients in the Machado *et al*. study who had acquired biallelic *TP53* mutations in the context of MDS transformation, showed substantial genome instability as evidenced by a multiple copy number aberrations including $ch5^-$, $6p^-$, 11^+, 18^- and X^-. This was further associated with an increased overall mutation burden, more than double that expected for age. *TP53*-mutated clones contained a higher total mutation burden than expected based on age and an increased proportion of the mutational signature 'SBS1', linked to rapid cell division, which notably was substantially lower in clones that pre-dated the most recently branched HSC clones, suggesting that the increase in SBS1 signature-related mutations was associated with the onset of MDS. In addition, the estimated clonal growth rate of 5200% per year for the biallelic *TP53* mutant clones (doubling roughly every 2 months), reinforces the requirement for frequent surveillance of SDS patients.[56] NGS results from BM and peripheral blood appear to be highly concordant, at least in the context of

hematological malignancy.[76] These data support the use of a combination of array comparative genomic hybridization and myeloid NGS performed on serial peripheral blood samples (perhaps 6-12 monthly), in conjunction with frequent (3-4 monthly) full blood count sampling as a potential surveillance strategy for SDS patients. That being said, the detection of mutations in IBMFS patients with severe peripheral blood cytopenias may prove to be challenging and requires further investigation. Bone marrow examination would be advisable if there is a significant alteration in blood counts or if rising VAFs for specific variants (especially in the *TP53* gene) were detected.

Future perspectives

The advent of the phylogeny building approach has begun to open new areas of research into the clonal dynamics of mutated HSCs during CH and disease progression, revealing the most common recurrent mutations and their relative risk with respect to clonal dominance and/or disease evolution. While the above section details its powerful use in a specific inherited bone marrow failure disorder, SDS, to comprehensively map the entire complement of genetic rescue mechanisms and how these might impact clinical evolution, the technology can be readily applied to other IBMFs and also to diseases driven entirely by somatic mutations such as VEXAS (vacuoles, E1 enzyme, X-linked, autoinflammatory, somatic) syndrome[77]. Once mapped, these data can be used in two main ways for clinical benefit: (1) to aid the design of targeted panels that cheaply and efficiently capture the mutational profile of an individual patient; and (2) to identify “adaptive” and “maladaptive” rescue mechanisms that may influence clinical outcomes. In both cases, this information can inform clinical decision-making and risk stratification. This is critically important since retrospective analysis of data in the North American SDS registry suggest a median survival of 7.7 years after MDS diagnosis and 1 year after AML diagnosis.[8]

The idea that inherited disorders that induce strong selective pressure will end up as composites of hundreds to thousands of differently mutated cells prompts us to also imagine how to use this information to help in clinical management. Could clones carrying adaptive mutations that restore ribosome homeostasis and are associated with low risk of transformation be “boosted” in some manner or those with high risk of transformation be targeted? Given the high frequency of *TP53* mutation acquisition in SDS HSCs, this prompts the challenge of how best to detect *TP53* mutations (particularly biallelic *TP53* mutations) in the clinical setting to inform decision-making (e.g., proceeding to transplant). Similarly, when high depth, high sensitivity methods permit reliable detection of biallelic *TP53* loss in clinical assays, how might this impact prognosis? One clear method would be to sample patients longitudinally to understand whether individually mutated clones are growing rapidly over time, but this approach could be costly and time-consuming for the researcher

and the patient and may be further confounded by new clones emerging with a stronger selective growth advantage. A potential route to avoid this latter challenge is to develop dedicated SDS/myeloid panels that include all the mutations described above in targeted panel design so that newly emerging clones bearing key SDS driver mutations can be tracked.

