

Clonal hematopoiesis and hematological malignancy

William G. Dunn,^{1,2} Matthew A. McLoughlin,¹ and George S. Vassiliou^{1,2}

¹Wellcome-MRC Cambridge Stem Cell Institute, University of Cambridge, Cambridge, United Kingdom. ²Department of Haematology, Addenbrooke's Hospital, Cambridge University Hospitals NHS Trust, Cambridge, United Kingdom.

Clonal hematopoiesis (CH), the expansion of hematopoietic stem cells and their progeny driven by somatic mutations in leukemia-associated genes, is a common phenomenon that rises in prevalence with advancing age to affect most people older than 70 years. CH remains subclinical in most carriers, but, in a minority, it progresses to a myeloid neoplasm, such as acute myeloid leukemia, myelodysplastic syndrome, or myeloproliferative neoplasm. Over the last decade, advances in our understanding of CH, its molecular landscape, and the risks associated with different driver gene mutations have culminated in recent developments that allow for a more precise estimation of myeloid neoplasia risk in CH carriers. In turn, this is leading to the development of translational and clinical programs to intercept and prevent CH from developing into myeloid neoplasia. Here, we give an overview of the spectrum of CH driver mutations, what is known about their pathophysiology, and how this informs the risk of incident myeloid malignancy.

Introduction

Throughout the human lifespan, blood cells are formed continuously through the division and differentiation of cells in a process termed hematopoiesis. At the apex of this hierarchy are a rare population of 50,000–200,000 multipotent hematopoietic stem cells (HSCs) (1), which give rise to more committed progenitors, the progeny of which ultimately generate all blood cell lineages. Somatic mutations are acquired throughout life by all cells, including HSCs. While most are inconsequential or sometimes detrimental, specific rare mutations can confer a fitness advantage that enables the host HSC and its mutation-bearing descendants to expand and occupy a large fraction of the HSC population, a process known as clonal hematopoiesis (CH).

The observation that some elderly healthy female individuals exhibit skewed X-chromosome inactivation patterns provided early evidence of CH (2). In the postgenomics era, several key studies demonstrated that CH is characterized by the presence of recurrent fitness-conferring “driver” mutations in a small number of genes, most of which are associated with leukemia, in individuals without any apparent blood neoplasm (3–6). This process was found to be increasingly prevalent with advancing age, affecting 10%–20% of those aged over 70 years (3, 4, 6), with a subsequent study utilizing next-generation sequencing (NGS) approaches able to detect very small clones identifying CH ubiquitously in people aged 50–60 years (7). A defining feature of CH is its association with an increased risk of incident hematological malignancy, and the magnitude of this risk varies according to the nature of the driver mutation(s), the clonal

size, and the context in which CH arises (3, 4, 8–10). This Review will consider what features define CH and how the molecular heterogeneity of CH informs the risk and type of malignant progression.

The nomenclature of CH

As our understanding of the causes and consequences of CH has grown, it has become clear that the phenomenon exists in several related but distinct forms or subtypes. The terms “clonal hematopoiesis” and “age-related clonal hematopoiesis” (ARCH) refer to any clonal expansion of HSCs, irrespective of the size of the clone. By contrast, “clonal hematopoiesis of indeterminate potential” (CHIP) describes the presence of a CH clone that has expanded to a variant allele fraction of $\geq 2\%$ (i.e., involving $\geq 4\%$ of nucleated blood cells), an arbitrary threshold chosen to signify clones large enough to be of potential clinical significance (11). The terms CH, ARCH, and CHIP usually refer to clones associated with somatic driver mutations in genes recurrently mutated in myeloid cancers, while lymphoid CHIP (L-CHIP) describes clonal expansions associated with mutations in lymphoid cancer genes (12). L-CHIP is less common than myeloid CHIP and may originate in committed lymphoid progenitors rather than HSCs. Myeloid and lymphoid clonal expansions do not only arise in association with somatic mutations in individual genes (the focus of this Review) but can also be driven by acquired chromosomal abnormalities, such as amplifications, deletions, and copy neutral loss-of-heterozygosity events. Depending on their specific nature, these mosaic chromosomal aberrations (mCAs) can be associated with an increased risk of myeloid malignancies, lymphoid malignancies, or both (12–18).

By definition, CHIP does not significantly affect blood cell counts but can progress to do so, with the terms clonal cytopenia of undetermined significance (CCUS) (11, 19) and clonal monocytosis of undetermined significance (CMUS) (20) used to describe cytopenia(s) or monocytosis, respectively, in the presence of a CH mutation, in individuals without overt neoplasia, such as a myelodysplastic syndrome (MDS) or chronic myelomonocytic leukemia (CMML).

Authorship note: WGD and MAM contributed equally to this work.

Conflict of interest: GSV is a consultant and scientific advisory board member for STRM.BIO and holds a research grant from AstraZeneca.

Copyright: © 2024, Dunn et al. This is an open access article published under the terms of the Creative Commons Attribution 4.0 International License.

Reference information: *J Clin Invest.* 2024;134(19):e180065.

<https://doi.org/10.1172/JCI180065>.

Table 1. Nomenclature of clonal hematopoiesis

Abbreviation	Term	VAF	Description
CH/ARCH	Clonal hematopoiesis/age-related clonal hematopoiesis	Any	The expansion of hematopoietic stem cells bearing a somatic, fitness-conferring gene mutation or mosaic chromosomal abnormality in an individual without evidence of hematological malignancy.
CHIP	Clonal hematopoiesis of indeterminate potential	≥2%	A subset of CH, in which the clone has reached a critical (arbitrary) size threshold of a variant allele frequency 2%.
L-CHIP	Lymphoid clonal hematopoiesis of indeterminate potential	≥2%	CH driven by somatic mutations or mosaic chromosomal abnormalities that are associated with lymphoid malignancies. Less prevalent than conventional CH and CHIP, in which gene mutations and mCAs are more strongly associated with myeloid cancers.
CCUS	Clonal cytopenia of undetermined significance	≥2%	An individual meeting definition of CHIP and cytopenia(s) (hemoglobin, <12/13 g/dL in females/males; neutrophils, <1.8 × 10 ⁹ /L; platelets, <150 × 10 ⁹ /L)
CMUS	Clonal monocytosis of undetermined significance	≥2%	An individual meeting definition of CHIP, with monocytes ≥0.5 × 10 ⁹ /L and ≥10% of the white cell count (with no evidence of a reactive cause).
CCMUS	Clonal cytopenia and monocytosis of undetermined significance	≥2%	An individual meeting definition of both CCUS and CMUS (i.e., CHIP, monocytosis, and cytopenias).

VAF, variant allele fraction.

While a key feature of CH is the presence of a somatic driver mutation, it is recognized that many individuals harbor clonal expansions in the absence of a known driver (“driverless” CH) (3, 21, 22). Studies utilizing somatic “passenger” mutations as barcodes to identify these driverless expansions have shown that driverless CH can be as prevalent as its driver-bearing counterpart but confers a lower risk of malignant progression (3, 23). The nomenclature of CH is summarized in Table 1.

This Review will focus on driver mutation-bearing CH (henceforth referred to as “CH”), examining how the molecular diversity of this entity and the context in which it arises influence the risk of progression to myeloid neoplasia.

The spectrum of driver mutations in CH

The majority of CH is caused by mutations in a limited set of genes (Figure 1), predominately those involved in epigenetic regulation (*DNMT3A*, *TET2*, and *ASXL1*), signal transduction (*JAK2*), splicing (*SF3B1*, *SRSF2*, and *U2AF1*), and the DNA damage response (DDR; *TP53* and *PPM1D*). This section will explore current knowledge regarding the molecular consequences of these mutations and their role in driving CH.

Epigenetic regulation. The production of mature blood cells requires tightly regulated changes in gene expression that are coordinated by epigenetic marks such as DNA methylation and histone modifications. Mutations in the epigenetic factors *DNMT3A*, *TET2*, and *ASXL1* are found in approximately 70% of CH cases (3, 4, 10) and drive aberrant programs of HSC self-renewal and differentiation.

DNMT3A is the most frequently mutated gene in CH, and it encodes a DNA methyltransferase of the *DNMT* family (*DNMT1*, *DNMT3A*, and *DNMT3B*). *DNMT3A* is the major de novo methyltransferase (24–26) and can form dimers or tetramers with substantially greater activity than the monomer (27–29). *DNMT3A* deposits methyl groups at the C5 position of cytosines, most commonly those adjacent to a guanine residue (a “CpG” motif), and is usually associated with reduced gene expression (particularly when spanning gene promoters), silencing of transposable ele-

ments, and heterochromatin formation. Noncanonical roles for *DNMT3A* have also been reported, including modulation of RNA splicing by promoting spliceosome recruitment (30).

CH-associated mutations in *DNMT3A* are missense or truncating loss-of-function mutations, which lead to global hypomethylation and increased HSC self-renewal (31–35). Approximately 20% of *DNMT3A*-CH involves substitution mutations affecting a single codon, R882 (31). Although *DNMT3A*-R882 mutations are almost always heterozygous, the mutant protein was proposed to exert a dominant-negative effect by dimerizing with its wild-type counterpart and preventing formation of the more active tetrameric complex (25, 36), although this has been disputed (37). The reduced methyltransferase activity in *DNMT3A*-mutated cells causes widespread hypomethylation, including formation of large hypomethylated regions known as methylation “canyons” (38). This prevents the silencing of multipotency genes during hematopoietic differentiation and drives persistent activation of HSC self-renewal programs (25, 33, 35, 39–41). In line with their enhanced self-renewal, *Dnmt3a*-mutated cells outcompete wild-type cells in competitive hematopoietic progenitor transplant experiments (33, 35, 41). Moreover, they retain their ability to produce multilineage output and are immunophenotypically indistinguishable from unmutated cells, although their contribution to mature progeny is lost following successive transplants (41). *DNMT3A* mutations have also been associated with hypermethylation at specific loci (24); however, the relevance of this to clonal expansion is poorly understood.

TET2 is the second most frequently mutated gene in CH. Belonging to the TET family of α -ketoglutarate (α -KG, also known as 2-oxoglutarate or 2-OG) dependent dioxygenases, *TET2* promotes DNA demethylation by converting 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC) (42, 43). 5hmC is subject to active demethylation via a series of intermediates (42, 44) and passive demethylation due to reduced binding of the maintenance methyltransferase DNMT1 (45, 46). *TET2* can also repress inflammatory signaling via HDAC2-associated histone deacetylation (47).

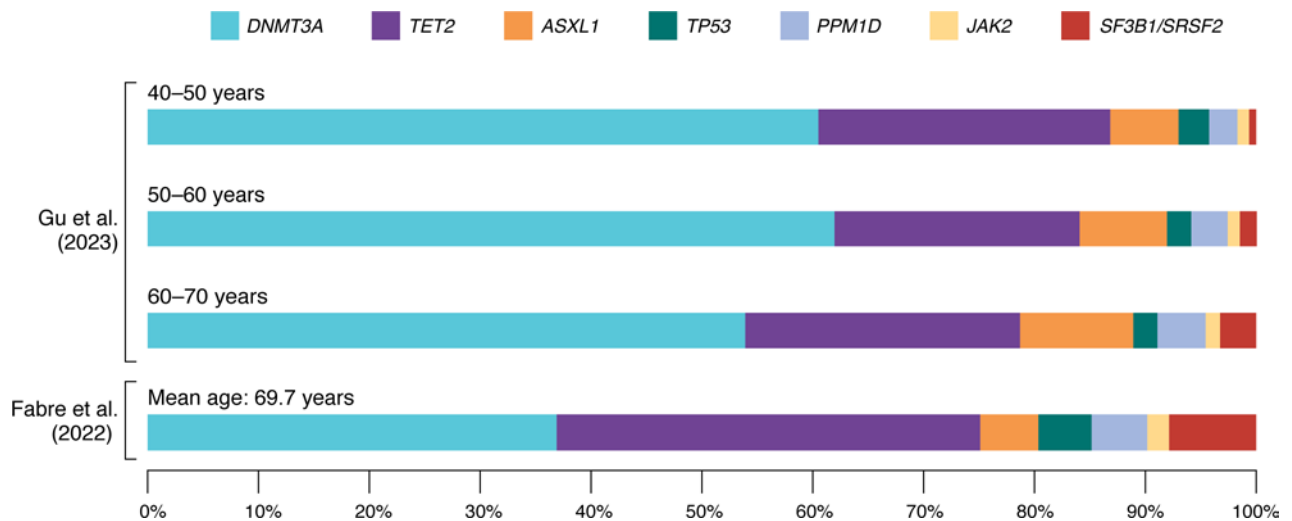


Figure 1. Proportions of commonly mutated CH driver genes in different age ranges. Proportions are based on published data from Gu et al., 2023 (144), broken down by age range, and Fabre et al., 2022 (80). The proportion of individuals with mutations in splicing factor genes (*SF3B1/SRSF2*) is very low before the age of 50 but increases markedly with advancing age. Similarly, the relative proportion of *DNMT3A*-CH decreases with advancing age, as the prevalence of *TET2*-CH increases. Of note, Fabre et al. use greater sequencing depth, enabling them to identify smaller clones than those in the Gu et al. study.

The loss-of-function mutations in *TET2* seen in CH have been studied extensively using mouse models. Similar to *Dnmt3a* mutations, *Tet2*^{-/-} cells exhibit a clonal advantage in competitive transplantation, increased replating capacity in vitro, and a granulocyte-monocyte lineage bias resembling CMML (43, 48–52). These features have been attributed to DNA hypermethylation and subsequent changes in transcription factor binding at loci involved in HSC self-renewal and lineage-specific differentiation (53–55). *Tet2*^{-/-} mice also possess elevated IL-6 levels and reduced apoptosis in response to inflammatory cytokines (47, 56, 57), suggesting an additional feedback loop through which *TET2* mutations may drive CH.

Reduced *TET2* function is also implicated in *IDH1*- and *IDH2*-mutant CH. Located in the cytoplasm (*IDH1*) or mitochondrial matrix (*IDH2*), *IDH1/2* are metabolic enzymes that convert isocitrate to α -KG. Hotspot mutations in either gene (*IDH1*-R132, *IDH2*-R140, and *IDH2*-R172) alter enzymatic function and promote the reduction of α -KG to the “oncometabolite” 2-hydroxyglutarate (2-HG) (58–60). The structural similarity between α -KG and 2-HG allows the latter to occupy substrate-binding sites on *TET2* and other α -KG-dependent dioxygenases, including prolyl hydroxylases and histone demethylases, inhibiting their enzymatic function (61, 62). Mutations in *IDH1/2* and *TET2* are mutually exclusive in hematological malignancy (60), supporting that inhibition of *TET2* activity plays a dominant role in *IDH1/2*-mutant myeloid malignancy. However, in contrast to *TET2*, *IDH1/2* mutations are rare in CH, which may reflect the fact that they are limited to hot spots and also in part that they may progress to myeloid neoplasms (MNs) more rapidly.

ASXL1 encodes a chromatin-binding protein that interacts with several histone-modifying complexes (63–66), including PRC2, which deposits H3K27me, and BAP1, a deubiquitinase that removes the repressive H2AK119Ub mark. Mutations in *ASXL1* drive approximately 10% of CH (3, 4, 10) and take the form of nonsense or frameshift mutations in the final or penultimate exon, resulting in a truncated protein that lacks the C-terminal his-

tone-binding domain. Whereas knockout of *Asxl1* in mice drives an MDS-like phenotype characterized by increased H3K27me and derepression of *Hoxa* genes (66, 67), *ASXL1* truncating mutations are thought to be gain of function (68) and stabilize interactions between *ASXL1* and BAP1 (69–71). This promotes BAP1 deubiquitinase activity, causing a reduction in H2AK119Ub levels and upregulation of a myeloid gene expression program (72). Mouse models of *Asxl1* mutations exhibit several overlapping phenotypes, including myeloid skewing, anemia, and thrombocytosis, and varying rates of progression to overt leukemia. However, the competitive advantage for *Asxl1*-mutated HSCs is more limited compared with *Dnmt3a*- or *Tet2*-mutant cells (73–79).

Mutations in *DNMT3A*, *TET2*, and *ASXL1* have widespread and divergent effects on gene expression yet are all capable of driving clonal expansion. This is particularly surprising with respect to *DNMT3A* and *TET2*, whereby mutations with opposing effects on DNA methylation promote similar increases in HSC fitness and/or “stemness.” Additionally, there are key differences between these CH subtypes, with *DNMT3A* dominating CH in early life and *TET2*-CH becomes increasingly common later in life, overtaking *DNMT3A*-CH by the eighth decade (80, 81). It is also notable that, while smoking increases the risk of most forms of CH, it is most strongly associated with *ASXL1*-CH (10, 82). Finally, lineage contribution studies have shown that *DNMT3A*-CH can more commonly contribute to nonmyeloid lineages compared with other CH subtypes (83).

Signal transduction. HSC behavior also depends on the integration of external stimuli, mediated by the binding of cytokines to extracellular receptors and subsequent activation of intracellular signaling. Mutations in components of the erythropoietin (EPO) and thrombopoietin (TPO) signaling pathways, specifically activating mutations in *JAK2*, *CALR*, or *MPL*, can drive aberrant HSC proliferation and clonal expansion and can eventually lead to the development of myeloproliferative neoplasms (MPNs) or MDS/MPN overlap syndromes.

JAK2 encodes a cytoplasmic kinase that associates with receptors for several cytokines, including EPO, TPO, and G-CSF. Following cytokine binding, *JAK2* activates pathways, including JAK/STAT, MAPK, and PI3K/AKT. *JAK2-V617F* hotspot mutations disrupt the pseudokinase domain, resulting in constitutively active *JAK2* and downstream activation of JAK/STAT signaling in the absence of cytokine engagement (84–88). This promotes myeloid cell expansion that can lead to the development of MPNs, such as polycythemia vera (PV) or essential thrombocythemia (ET), characterized by the overproduction of red blood cells and/or platelets or primary myelofibrosis (PMF) linked to the proliferation of abnormal profibrotic megakaryocytes (89).