One thing is clear from the Kennedy and Machado studies - the complexity of the mutational landscape in SDS is enormous. This presents particular challenges for therapeutic approaches such as corrective gene therapy for a number of reasons. Firstly, the somatic rescue mutations that have arisen to compensate for the absence of SBDS are already present from a very early age. Initiatives like the Newborn Genomes Programme (https://www.genomicsengland.co.uk/initiatives/newborns) are investigating the potential benefits of early patient diagnosis. In the UK, the Generation Study will sequence the genomes of 100,000 newborn babies in England in partnership with the National Health Service to understand whether improvements can be made in diagnosing and treating genetic conditions. The full complement of causative genes in SDS patients are included in this study. In future, it also might become possible to screen for cancer-associated mutations as these may be acquired in utero.[56,68] Secondly, correcting a specific genetic change (e.g., mutant *TP53*) will still leave behind a wide range of other mutated HSC clones, so any gene therapy "fix" may only be temporary. One approach might involve selectively steering clonal dynamics to promote the outgrowth of adaptive clones (e.g., specific *EIF6*, *RPL22* or *RPL5* mutations, etc.) or restrict the development, growth or persistence of biallelic (or indeed heterozygous) *TP53* mutant clones.

The recent new biological insights into the molecular mechanisms underpinning SDS and evolution to MDS and AML generated using these new technologies highlight the importance of routinely integrating serial next generation sequencing and microarray analysis into hematological surveillance strategies for patients. Tools to study the relative competitive ability and potential leukemogenicity of individual mutant clones are not currently well-developed and the WGS approach is not practical to undertake at a large scale. Targeted sequencing panels may be useful to detect and track the most common genes/mutations and variants when they reach a clonal burden of sufficient size, but understanding the mechanism of each mutation will require more sophisticated genetic tools and functional studies. As most myeloid malignancies in SDS involve the acquisition of biallelic *TP53* mutations, there is an urgent need to develop robust tools to detect such clones at an early stage, at very low allele frequency, prior to progression to overt MDS or AML, to allow for early consideration of HSC transplantation. In this context, it also becomes vital to understand whether early transplant for SDS patients who have acquired biallelic *TP53* mutations will improve the otherwise poor long term outcomes.[8,75,78,79] A better understanding of the consequences of somatic mutation acquisition on clonal dynamics in SDS may also guide the development of therapeutics that mimic the effects of adaptive mutations in restoring ribosome homeostasis, while avoiding the potential leukemic risk of

heterozygous *TP53* mutation acquisition and the leukemogenic potential of such clones if they arise. More generally, the identification of somatic rescue mutations across IBMFSs offers the potential to identify new targets for the design of disease modifying therapeutics.

References

1. Alter BP, Giri N, Savage SA, Rosenberg PS. Cancer in the National Cancer Institute inherited bone marrow failure syndrome cohort after fifteen years of follow-up. *Haematologica*. 2018;103(1):30–39.
2. Warren AJ. Molecular basis of the human ribosomopathy Shwachman-Diamond syndrome. *Adv. Biol. Regul.* 2018;67:109–127.
3. Boocock GRB, Morrison JA, Popovic M, et al. Mutations in SBDS are associated with Shwachman-Diamond syndrome. *Nat. Genet.* 2003;33(1):97–101.
4. Woloszynek JR, Rothbaum RJ, Rawls AS, et al. Mutations of the SBDS gene are present in most patients with Shwachman-Diamond syndrome. *Blood*. 2004;104(12):3588–3590.
5. Dror Y, Durie P, Ginzberg H, et al. Clonal evolution in marrows of patients with Shwachman-Diamond syndrome: a prospective 5-year follow-up study. *Exp. Hematol.* 2002;30(7):659–669.
6. Donadieu J, Fenneteau O, Beaupain B, et al. Classification of and risk factors for hematologic complications in a French national cohort of 102 patients with Shwachman-Diamond syndrome. *Haematologica*. 2012;97(9):1312–1319.
7. Cesaro S, Pegoraro A, Sainati L, et al. A Prospective Study of Hematologic Complications and Long-Term Survival of Italian Patients Affected by Shwachman-Diamond Syndrome. *J. Pediatr.* 2020;219:196–201.e1.
8. Myers KC, Furutani E, Weller E, et al. Clinical features and outcomes of patients with Shwachman-Diamond syndrome and myelodysplastic syndrome or acute myeloid leukaemia: a multicentre, retrospective, cohort study. *Lancet Haematol*. 2020;7(3):e238–e246.
9. Furutani E, Liu S, Galvin A, et al. Hematologic complications with age in Shwachman-Diamond syndrome. *Blood Adv*. 2022;6(1):297–306.
10. Bernard E, Nannya Y, Hasserjian RP, et al. Implications of TP53 allelic state for genome stability, clinical presentation and outcomes in myelodysplastic syndromes. *Nat. Med.* 2020;26(10):1549–1556.
11. Valli R, Pressato B, Marletta C, et al. Different loss of material in recurrent chromosome 20 interstitial deletions in Shwachman-Diamond syndrome and in myeloid neoplasms. *Mol. Cytogenet.* 2013;6(1):56.
12. Revy P, Kannengiesser C, Fischer A. Somatic genetic rescue in Mendelian haematopoietic diseases. *Nat. Rev. Genet.* 2019;20(10):582–598.
13. Reilly CR, Shimamura A. Predisposition to myeloid malignancies in Shwachman-Diamond syndrome: biological insights and clinical advances. *Blood*. 2023;141(13):1513–1523.
14. Dragon F, Gallagher JEG, Compagnone-Post PA, et al. A large nucleolar U3 ribonucleoprotein required for 18S ribosomal RNA biogenesis. *Nature*. 2002;417(6892):967–970.
15. Grandi P, Rybin V, Bassler J, et al. 90S pre-ribosomes include the 35S pre-rRNA, the U3 snoRNP, and 40S subunit processing factors but predominantly lack 60S synthesis factors. *Mol. Cell*. 2002;10(1):105–115.
16. Kornprobst M, Turk M, Kellner N, et al. Architecture of the 90S Pre-ribosome: A Structural View on the Birth of the Eukaryotic Ribosome. *Cell*. 2016;166(2):380–393.
17. Chaker-Margot M, Barandun J, Hunziker M, Klinge S. Architecture of the yeast small subunit processome. *Science*. 2017;355:6321.
18. Sun Q, Zhu X, Qi J, et al. Molecular architecture of the 90S small subunit pre-ribosome. *Elife*. 2017;6:e22086.
19. Menne TF, Goyenechea B, Sánchez-Puig N, et al. The Shwachman-Bodian-Diamond syndrome protein mediates translational activation of ribosomes in yeast. *Nat. Genet.* 2007;39(4):486–495.
20. Finch AJ, Hilcenko C, Basse N, et al. Uncoupling of GTP hydrolysis from eIF6 release on the