Substitution mutations in *MPL*, encoding the TPO receptor, promote conformational changes resembling those induced by TPO binding and subsequently activate *JAK2* (90, 91). Aberrant *MPL* activation is also seen with frameshift mutations in the final exon of *CALR*, an ER-associated chaperone protein whose mutant form binds *MPL* and activates JAK/STAT signaling in a *MPL*-dependent manner (92–99). Whereas *JAK2* mutations mimic signaling from several cytokine receptors and lead to diverse phenotypes, *MPL* and *CALR* mutations are thought to act solely through TPO and are almost exclusively associated with ET and PMF (90, 100, 101). Unlike *JAK2*, *CALR/MPL* mutations are rarely detected in CH, possibly because they promote thrombocytosis, even at low clonal fractions, increasing the likelihood of meeting the diagnostic threshold for ET.

There is a well-established trade-off between proliferation and long-term self-renewal in HSCs, raising questions about how myeloproliferative mutations in *JAK2*, *CALR*, and *MPL* can simultaneously drive HSC expansion over decades. This antagonism is captured in mouse HSCs expressing human *JAK2-V617F*, which exhibit enhanced colony formation at the expense of impaired long-term reconstitution capacity (102, 103). The absence of a long-term competitive advantage over wild-type cells is also seen in *Calr*-mutant mice, despite their development of ET-like disease (97–99). A possible explanation for this is that mutations in *JAK2*, *CALR*, and *MPL* may promote clonal expansion in specific genetic backgrounds or environmental contexts not captured in current models.

Splicing factors. The splicing factors *SF3B1*, *SRSF2*, and *U2AF1* are another class of genes frequently mutated in CH, particularly in elderly individuals (6). Splicing describes the removal of intronic sequences from nascent pre-mRNAs to generate a mature mRNA transcript. This process is carried out by the spliceosome, a series of multisubunit ribonucleoprotein complexes. *SF3B1*, *SRSF2*, and *U2AF1* encode core factors within the spliceosome, with *SRSF2* involved in exon recognition by binding to exonic splicing enhancers, *SF3B1* in branch point recognition and *U2AF1* in the recognition of the 3' splice site (104, 105). Splicing factor gene mutations typically result in single amino acid substitutions at hot spots such as *SRSF2-P95*, *U2AF1-S34*, and *U2AF1-Q157* and several positions within *SF3B1*, including *SF3B1-K700*, *SF3B1-K666*, and *SF3B1-R625* (105, 106). These mutations alter protein function, leading to aberrant splicing of large numbers of transcripts and the generation of novel mRNA isoforms. Additional mutations have also been detected in the splicing factor *ZRSR2*, albeit more rarely and as nonsense, frameshift, or substitution

mutations that disrupt the function of the minor spliceosome and cause aberrant retention of U12-type introns (106, 107).

Initial investigations of splicing factor mutations deployed knockin mouse models, which develop an MDS-like disease (108–112) but lack some of the characteristic features seen in patients, such as the presence of ring sideroblasts in *SF3B1*-mutated MDS (113, 114). This is in large part due to the limited overlap in misspliced transcripts between mice and humans (111), such as the *ABCB7* and *TMEM14C* mRNAs that play important roles in the generation of ring sideroblasts in humans (115, 116). Nevertheless, the presence of an MDS-like phenotype in mice suggests that dysregulated splicing per se may play a role in pathogenesis, regardless of the specific transcripts affected. Recent studies have also shown that splicing factor mutations lead to the generation of R-loops, DNA:RNA hybrids that can activate DDR signaling (117–119). Notably, current models of splicing factor mutation exhibit a reduction in cellular fitness, as assessed by competitive transplantation (108–112), so the mechanism through which splicing factor gene mutations drive clonal expansion remains elusive.

The absence of clonal expansion in models of *SF3B1*, *SRSF2*, and *U2AF1* mutations may relate to the unique relationship these mutations exhibit with age. Unlike other CH drivers, splicing factor gene mutations are rarely detected prior to middle age and show a dramatic increase in prevalence in elderly individuals (6). Going forward, understanding interactions between the pleiotropic effects of splicing factor mutations and an aging hematopoietic system will be critical in characterizing their role in CH.

DNA damage response. HSCs are exposed to several forms of DNA damage that can lead to oncogenic mutations and cancer development (120). To safeguard against this, the DDR enables cells to sense genomic insults and activate protective mechanisms. These include cell cycle arrest, giving time for DNA repair, and the activation of apoptosis or senescence if DDR signaling persists. Given the antiproliferative and tumor-suppressive role of the DDR, mutations in this pathway (*TP53*, *PPM1D*) are recognized drivers of CH.

TP53 encodes a transcription factor, p53, that is typically maintained at low levels in cells via proteasomal degradation (121, 122). Cellular stress such as DNA damage causes phosphorylation and stabilization of p53 (123), which then activates transcriptional programs associated with cell cycle arrest, apoptosis, and senescence (124). In CH, *TP53* mutations are predominantly truncating or missense mutations affecting the DNA-binding domain of the protein (125). These p53 variants are less able to bind DNA or activate transcription, and missense mutations can exhibit a dominant-negative effect by oligomerizing with wild-type p53 and partially reducing its activity (125). As a result, *TP53*-mutant cells exhibit a reduction in apoptosis when exposed to radiation or cytotoxic chemotherapy (120, 121). The antiapoptotic effects of *TP53* mutations explain the increased prevalence of *TP53*-mutant CH after cancer chemo-/radiotherapy and its association with therapy-related MNs (tMNs) (126–129).

Similar prevalence patterns are seen with mutations in *PPM1D* (130), another DNA damage response gene frequently mutated in CH. *PPM1D* expression is induced by p53, and the protein acts in a negative feedback loop, dephosphorylating p53 and other DDR factors and suppressing DDR-mediated apoptosis (131, 132). In

CH, frameshift or nonsense mutations distributed within the final exon of *PPM1D* create a C-terminally truncated protein that lacks a ubiquitination domain (9). As a result, truncated PPM1D is several-fold more stable than wild-type PPM1D but maintains its phosphatase activity. The increased levels of the stabilized mutant PPM1D promote dephosphorylation of several DDR factors, suppressing activation of the DDR and reducing apoptosis following cytotoxic chemotherapy or radiation (9, 133). Similar mechanisms likely explain why CH with mutations in other DDR factors (*CHEK2*, *ATM*, and *SRCAP*) are also observed in patients following chemotherapy (127–129, 134).

Notably, a substantial proportion of people with *TP53*- and *PPM1D*-mutant CH lack prior therapy exposure (128). It is possible that the antiapoptotic effects of these mutations promote survival following alternative sources of cellular stress or that these mutations have pleiotropic effects that increase HSC fitness via alternative mechanisms, as suggested for *TP53* (135, 136).

Associations between CH and hematological malignancy

In 2014, landmark papers from Genovese et al. and Jaiswal et al. established an association between CH and incident hematological malignancies (3, 4). Several studies have since confirmed this association and demonstrated that this is predominantly related to progression to myeloid cancers (8, 10, 137–139). MNs, which encompass acute myeloid leukemia (AML), MDS, and MPNs, represent a family of cancers with a combined annual incidence of approximately 12 cases per 100,000 persons (140). Molecular heterogeneity is a feature of all MN subtypes (141–143), and many of the mutated genes and genomic aberrations found across the spectrum of MN are drivers of CH. As such, CH is regarded as a premalignant lesion, and the relationship between CH and MN can be considered analogous to that between monoclonal gammopathy of undetermined significance (MGUS) and multiple myeloma (MM) or monoclonal B lymphocytosis (MBL) and chronic lymphocytic leukemia (CLL). Similar to MGUS and MBL, CH can precede the development of MN by many years, with an overall annual rate of progression of around 0.5%–1% (4). However, the precise somatic mutations driving each clone strongly influence the likelihood and nature of progression to MN (144).

While *DNMT3A* is the most frequently mutated gene in CH, the risk of transformation of *DNMT3A*-CH is low compared with other CH subtypes. However, this risk is markedly higher for R882 hotspot mutations that are highly fit (145) and 8-fold more likely to progress to AML than non-R882 *DNMT3A*-CH (144). It is also notable that *DNMT3A*-CH was not found to be significantly associated with progression to MDS or MPN (144). Of the other commonly mutated epigenetic modifiers, *TET2*-CH confers a modestly increased risk of incident AML or MPN, while *ASXL1*-CH shows a stronger association with all major subtypes of MN. An important exception is the striking association between *TET2*-CH and incident CMML (hazard ratio, 91.5; based on analyses by Kar, Quiros et al., ref. 10), although the absolute risk of developing CMML from *TET2*-CH is low.

In contrast to CH driven by mutations in epigenetic modifiers, CH driven by mutations in splicing factor genes *SF3B1*, *SRSF2*, and *U2AF1* engenders a much higher risk of progression to MN, partic-

ularly MDS (144). Analogously, driver mutations in signal transduction genes *JAK2*, *CALR*, and *MPL* confer a high risk of progression to MPNs (144). These predilections are unsurprising, since these mutations are defining features of the associated malignancies: for example, *SF3B1* mutations are characteristic of the MDS subtype refractory anemia with ring sideroblasts (RARS), whereas mutations occurring in *JAK2*, *CALR*, and *MPL* are MPN-defining mutations (84, 86, 87, 90, 91, 100, 101, 113, 146, 147).