ribosome causes Shwachman-Diamond syndrome. *Genes Dev.* 2011;25(9):917–929.
21. Weis F, Giudice E, Churcher M, et al. Mechanism of eIF6 release from the nascent 60S ribosomal subunit. *Nat. Struct. Mol. Biol.* 2015;22(11):914–919.
22. Jaako P, Faille A, Tan S, et al. eIF6 rebinding dynamically couples ribosome maturation and translation. *Nat. Commun.* 2022;13(1):1562.
23. Stepensky P, Chacón-Flores M, Kim KH, et al. Mutations in EFL1, an SBDS partner, are associated with infantile pancytopenia, exocrine pancreatic insufficiency and skeletal anomalies in aShwachman-Diamond like syndrome. *J. Med. Genet.* 2017;54(8):558–566.
24. Tan QK-G, Cope H, Spillmann RC, et al. Further evidence for the involvement of EFL1 in a Shwachman-Diamond-like syndrome and expansion of the phenotypic features. *Cold Spring Harb Mol Case Stud*. 2018;4(5):a003046.
25. Tan S, Kermasson L, Hoslin A, et al. EFL1 mutations impair eIF6 release to cause Shwachman-Diamond syndrome. *Blood*. 2019;134(3):277–290.
26. Tummala H, Walne AJ, Williams M, et al. DNAJC21 Mutations Link a Cancer-Prone Bone Marrow Failure Syndrome to Corruption in 60S Ribosome Subunit Maturation. *Am. J. Hum. Genet.* 2016;99(1):115–124.
27. Dhanraj S, Matveev A, Li H, et al. Biallelic mutations in DNAJC21 cause Shwachman-Diamond syndrome. *Blood*. 2017;129(11):1557–1562.
28. Bellanné-Chantelot C, Schmaltz-Panneau B, Marty C, et al. Mutations in the SRP54 gene cause severe congenital neutropenia as well as Shwachman-Diamond-like syndrome. *Blood*. 2018;132(12):1318–1331.
29. Carapito R, Konantz M, Paillard C, et al. Mutations in signal recognition particle SRP54 cause syndromic neutropenia with Shwachman-Diamond-like features. *J. Clin. Invest.* 2017;127(11):4090–4103.
30. Juaire KD, Lapouge K, Becker MMM, et al. Structural and Functional Impact of SRP54 Mutations Causing Severe Congenital Neutropenia. *Structure*. 2021;29(1):15–28.e7.
31. Tan S, Kermasson L, Hilcenko C, et al. Somatic genetic rescue of a germline ribosome assembly defect. *Nat. Commun.* 2021;12(1):5044.
32. Kennedy AL, Myers KC, Bowman J, et al. Distinct genetic pathways define pre-malignant versus compensatory clonal hematopoiesis in Shwachman-Diamond syndrome. *Nat. Commun.* 2021;12(1):1334.
33. In K, Zaini MA, Müller C, et al. Shwachman-Bodian-Diamond syndrome (SBDS) protein deficiency impairs translation re-initiation from C/EBPα and C/EBPβ mRNAs. *Nucleic Acids Res.* 2016;44(9):4134–4146.
34. Rawls AS, Gregory AD, Woloszynek JR, Liu F, Link DC. Lentiviral-mediated RNAi inhibition of Sbds in murine hematopoietic progenitors impairs their hematopoietic potential. *Blood*. 2007;110(7):2414–2422.
35. Orelio C, Verkuijlen P, Geissler J, van den Berg TK, Kuijpers TW. SBDS expression and localization at the mitotic spindle in human myeloid progenitors. *PLoS One*. 2009;4(9):e7084.
36. Zambetti NA, Bindels EMJ, Van Strien PMH, et al. Deficiency of the ribosome biogenesis gene Sbds in hematopoietic stem and progenitor cells causes neutropenia in mice by attenuating lineage progression in myelocytes. *Haematologica*. 2015;100(10):1285–1293.
37. Hirschhorn R, Yang DR, Israni A, Huie ML, Ownby DR. Somatic mosaicism for a newly identified splice-site mutation in a patient with adenosine deaminase-deficient immunodeficiency and spontaneous clinical recovery. *Am. J. Hum. Genet.* 1994;55(1):59–68.
38. Buonocore F, Kühnen P, Suntharalingham JP, et al. Somatic mutations and progressive monosomy modify SAMD9-related phenotypes in humans. *J. Clin. Invest.* 2017;127(5):1700–1713.
39. Narumi S, Amano N, Ishii T, et al. SAMD9 mutations cause a novel multisystem disorder, MIRAGE syndrome, and are associated with loss of chromosome 7. *Nat. Genet.* 2016;48(7):792–797.
40. Tesi B, Davidsson J, Voss M, et al. Gain-of-function SAMD9L mutations cause a syndrome of cytopenia, immunodeficiency, MDS, and neurological symptoms. *Blood*. 2017;129(16):2266–2279.
41. Pastor VB, Sahoo SS, Boklan J, et al. Constitutional SAMD9L mutations cause familial