TP53-mutant CH exhibits a variable association with MN in the literature, with some studies demonstrating a high risk of progression (8, 137) and others finding no significant association (144, 148). This heterogeneous association appears counterintuitive given that *TP53* mutations confer a dismal prognosis in MDS and AML, and it is likely to be a product of the multitude of confounding factors modulating the behavior and fitness of *TP53*-mutated clones, such as previous exposure to cytotoxic chemotherapy (128), presence of germline pathogenic variants (Li-Fraumeni syndrome) (149), biallelic *TP53* loss (150, 151), and mutation type (with missense variants behaving in a dominant-negative manner) (125). Similarly, gain-of-function mutations in *PPM1D* are present in approximately 20% of cases of therapy-related AML and MDS but are rare in primary (de novo) AML, consistent with prior chemotherapy exposure strongly augmenting the fitness advantage of *PPM1D*-mutant CH clones (9). Where *PPM1D* mutations are present in tMNs, they are often subclonal, suggesting *PPM1D* mutations in this context may indirectly reflect genotoxic damage, rather than being true drivers of the tMN (152).

Beyond the individual mutated gene(s), the number of driver mutations and CH clone size also correlate strongly with progression (8, 137, 144, 148). Furthermore, the propensity to transform is proportional to the relative growth rate associated with individual driver genes: for example, splicing factor mutations, which are among the fittest CH drivers, are associated with a high risk of progression to MN, whereas non-R882 *DNMT3A* mutations typically have a lower fitness effect and progress infrequently (80, 145). An additional factor is the mode of progression from CH to MN. Although CH is the shared precursor of most cases of MN, the trajectory linking asymptomatic clonal expansion (i.e., CH) to frank malignancy varies between different types of MN. De novo AML is usually explosive in its clinical presentation, with previously healthy patients often presenting with complications of fulminant bone marrow failure or leukostasis. Two of the most common drivers of de novo AML are small duplications/insertions in the final exon of *NPM1* and internal tandem duplications of *FLT3* (*FLT3*-ITD). Although these mutations are highly recurrent in AML, they are conspicuously absent in CH. This suggests that acquisition of these mutations leads to rapid clonal expansion and transformation to AML in a short time frame, a trajectory exemplified in a study by Quiros et al., who employed sensitive variant calling to identify individuals with *NPM1* mutations in the United Kingdom Biobank (UKB) (153). Their analysis of 200,453 individuals with whole-exome sequencing (WES) data identified only two with *NPM1* mutations, both of whom developed AML within five months. It is also notable that both these individuals had large antecedent *DNMT3A*-CH clones and that

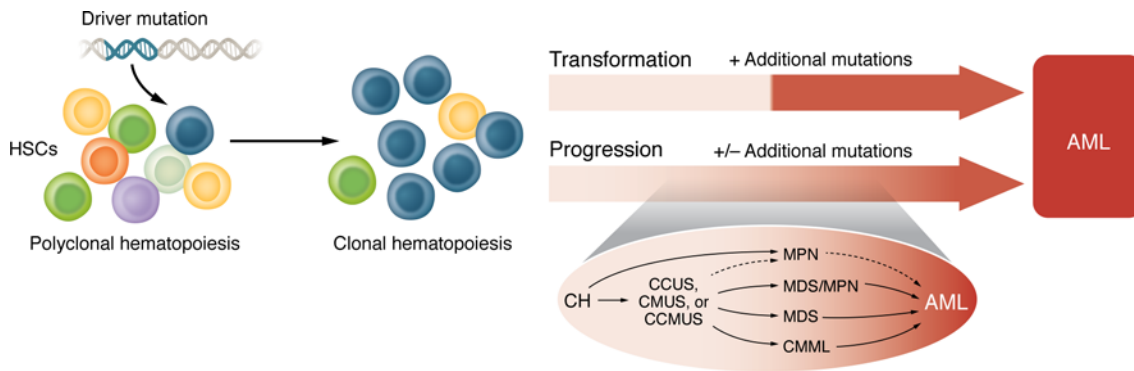


Figure 2. Malignant progression or transformation of CH. The transition between CH and myeloid neoplasia can occur via different routes, including direct “transformation” to de novo acute myeloid leukemia (AML) or “progression” through clonal myeloid disorders, such as clonal cytopenia of undetermined significance (CCUS), clonal monocytosis of undetermined significance (CMUS), clonal cytopenia and monocytosis of undetermined significance (CCMUS), and established myeloid neoplasms, such as the myelodysplastic syndrome (MDS), myeloproliferative neoplasms (MPN), and chronic myelomonocytic leukemia (CMML), or other MDS/MPN overlap syndromes, each of which has potential to progress to (secondary) AML. Dotted lines in the progression inset indicate a less frequent progression path: for example, CCUS predominantly progresses to MDS, but in a minority of cases progresses to MPN (specifically, primary myelofibrosis), while MPNs progress to AML only infrequently. Transformation from CH to AML requires additional leukemia-associated mutations, such as *NPM1*/*FLT3*-ITD mutations, whereas acquisition of additional driver mutations is not always observed at progression from CH to MDS or MPN.

it is common for *DNMT3A* (particularly R882), *NPM1*, and *FLT3* mutations to co-occur in AML (142, 154); indeed, *NPM1* mutations in isolation are known to be insufficient for transformation to AML (155). Taken together, this implies that CH mutations such as *DNMT3A*-R882 may either facilitate the driver effects of *NPM1* mutations or increase the likelihood of acquiring these and potentially also of *FLT3*-ITD mutations, ostensibly through a shared mutagenic process, since both are insertion mutations with similar molecular anatomies (156, 157).

In comparison to de novo AML, the distinction between CH and MDS or MPN is less dichotomous (Figure 2). In fact, a proportion of individuals with CH driven by mutations in splicing or signal transduction pathways who progress to MDS or MPN appear to do so without acquiring a further driver event: for example, approximately one-quarter of *SF3B1*-mutated MDS cases display no additional driver mutations (141), and 45% of MPNs bear only a single driver, namely mutant *JAK2*, *CALR*, or *MPL* (143). In this context, the transition between CH and MDS or MPN can be indistinct and may reflect either the clone reaching a critical mass at which it appreciably perturbs blood counts, leading to clinical presentation, or the accumulation of heritable cellular changes that progressively perturb hematopoiesis.

Because of its more gradual nature, the transition from CH to MDS can frequently be identified at the intermediate state of CCUS, where the morphological and cytopenic changes do not meet defined criteria for MDS diagnosis. Similarly, the intermediate phase between CH and CMML has recently been enshrined by the International Consensus Classification as CMUS (see Table 1) (20). By contrast, the diagnostic criteria of PV and ET do not include an equivalent “clonal cytosis” entity. Instead, a small rise in hemoglobin concentration or platelet count that leads to crossing predefined thresholds results in an arbitrary “jump” from CH to MPN (20, 146). This highlights how prevailing pathological classifications do not adequately capture the biological continuum and progression from CH to MDS or MPN.

Risk stratification of CH and prediction of MN

Despite recent therapeutic advances, survival remains poor for patients with AML and MDS. The observation that myeloid malignancies are preceded by a long, detectable preclinical phase of clonal expansion raises the prospect of intervening within this window to prevent or delay malignant transformation. The current lack of premalignant intervention(s) notwithstanding, a key obstacle to realizing the goal of MN prevention is the high prevalence but low overall progression rate of CH: consequently, when CH is detected, accurate prediction of MN risk is of paramount importance in understanding which patients could benefit from clinical follow-up and/or entry into interventional studies.

Two groups independently sought to address this challenge utilizing the wealth of data available in the UKB. Weeks et al. (148) identified 11,337 cases of CHIP and CCUS among 193,743 UKB participants, labeled each according to whether they developed incident MN within ten years of follow-up, and then used demographic, molecular, and blood count data as features in a decision tree–based machine learning model. Having generated binarized features from their final model, they used these as input to a Cox regression model to generate coefficients from which they derived their CH risk score (CHRS), which designated individuals as being at low, intermediate, or high risk of MN. Through validation in an unseen internal and two external cohorts, they showed that this simple, interpretable score was strongly predictive of incident MN, with individuals with low CHRS having an event rate of 0.6% versus 15.2% for those with high CHRS in one of the external CH cohorts.

In comparison, Gu et al. (144) identified participants with CH and performed forward stepwise Cox regression to fit a parsimonious model (termed MN-predict) on a training cohort of 207,035 individuals to predict MN-free survival using demographic, molecular, blood count, and biochemical data and similarly validated their model in an unseen internal test set and two external CCUS cohorts. In contrast with the approach of Weeks et al., Gu et al. modeled each MN subtype separately;

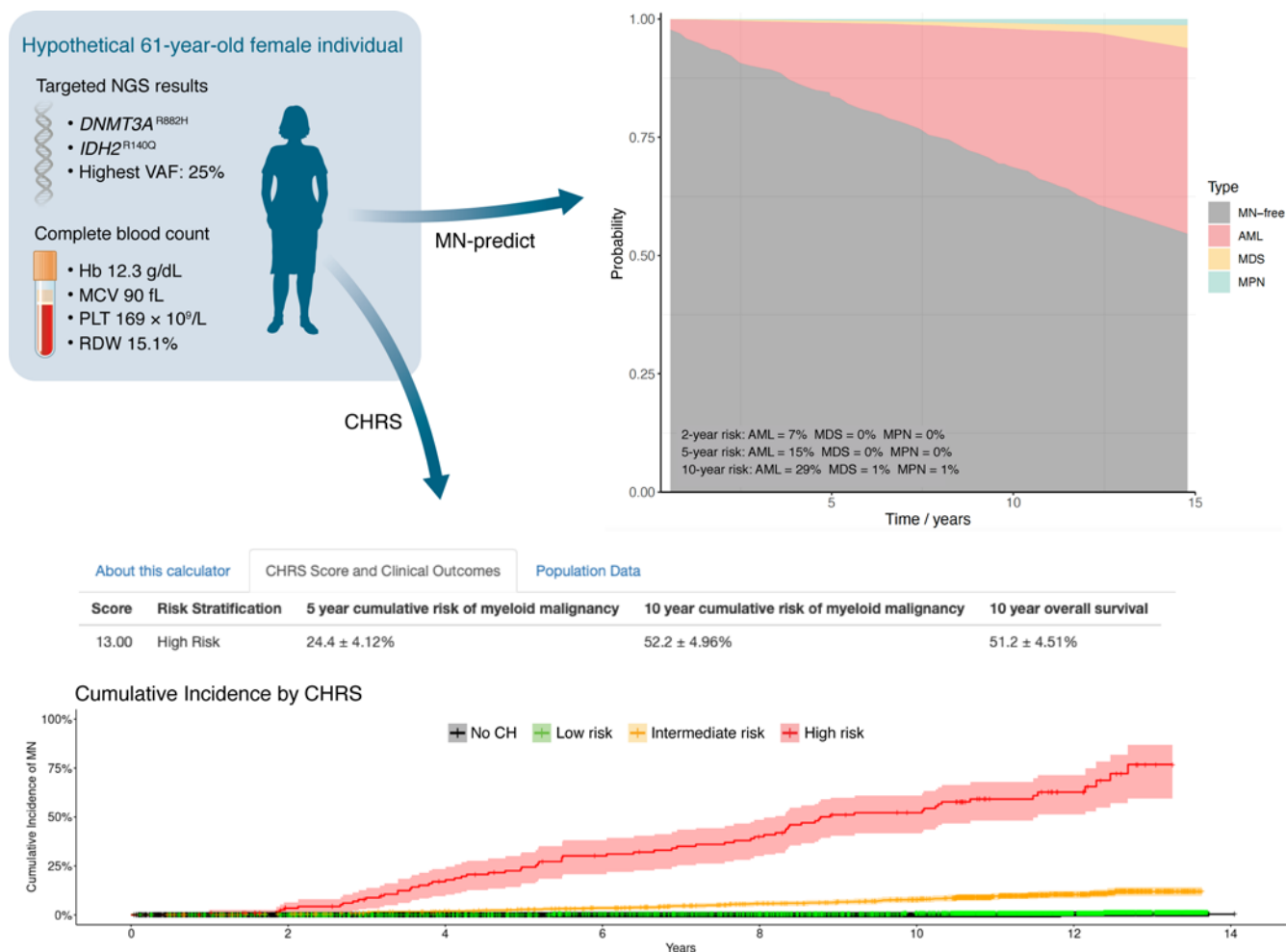


Figure 3. Risk stratification of CH. A hypothetical individual with *DNMT3A* and *IDH2*-mutated CH is used as an exemplar. NGS, demographic, and complete blood count data are used as inputs into two CH risk stratification tools (144,148), and both identify the individual as being at high risk of incident MN. The clonal hematopoiesis risk score (CHRS) (148) (available online at <http://www.chrsapp.com/>) designates the individual as “high risk” and depicts the cumulative incidence of myeloid neoplasia for this subset. The output of MN-predict (144) (available online at https://bioinf.stemcells.cam.ac.uk/shiny/vassiliou/MN_predict) shows the probability of progression to each MN subtype, highlighting that the risk is predominantly related to transformation to AML in this sample case. Images from CHRS and MN-predict tools were reproduced with permission. Hb, hemoglobin; MCV, mean corpuscular volume; PLT, platelet count; RDW, red cell distribution width.

made no distinction among CH, CHIP, or CCUS; and generated quantitative, time-dependent predictions of MN-free survival for each MN subtype (AML, MDS, MPN).

Each approach has its own merits: the CHRS developed by Weeks et al. offers a readily interpretable score and categorization of the risk of any MN that will be easily understood by clinicians (148), whereas the MN-predict approach adopted by Gu et al. offers nuanced, continuous estimates of the likelihood of developing each type of the three main types of MN (144) (Figure 3). CHRS takes as input a smaller number of variables commonly used in the clinic, whereas MN-predict can also incorporate other variables that may not be routinely measured in clinical practice, although it can still generate robust predictions in their absence. Furthermore, only MN-predict offers predictions for CH driven by mutations in *U2AF1*, since issues with this locus in the human reference genome (158) precluded its inclusion in the CHRS approach. Both approaches are readily accessible through user-friendly web interfaces (<http://www.chrsapp.com>

and https://bioinf.stemcells.cam.ac.uk/shiny/vassiliou/MN_predict/) and will be important tools for risk-stratifying individuals identified as having CH, CHIP, or CCUS in the future (though currently neither tool has been prospectively validated nor reviewed by the United States Food and Drug Administration or an equivalent national body). In the absence of a unified risk-stratification strategy for CH, the use of both models, or the development of a model incorporating the best aspects of both may be beneficial in assessing MN risk in individuals harboring CH.

Clinical management of CH and preventative medicine

With increasingly widespread use of NGS for diagnostics and research, CH is a frequent incidental finding, leading to the establishment of clinics dedicated to managing CH. As CH is a highly prevalent phenomenon in advanced age, adoption of the aforementioned risk-stratification tools in clinical decision-making will be key to ensuring that individuals with high-risk CH are

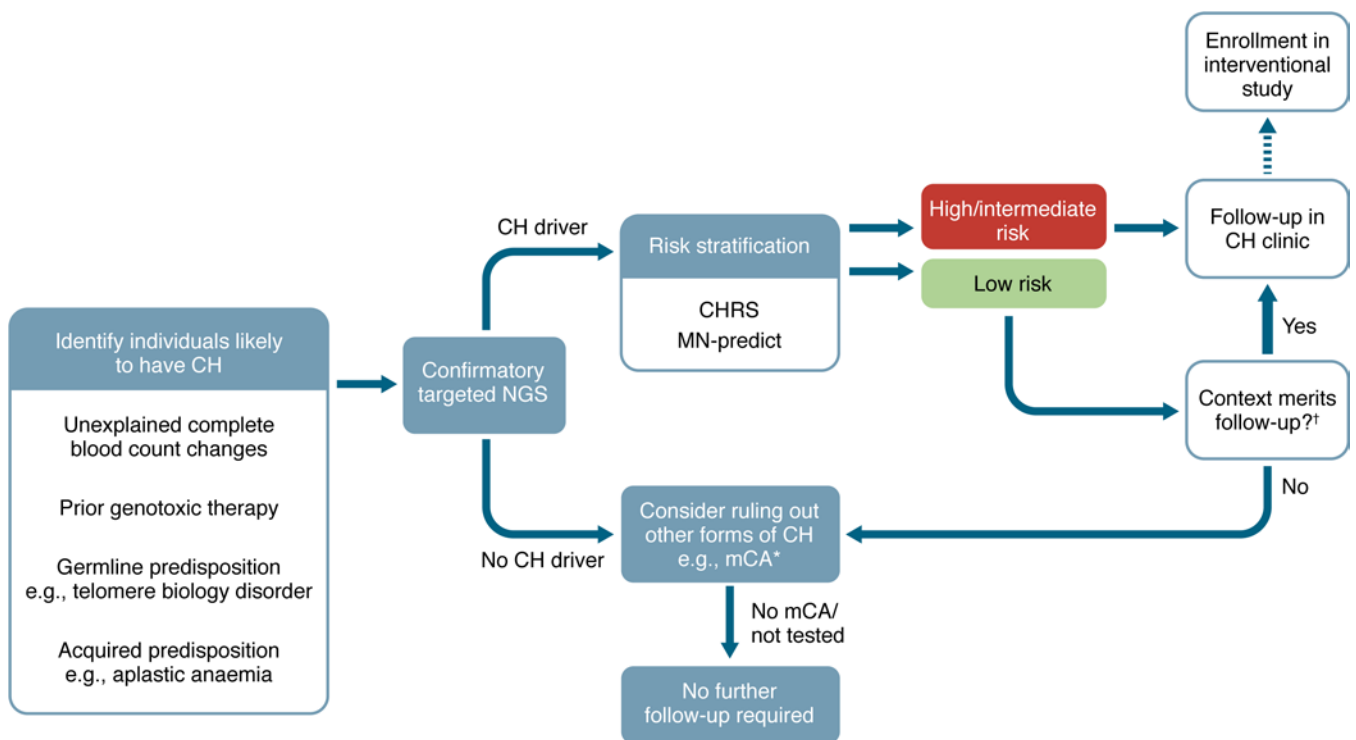


Figure 4. Identification and management of CH. In the absence of population-level screening for CH, testing can be targeted to individuals at higher a priori risk of CH. These include individuals with germline predisposition, such as those with telomere biology disorders, and conditions associated with acquired predisposition to CH, such as acquired aplastic anemia (in which clonal selection occurs as a means of escaping immune-mediated destruction). Furthermore, individuals with certain unexplained complete blood count changes may harbor particular forms of CH, while exposure to cytotoxic chemotherapy or radiotherapy favors the expansion of *PPM1D*- and *TP53*-mutated clones. As most individuals with CH will never develop MN, clinical follow-up and consideration for enrollment into interventional studies should be reserved for individuals at higher risk of malignant progression. *In individuals with no evidence of CH on targeted NGS, testing for mosaic chromosomal abnormalities (mCA) may be considered where there remains a strong suspicion of this form of CH, e.g., in an individual with unexplained cytopenias or monocytosis. †In some cases, individuals designated as “low-risk CH” using risk-stratification tools, may merit follow-up owing to the particular clinical context: a typical example would be an individual with a small (low variant allele fraction [VAF]) *TP53*- or *PPM1D*-mutant clone who is scheduled to undergo chemo- or radiotherapy.