myelodysplastic syndrome and transient monosomy 7. *Haematologica*. 2018;103(3):427–437.
42. Wong JC, Bryant V, Lamprecht T, et al. Germline SAMD9 and SAMD9L mutations are associated with extensive genetic evolution and diverse hematologic outcomes. *JCI Insight*. 2018;3(14):e121086.
43. Sahoo SS, Pastor VB, Goodings C, et al. Clinical evolution, genetic landscape and trajectories of clonal hematopoiesis in SAMD9/SAMD9L syndromes. *Nat. Med.* 2021;27(10):1806–1817.
44. Davidsson J, Puschmann A, Tedgård U, et al. SAMD9 and SAMD9L in inherited predisposition to ataxia, pancytopenia, and myeloid malignancies. *Leukemia*. 2018;32(5):1106–1115.
45. Inaba T, Honda H, Matsui H. The enigma of monosomy 7. *Blood*. 2018;131(26):2891–2898.
46. Elghetany MT, Alter BP. p53 protein overexpression in bone marrow biopsies of patients with Shwachman-Diamond syndrome has a prevalence similar to that of patients with refractory anemia. *Arch. Pathol. Lab. Med.* 2002;126(4):452–455.
47. Sieff CA, Yang J, Merida-Long LB, Lodish HF. Pathogenesis of the erythroid failure in Diamond Blackfan anaemia. *Br. J. Haematol.* 2010;148(4):611–622.
48. Jaako P, Debnath S, Olsson K, et al. Disruption of the 5S RNP-Mdm2 interaction significantly improves the erythroid defect in a mouse model for Diamond-Blackfan anemia. *Leukemia*. 2015;29(11):2221–2229.
49. Frattini A, Bolamperti S, Valli R, et al. Enhanced p53 Levels Are Involved in the Reduced Mineralization Capacity of Osteoblasts Derived from Shwachman-Diamond Syndrome Subjects. *Int. J. Mol. Sci.* 2021;22(24):13331.
50. Barlow JL, Drynan LF, Hewett DR, et al. A p53-dependent mechanism underlies macrocytic anemia in a mouse model of human 5q- syndrome. *Nat. Med.* 2010;16(1):59–66.
51. Tourlakis ME, Zhang S, Ball HL, et al. In Vivo Senescence in the Sbds-Deficient Murine Pancreas: Cell-Type Specific Consequences of Translation Insufficiency. *PLoS Genet.* 2015;11(6):e1005288.
52. Sloan KE, Bohnsack MT, Watkins NJ. The 5S RNP couples p53 homeostasis to ribosome biogenesis and nucleolar stress. *Cell Rep.* 2013;5(1):237–247.
53. Xia J, Miller CA, Baty J, et al. Somatic mutations and clonal hematopoiesis in congenital neutropenia. *Blood*. 2018;131(4):408–416.
54. Valli R, Minelli A, Galbiati M, et al. Shwachman-Diamond syndrome with clonal interstitial deletion of the long arm of chromosome 20 in bone marrow: haematological features, prognosis and genomic instability. *Br. J. Haematol.* 2019;184(6):974–981.
55. Lee S, Shin CH, Lee J, et al. Somatic uniparental disomy mitigates the most damaging EFL1 allele combination in Shwachman-Diamond syndrome. *Blood*. 2021;138(21):2117–2128.
56. Machado HE, Øbro NF, Williams N, et al. Convergent somatic evolution commences in utero in a germline ribosomopathy. *Nat. Commun.* 2023;14(1):1–14.
57. Kennedy SR, Schmitt MW, Fox EJ, et al. Detecting ultralow-frequency mutations by Duplex Sequencing. *Nat. Protoc.* 2014;9(11):2586–2606.
58. Ip BBK, Wong ATC, Law JHY, et al. Application of droplet digital PCR in minimal residual disease monitoring of rare fusion transcripts and mutations in haematological malignancies. *Sci. Rep.* 2024;14(1):6400.
59. Salk JJ, Schmitt MW, Loeb LA. Enhancing the accuracy of next-generation sequencing for detecting rare and subclonal mutations. *Nat. Rev. Genet.* 2018;19(5):269–285.
60. Schmitt MW, Kennedy SR, Salk JJ, et al. Detection of ultra-rare mutations by next-generation sequencing. *Proc. Natl. Acad. Sci. U. S. A.* 2012;109(36):14508–14513.
61. Abascal F, Harvey LMR, Mitchell E, et al. Somatic mutation landscapes at single-molecule resolution. *Nature*. 2021;593(7859):405–410.
62. Dentro SC, Leshchiner I, Haase K, et al. Characterizing genetic intra-tumor heterogeneity across 2,658 human cancer genomes. *Cell*. 2021;184(8):2239–2254.e39.
63. Dremsek P, Schwarz T, Weil B, et al. Optical Genome Mapping in Routine Human Genetic Diagnostics-Its Advantages and Limitations. *Genes*. 2021;12(12):1958.
64. Lee-Six H, Øbro NF, Shepherd MS, et al. Population dynamics of normal human blood inferred from somatic mutations. *Nature*. 2018;561(7724):473–478.
65. Osorio FG, Rosendahl Huber A, Oka R, et al. Somatic Mutations Reveal Lineage Relationships and Age-Related Mutagenesis in Human Hematopoiesis. *Cell Rep.* 2018;25(9):2308–2316.e4.

66. Welch JS, Ley TJ, Link DC, et al. The origin and evolution of mutations in acute myeloid leukemia. *Cell*. 2012;150(2):264–278.
67. Mitchell E, Spencer Chapman M, Williams N, et al. Clonal dynamics of haematopoiesis across the human lifespan. *Nature*. 2022;606(7913):343–350.
68. Williams N, Lee J, Mitchell E, et al. Life histories of myeloproliferative neoplasms inferred from phylogenies. *Nature*. 2022;602(7895):162–168.
69. Fabre MA, de Almeida JG, Fiorillo E, et al. The longitudinal dynamics and natural history of clonal haematopoiesis. *Nature*. 2022;606(7913):335–342.
70. Spencer Chapman M, Cull AH, Ciuculescu MF, et al. Clonal selection of hematopoietic stem cells after gene therapy for sickle cell disease. *Nat. Med.* 2023;29(12):3175–3183.
71. Lee-Six H, Kent DG. Tracking hematopoietic stem cells and their progeny using whole-genome sequencing. *Exp. Hematol.* 2020;83:12–24.
72. Miller TE, Lareau CA, Verga JA, et al. Mitochondrial variant enrichment from high-throughput single-cell RNA sequencing resolves clonal populations. *Nat. Biotechnol.* 2022;40(7):1030–1034.
73. Hård J, Mold JE, Eisfeldt J, et al. Long-read whole-genome analysis of human single cells. *Nat. Commun.* 2023;14(1):5164.
74. Gawad C, Koh W, Quake SR. Single-cell genome sequencing: current state of the science. *Nat. Rev. Genet.* 2016;17(3):175–188.
75. Lindsley RC, Saber W, Mar BG, et al. Prognostic Mutations in Myelodysplastic Syndrome after Stem-Cell Transplantation. *N. Engl. J. Med.* 2017;376(6):536–547.
76. Jansko-Gadermeir B, Leisch M, Gassner FJ, et al. Myeloid NGS Analyses of Paired Samples from Bone Marrow and Peripheral Blood Yield Concordant Results: A Prospective Cohort Analysis of the AGMT Study Group. *Cancers*. 2023;15(8):2305.
77. Beck DB, Ferrada MA, Sikora KA, et al. Somatic Mutations in UBA1 and Severe Adult-Onset Autoinflammatory Disease. *N. Engl. J. Med.* 2020;383(27):2628–2638.
78. Cesaro S, Pillon M, Sauer M, et al. Long-term outcome after allogeneic hematopoietic stem cell transplantation for Shwachman–Diamond syndrome: a retrospective analysis and a review of the literature by the Severe Aplastic Anemia Working Party of the European Society for Blood and Marrow Transplantation (SAAWP-EBMT). *Bone Marrow Transplant.* 2020;55(9):1796–1809.
79. Myers K, Hebert K, Antin J, et al. Hematopoietic Stem Cell Transplantation for Shwachman-Diamond Syndrome. *Biol. Blood Marrow Transplant.* 2020;26(8):1446–1451.