prioritized for closer follow-up and consideration for recruitment to interventional studies. Importantly, knowledge of the clinical context, which cannot be an input feature in predictive models, will be critical in guiding nuanced management of individuals whose age, fitness, perception of risk, comorbidities, treatment history, and other factors will need to be considered prior to deciding their clinical management.

At present, there are no approved therapies to prevent progression of CH to MN, and the fact that individuals with CH are largely asymptomatic dictates that any such intervention must be well-tolerated with minimal toxicity. Furthermore, given the protracted window between CH detection and onset of MN, use of MN-free survival may be an impractical endpoint for initial interventional studies, necessitating the use of surrogate endpoints such as reduction in clone size or improvement in blood indices in the case of CCUS. Recently initiated clinical trials are investigating potential treatments for CH/CCUS, including anti-IL-1 β antibody canakinumab for high-risk CCUS (to attenuate inflammation-driven clonal expansion) (159) (NCT05641831), the administration of vitamin C (a cofactor of TET2) for TET2-CCUS (NCT03418038), and the exploration of targeted therapies such as *IDH1/2* inhibitors in *IDH1/2*-CCUS (NCT05030441 and NCT05102370). In the absence of established therapies, an

important role of CH clinics is to identify patients for entry into clinical studies of MN prevention. Furthermore, in practice, CH clinics will establish the mode and frequency of monitoring high-risk CH/CCUS to identify patients who have transitioned to an established MN at the earliest possible stage, until evidence for treating specific subtypes of CH/CCUS is firmly established. A proposed structure for the clinical management of CH focused on MN prevention is outlined in Figure 4.

Unanswered questions and challenges

Despite the huge advances over the last decade, there are many unknowns in the field of driver gene CH and myeloid cancer prevention. A key unsolved problem is how to detect CH at scale. At present CH is identified by NGS, which is both expensive and impractical to apply at scale to unselected populations to identify the few individuals at high risk of MN. In an attempt to address this, efforts have been made to reduce the cost of CH-targeted NGS (160, 161), but applying this at scale in asymptomatic individuals would still be practically challenging. An alternative approach is to target NGS testing to individuals with a higher a priori risk of harboring high-risk CH or CCUS on the basis of simpler, more scalable tests. Since CH is known to perturb blood indices, such as red cell distribution width and mean cell volume, one such approach might

be to leverage complete blood count data to identify individuals likely to harbor CH who can be prioritized for genetic testing.

Prediction of progression from CH to de novo AML also remains challenging, despite the advances culminating in the CHRS and MN-predict risk stratification tools. For example, the predictive performance of MN-predict was lower for AML than MDS and MPN, and this likely reflects the varying trajectories of CH progression to different MN subtypes, with transformation to de novo AML being more rapid. Since outcomes for AML remain poor, refining risk stratification tools to specifically improve de novo AML prediction could be a priority in coming years.

Finally, while the landscape of recurrent driver mutations has been established by applying NGS to large cohorts, the precise mechanism by which many of these mutations engender a fitness advantage to mutant HSCs is unclear, and this will be of paramount importance in designing targeted treatments to avert progression of CH to MN. Most strikingly, the fitness effect of splicing factor mutations, which drive rapid clonal expansion and confer high risk of progression while perturbing a critical cellular process, is yet to be established: paradoxically, these clones do not expand well in vitro or in vivo and seem to expand preferentially in advanced age, as hematopoiesis becomes increasingly oligoclonal (6, 21, 110, 111). Ascertaining the mechanisms through which disruption of critical molecular processes (RNA splicing and DNA methylation) allows cells to expand clonally is likely to be a critical step toward developing rational strategies for myeloid cancer prevention.

Concluding remarks

The past decade has borne witness to unprecedented advances in our understanding of the natural history and evolution of myeloid cancers, aided substantially by the description and

comprehensive molecular characterization of CH. As the use of NGS becomes increasingly commonplace in clinical practice, the number of individuals identified with CH will continue to grow. Only a minority of these individuals will go on to develop MN, and advances in risk stratification are enabling clinicians to quantify the magnitude and nature of this risk. Further improvements in our ability to identify those at highest risk of progression will be important, but the most important hurdle in the successful development of the field of myeloid cancer prevention is the lack of an approved therapy to avert or delay MN development. As the second decade of research into CH commences, the focus is turning toward an improved understanding of the basis of clonal expansion and malignant progression, with the ultimate aim of developing effective interventions that can change the natural history of high-risk CH.

Acknowledgments

WGD is funded by a Clinical Research Fellowship from the Cancer Research UK Cambridge Centre (CTRQR-2021\100012). MAM is funded by a Wellcome Trust PhD Studentship (102160/B/13/Z). GSV is supported by a Cancer Research UK Senior Cancer Fellowship (C22324/A23015), and work in his laboratory is also funded by the Leukemia Lymphoma Society, Blood Cancer UK, European Research Council, Cancer Research UK, Kay Kendall Leukemia Fund, AstraZeneca, and Wellcome Trust.

Address correspondence to: George S. Vassiliou, Cambridge Stem Cell Institute, Jeffrey Cheah Biomedical Centre, Cambridge Biomedical Campus, Puddicombe Way, Cambridge, CB2 0AW. Email: gsv20@cam.ac.uk.

- Lee-Six H, et al. Population dynamics of normal human blood inferred from somatic mutations. *Nature*. 2018;561(7724):473–478.
- Busque L, et al. Nonrandom X-inactivation patterns in normal females: lyonization ratios vary with age. *Blood*. 1996;88(1):59–65.
- Genovese G, et al. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N Engl J Med*. 2014;371(26):2477–2487.
- Jaiswal S, et al. Age-related clonal hematopoiesis associated with adverse outcomes. *N Engl J Med*. 2014;371(26):2488–2498.
- Xie M, et al. Age-related mutations associated with clonal hematopoietic expansion and malignancies. *Nat Med*. 2014;20(12):1472–1478.
- McKerrell T, et al. Leukemia-associated somatic mutations drive distinct patterns of age-related clonal hemopoiesis. *Cell Rep*. 2015;10(8):1239–1245.
- Young AL, et al. Clonal haematopoiesis harbouring AML-associated mutations is ubiquitous in healthy adults. *Nat Commun*. 2016;7:12484.
- Abelson S, et al. Prediction of acute myeloid leukaemia risk in healthy individuals. *Nature*. 2018;559(7714):400–404.
- Hsu JI, et al. PPM1D mutations drive clonal hematopoiesis in response to cytotoxic chemotherapy. *Cell Stem Cell*. 2018;23(5):700–713.
- Kar SP, et al. Genome-wide analyses of 200,453 individuals yield new insights into the causes and consequences of clonal hematopoiesis. *Nat Genet*. 2022;54(8):1155–1166.
- Steensma DP, et al. Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes. *Blood*. 2015;126(1):9–16.
- Niroula A, et al. Distinction of lymphoid and myeloid clonal hematopoiesis. *Nat Med*. 2021;27(11):1921–1927.
- Loh PR, et al. Insights into clonal haematopoiesis from 8,342 mosaic chromosomal alterations. *Nature*. 2018;559(7714):350–355.
- Loh PR, et al. Monogenic and polygenic inheritance become instruments for clonal selection. *Nature*. 2020;584(7819):136–141.
- Jacobs KB, et al. Detectable clonal mosaicism and its relationship to aging and cancer. *Nat Genet*. 2012;44(6):651–658.
- Laurie CC, et al. Detectable clonal mosaicism from birth to old age and its relationship to cancer. *Nat Genet*. 2012;44(6):642–650.
- Schick UM, et al. Confirmation of the reported association of clonal chromosomal mosaicism with an increased risk of incident hematologic cancer. *PLoS One*. 2013;8(3):e59823.
- Saiki R, et al. Combined landscape of single-nucleotide variants and copy number alterations in clonal hematopoiesis. *Nat Med*. 2021;27(7):1239–1249.
- Kwok B, et al. MDS-associated somatic mutations and clonal hematopoiesis are common in idiopathic cytopenias of undetermined significance. *Blood*. 2015;126(21):2355–2361.
- Arber DA, et al. International consensus classification of myeloid neoplasms and acute leukemias: integrating morphologic, clinical, and genomic data. *Blood*. 2022;140(11):1200–1228.
- Mitchell E, et al. Clonal dynamics of haematopoiesis across the human lifespan. *Nature*. 2022;606(7913):343–350.
- Zink F, et al. Clonal hematopoiesis, with and without candidate driver mutations, is common in the elderly. *Blood*. 2017;130(6):742–752.
- Stacey SN, et al. Genetics and epidemiology of mutational barcode-defined clonal hematopoiesis. *Nat Genet*. 2023;55(12):2149–2159.
- Challen GA, et al. Dnmt3a and Dnmt3b have overlapping and distinct functions in hematopoietic stem cells. *Cell Stem Cell*. 2014;15(3):350–364.
- Kim SJ, et al. A DNMT3A mutation common in AML exhibits dominant-negative effects in murine ES cells. *Blood*. 2013;122(25):4086–4089.
- Okano M, et al. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell*. 1999;99(3):247–257.
- Jia D, et al. Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA

- methylation. *Nature*. 2007;449(7159):248–251.
28. Emperle M, et al. Cooperative DNA binding and protein/DNA fiber formation increases the activity of the Dnmt3a DNA methyltransferase. *J Biol Chem*. 2014;289(43):29602–29613.
 29. Holz-Schietinger C, et al. Mutations in DNA methyltransferase (DNMT3A) observed in acute myeloid leukemia patients disrupt processive methylation. *J Biol Chem*. 2012;287(37):30941–30951.
 30. Ramabadrán R, et al. DNMT3A-coordinated splicing governs the stem state switch towards differentiation in embryonic and haematopoietic stem cells. *Nat Cell Biol*. 2023;25(4):528–539.
 31. Brunetti L, et al. DNMT3A in leukemia. *Cold Spring Harb Perspect Med*. 2017;7(2):a030320.
 32. Yang L, et al. DNMT3A in haematological malignancies. *Nat Rev Cancer*. 2015;15(3):152–165.
 33. Challen GA, et al. Dnmt3a is essential for hematopoietic stem cell differentiation. *Nat Genet*. 2012;44(1):23–31.
 34. Scheller M, et al. Hotspot DNMT3A mutations in clonal hematopoiesis and acute myeloid leukemia sensitize cells to azacytidine via viral mimicry response. *Nat Cancer*. 2021;2(5):527–544.
 35. Loberg MA, et al. Sequentially inducible mouse models reveal that Npm1 mutation causes malignant transformation of Dnmt3a-mutant clonal hematopoiesis. *Leukemia*. 2019;33(7):1635–1649.
 36. Russler-Germain DA, et al. The R882H DNMT3A mutation associated with AML dominantly inhibits wild-type DNMT3A by blocking its ability to form active tetramers. *Cancer Cell*. 2014;25(4):442–454.
 37. Emperle M, et al. The DNMT3A R882H mutation does not cause dominant negative effects in purified mixed DNMT3A/R882H complexes. *Sci Rep*. 2018;8(1):13242.
 38. Jeong M, et al. Large conserved domains of low DNA methylation maintained by Dnmt3a. *Nat Genet*. 2014;46(1):17–23.
 39. Yan XJ, et al. Exome sequencing identifies somatic mutations of DNA methyltransferase gene DNMT3A in acute monocytic leukemia. *Nat Genet*. 2011;43(4):309–315.
 40. Qu Y, et al. Differential methylation in CN-AML preferentially targets non-CGI regions and is dictated by DNMT3A mutational status and associated with predominant hypomethylation of HOX genes. *Epigenetics*. 2014;9(8):1108–1119.
 41. Jeong M, et al. Loss of Dnmt3a immortalizes hematopoietic stem cells in vivo. *Cell Rep*. 2018;23(1):1–10.
 42. Ito S, et al. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science*. 2011;333(6047):1300–1303.
 43. Ko M, et al. Ten-eleven-translocation 2 (TET2) negatively regulates homeostasis and differentiation of hematopoietic stem cells in mice. *Proc Natl Acad Sci U S A*. 2011;108(35):14566–14571.
 44. Maiti A, Drohat AC. Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites. *J Biol Chem*. 2011;286(41):35334–35338.
 45. Hashimoto H, et al. Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation. *Nucleic Acids Res*. 2012;40(11):4841–4849.
 46. Inoue A, Zhang Y. Replication-dependent loss of 5-hydroxymethylcytosine in mouse preimplantation embryos. *Science*. 2011;334(6053):194.
 47. Zhang Q, et al. Tet2 is required to resolve inflammation by recruiting Hdac2 to specifically repress IL-6. *Nature*. 2015;525(7569):389–393.
 48. Li Z, et al. Deletion of Tet2 in mice leads to dysregulated hematopoietic stem cells and subsequent development of myeloid malignancies. *Blood*. 2011;118(17):4509–4518.
 49. Moran-Crusio K, et al. Tet2 loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. *Cancer Cell*. 2011;20(1):11–24.
 50. Ostrand EL, et al. Divergent effects of Dnmt3a and Tet2 mutations on hematopoietic progenitor cell fitness. *Stem Cell Reports*. 2020;14(4):551–560.
 51. Quivoron C, et al. TET2 inactivation results in pleiotropic hematopoietic abnormalities in mouse and is a recurrent event during human lymphomagenesis. *Cancer Cell*. 2011;20(1):25–38.
 52. Shide K, et al. TET2 is essential for survival and hematopoietic stem cell homeostasis. *Leukemia*. 2012;26(10):2216–2223.
 53. Izzo F, et al. DNA methylation disruption reshapes the hematopoietic differentiation landscape. *Nat Genet*. 2020;52(4):378–387.
 54. Rasmussen KD, et al. Loss of TET2 in hematopoietic cells leads to DNA hypermethylation of active enhancers and induction of leukemogenesis. *Genes Dev*. 2015;29(9):910–922.
 55. Tulstrup M, et al. TET2 mutations are associated with hypermethylation at key regulatory enhancers in normal and malignant hematopoiesis. *Nat Commun*. 2021;12(1):6061.
 56. Abegunde SO, et al. An inflammatory environment containing TNF α favors Tet2-mutant clonal hematopoiesis. *Exp Hematol*. 2018;59:60–65.
 57. Cai Z, et al. Inhibition of inflammatory signaling in Tet2 mutant preleukemic cells mitigates stress-induced abnormalities and clonal hematopoiesis. *Cell Stem Cell*. 2018;23(6):833–849.
 58. Dang L, et al. Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. *Nature*. 2009;462(7274):739–744.
 59. Ward PS, et al. The common feature of leukemia-associated IDH1 and IDH2 mutations is a neomorphic enzyme activity converting α -ketoglutarate to 2-hydroxyglutarate. *Cancer Cell*. 2010;17(3):225–234.
 60. Figueroa ME, et al. Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. *Cancer Cell*. 2010;18(6):553–567.
 61. Foskolou IP, et al. The two enantiomers of 2-hydroxyglutarate differentially regulate cytotoxic T cell function. *Cell Rep*. 2023;42(9):113013.
 62. Chowdhury R, et al. The oncometabolite 2-hydroxyglutarate inhibits histone lysine demethylases. *EMBO Rep*. 2011;12(5):463–469.
 63. Fujino T, Kitamura T. ASXL1 mutation in clonal hematopoiesis. *Exp Hematol*. 2020;83:74–84.
 64. Inoue D, et al. A novel ASXL1-OGT axis plays roles in H3K4 methylation and tumor suppression in myeloid malignancies. *Leukemia*. 2018;32(6):1327–1337.
 65. Medina EA, et al. ASXL1/2 mutations and myeloid malignancies. *J Hematol Oncol*. 2022;15(1):127.
 66. Abdel-Wahab O, et al. Deletion of Asxl1 results in myelodysplasia and severe developmental defects in vivo. *J Exp Med*. 2013;210(12):2641–2659.
 67. Wang J, et al. Loss of Asxl1 leads to myelodysplastic syndrome-like disease in mice. *Blood*. 2014;123(4):541–553.
 68. Inoue D, et al. Truncation mutants of ASXL1 observed in myeloid malignancies are expressed at detectable protein levels. *Exp Hematol*. 2016;44(3):172–176.
 69. Asada S, et al. Mutant ASXL1 cooperates with BAP1 to promote myeloid leukaemogenesis. *Nat Commun*. 2018;9(1):2733.
 70. Balasubramani A, et al. Cancer-associated ASXL1 mutations may act as gain-of-function mutations of the ASXL1-BAP1 complex. *Nat Commun*. 2015;6(1):7307.
 71. Wang L, et al. Epigenetic targeted therapy of stabilized BAP1 in ASXL1 gain-of-function mutated leukemia. *Nat Cancer*. 2021;2(5):515–526.
 72. Kohnke T, et al. Human ASXL1-mutant hematopoiesis is driven by a truncated protein associated with aberrant de-ubiquitination of H2AK119. *Blood Cancer Discov*. 2024;5(3):202–223.
 73. Inoue D, et al. Myelodysplastic syndromes are induced by histone methylation-altering ASXL1 mutations. *J Clin Invest*. 2013;123(11):4627–4640.
 74. Hsu YC, et al. The distinct biological implications of Asxl1 mutation and its roles in leukemogenesis revealed by a knock-in mouse model. *J Hematol Oncol*. 2017;10(1):139.
 75. Nagase R, et al. Expression of mutant Asxl1 perturbs hematopoiesis and promotes susceptibility to leukemic transformation. *J Exp Med*. 2018;215(6):1729–1747.
 76. Yang H, et al. Gain of function of ASXL1 truncating protein in the pathogenesis of myeloid malignancies. *Blood*. 2018;131(3):328–341.
 77. Gao X, et al. Role of ASXL1 in hematopoiesis and myeloid diseases. *Exp Hematol*. 2022;115:14–19.
 78. Fujino T, et al. Mutant ASXL1 induces age-related expansion of phenotypic hematopoietic stem cells through activation of Akt/mTOR pathway. *Nat Commun*. 2021;12(1):1826.
 79. Uni M, et al. Modeling ASXL1 mutation revealed impaired hematopoiesis caused by derepression of p16Ink4a through aberrant PRC1-mediated histone modification. *Leukemia*. 2019;33(1):191–204.
 80. Fabre MA, et al. The longitudinal dynamics and natural history of clonal haematopoiesis. *Nature*. 2022;606(7913):335–342.
 81. Williams N, et al. Driver mutation acquisition in utero and childhood followed by lifelong clonal evolution underlie myeloproliferative neoplasms. *Blood*. 2020;136(suppl 2):LBA-1.
 82. Dawoud AAZ, et al. Clonal myelopoiesis in the UK Biobank cohort: ASXL1 mutations are strongly associated with smoking. *Leukemia*. 2020;34(10):2660–2672.
 83. Buscarlet M, et al. Lineage restriction analyses in CHIP indicate myeloid bias for TET2 and multipotent stem cell origin for DNMT3A. *Blood*. 2018;132(3):277–280.
 84. Kralovics R, et al. A gain-of-function mutation of JAK2 in myeloproliferative disorders. *N Engl J Med*. 2005;352(17):1779–1790.