Figure captions

Figure 1. Using somatic mutations to understand historic HSC clonal relationships. (A) A simplified phylogenetic tree of HSC relatedness. Based on the mutations detected in each stem cell by WGS, we can build a "family tree" of HSCs using unique and shared somatic mutations. Shapes within the circles represent unique somatic mutations that are inherited by a cell's progeny. (B) A hypothetical phylogenetic tree. The terminal point of each line represents the stem cell that was sampled. Traveling up these lines, any points of convergence (termed a "coalescence") indicate an ancestral HSC where two progeny stem cells and their subsequent progeny can be traced. The uppermost branch points represent stem cell divisions that likely occurred in embryogenesis. The dotted box highlights how the structure in (A) would appear within a phylogenetic tree. The presence and distribution of coalescences is related to both aging and disease state. Example tree structures representing HSC relatedness in three different contexts are shown in (C-F). (C) Phylogenies from younger healthy adults appear highly polyclonal with many HSCs contributing to the overall

structure,[67] as indicated by the back of branch points beyond early development. (D) By the time individuals reach their seventh decade of life, tree structures begin to show evidence of decreased clonal diversity,[67] with a greater number of sampled HSCs being part of large clones. These expansions are often (but not always) associated with non-pathogenic mutations that provide HSCs with a survival advantage. (E) This skewed clonal structure is even more exaggerated in the case of myeloid malignancy. In this context, pathogenic mutations like *JAK2V617F* cause large expansions of specific HSCs. (F) The phylogenetic trees of patients with SDS show evidence of clonal expansions more characteristic of older adults. Colored branches indicate which clones carry driver mutations (trees taken from Machado *et al.*[56]).

Figure 2. Information generated using the WGS approach to studying HSC dynamics. Aside from creating phylogenetic trees using the patterns of shared and unique SNVs generated using the WGS approach, a variety of different information can be mined from this type of experiment. For example, the timing of mutation acquisition can be determined for expanded clones. This potentially has significant implications for when and how myeloid cancer-associated mutations should be monitored and when patients could receive therapeutic intervention. Additionally, overall mutation burdens can be estimated and, along with information on mutational signatures, overall clonality and the range of somatic genetic rescue mutations, these data can help to build a more comprehensive picture of the pathways and genes involved in disease-related clonal dynamics. Importantly, the source of cellular material (bone marrow vs. peripheral blood) was not found to significantly influence these results. Lee-Six *et al.*[64] sorted single $Lin^-CD34^+CD38^-CD90^-CD45RA^+$ HSCs and bulk HSPCs (MEPs, GMPs, CMPs) from both peripheral blood and bone marrow of the same individual and did not observe any significant differences in mutational burden between the two compartments. Also, the phylogenetic tree built exclusively from one source versus the other was not significantly different. This is because sampled HSCs and HSPCs all report on their parent HSCs.

Acknowledgements

Work in the D.G.K. laboratory is supported by the Bill and Melinda Gates Foundation (INV-002189 and INV-038816, a Cancer Research UK Programme Foundation Award (DCRPGF\100008), the UK Medical Research Council (MC_PC_21043, MR/V005502/1, MR/Y011945/1), and the National Institute for Health and Care Research (NIHR) Leeds Biomedical Research Centre (BRC) (NIHR203331). Work in the AJW laboratory is supported by the UK MRC (MR/T012412/1), Blood Cancer UK (21002), Rosetrees Trust (PGL22/100032), Isaac Newton Trust (G125517), Addenbrookes Charitable Trust (900426), SDSUK, SDS Foundation, Butterfly Guild, Connor Wright Project, the Cambridge National Institute for Health Research Biomedical Research Centre, European

639 COST Actions CA18233 and CA21154.

640

641 Authorship and conflict-of-interest statements

642

643 A.H.C., D.G.K. and A.J.W. contributed equally to the text of this paper and A.H.C. created the
644 figures. A.J.W. is a consultant for SDS Therapeutics. The remaining authors declare no competing
645 interests.