85. Levine RL, et al. Activating mutation in the tyrosine kinase JAK2 in polycythemia vera, essential thrombocythemia, and myeloid metaplasia with myelofibrosis. *Cancer Cell*. 2005;7(4):387–397.
86. Lippert E, et al. The JAK2-V617F mutation is frequently present at diagnosis in patients with essential thrombocythemia and polycythemia vera. *Blood*. 2006;108(6):1865–1867.
87. Baxter EJ, et al. Acquired mutation of the tyrosine kinase JAK2 in human myeloproliferative disorders. *Lancet*. 2005;365(9464):1054–1061.
88. Scott LM, et al. JAK2 exon 12 mutations in polycythemia vera and idiopathic erythrocytosis. *N Engl J Med*. 2007;356(5):459–468.
89. Nangalia J, Green AR. Myeloproliferative neoplasms: from origins to outcomes. *Blood*. 2017;130(23):2475–2483.
90. Pardanani AD, et al. MPLS15 mutations in myeloproliferative and other myeloid disorders: a study of 1182 patients. *Blood*. 2006;108(10):3472–3476.
91. Pikman Y, et al. MPLW515L is a novel somatic activating mutation in myelofibrosis with myeloid metaplasia. *PLoS Med*. 2006;3(7):e270.
92. Chachoua I, et al. Thrombopoietin receptor activation by myeloproliferative neoplasm associated calreticulin mutants. *Blood*. 2016;127(10):1325–1335.
93. Elf S, et al. Mutant calreticulin requires both its mutant C-terminus and the thrombopoietin receptor for oncogenic transformation. *Cancer Discov*. 2016;6(4):368–381.
94. Nivarthi H, et al. Thrombopoietin receptor is required for the oncogenic function of CALR mutants. *Leukemia*. 2016;30(8):1759–1763.
95. Araki M, et al. Activation of the thrombopoietin receptor by mutant calreticulin in CALR-mutant myeloproliferative neoplasms. *Blood*. 2016;127(10):1307–1316.
96. Balligand T, et al. Pathologic activation of thrombopoietin receptor and JAK2-STAT5 pathway by frameshift mutants of mouse calreticulin. *Leukemia*. 2016;30(8):1775–1778.
97. Marty C, et al. Calreticulin mutants in mice induce an MPL-dependent thrombocytosis with frequent progression to myelofibrosis. *Blood*. 2016;127(10):1317–1324.
98. Shide K, et al. Calreticulin mutant mice develop essential thrombocythemia that is ameliorated by the JAK inhibitor ruxolitinib. *Leukemia*. 2017;31(5):1136–1144.
99. Li J, et al. Mutant calreticulin knockin mice develop thrombocytosis and myelofibrosis without a stem cell self-renewal advantage. *Blood*. 2018;131(6):649–661.
100. Nangalia J, et al. Somatic CALR mutations in myeloproliferative neoplasms with nonmutated JAK2. *N Engl J Med*. 2013;369(25):2391–2405.
101. Klampfl T, et al. Somatic mutations of calreticulin in myeloproliferative neoplasms. *N Engl J Med*. 2013;369(25):2379–2390.
102. Kent DG, et al. Self-renewal of single mouse hematopoietic stem cells is reduced by JAK2V617F without compromising progenitor cell expansion. *PLoS Biol*. 2013;11(6):e1001576.
103. Li J, et al. JAK2 V617F impairs hematopoietic stem cell function in a conditional knock-in mouse model of JAK2 V617F-positive essential thrombocytopenia. *Blood*. 2010;116(9):1528–1538.
104. Scotti MM, Swanson MS. RNA mis-splicing in disease. *Nat Rev Genet*. 2016;17(1):19–32.
105. Chen S, et al. Splicing factor mutations in hematologic malignancies. *Blood*. 2021;138(8):599–612.
106. Yoshida K, et al. Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature*. 2011;478(7367):64–69.
107. Madan V, et al. Aberrant splicing of U12-type introns is the hallmark of ZRSR2 mutant myelodysplastic syndrome. *Nat Commun*. 2015;6(1):6042.
108. Shirai CL, et al. Mutant U2AF1 expression alters hematopoiesis and pre-mRNA splicing in vivo. *Cancer Cell*. 2015;27(5):631–643.
109. Kim E, et al. SRSF2 mutations contribute to myelodysplasia by mutant-specific effects on exon recognition. *Cancer Cell*. 2015;27(5):617–630.
110. Obeng EA, et al. Physiologic expression of SF3B1K(700E) causes impaired erythropoiesis, aberrant splicing, and sensitivity to pharmacologic spliceosome modulation. *Cancer Cell*. 2016;30(3):404–417.
111. Mupo A, et al. Hemopoietic-specific Sf3b1-K700E knock-in mice display the splicing defect seen in human MDS but develop anemia without ring sideroblasts. *Leukemia*. 2017;31(3):720–727.
112. Smeets MF, et al. Srsf2^{PSH} initiates myeloid bias and myelodysplastic/myeloproliferative syndrome from hemopoietic stem cells. *Blood*. 2018;132(6):608–621.
113. Papaemmanuil E, et al. Somatic SF3B1 mutation in myelodysplasia with ring sideroblasts. *N Engl J Med*. 2011;365(15):1384–1395.
114. Malcovati L, et al. SF3B1 mutation identifies a distinct subset of myelodysplastic syndrome with ring sideroblasts. *Blood*. 2015;126(2):233–241.
115. Clough CA, et al. Coordinated missplicing of TMEM14C and ABCB7 causes ring sideroblast formation in SF3B1-mutant myelodysplastic syndrome. *Blood*. 2022;139(13):2038–2049.
116. Dolatshad H, et al. Cryptic splicing events in the iron transporter ABCB7 and other key target genes in SF3B1-mutant myelodysplastic syndromes. *Leukemia*. 2016;30(12):2322–2331.
117. Chen L, et al. The augmented R-loop is a unifying mechanism for myelodysplastic syndromes induced by high-risk splicing factor mutations. *Mol Cell*. 2018;69(3):412–425.
118. Nguyen HD, et al. Spliceosome mutations induce R-loop-associated sensitivity to ATR inhibition in myelodysplastic syndromes. *Cancer Res*. 2018;78(18):5363–5374.
119. Singh S, et al. SF3B1 mutations induce R-loop accumulation and DNA damage in MDS and leukemia cells with therapeutic implications. *Leukemia*. 2020;34(9):2525–2530.
120. Beeram I. Accumulation of DNA damage in the aged hematopoietic stem cell compartment. *Semin Hematol*. 2017;54(1):12–18.
121. Haupt Y, et al. Mdm2 promotes the rapid degradation of p53. *Nature*. 1997;387(6630):296–299.
122. Honda R, et al. Oncoprotein MDM2 is a ubiquitin ligase E3 for tumor suppressor p53. *FEBS Lett*. 1997;420(1):25–27.
123. Shieh SY, et al. DNA damage-induced phosphorylation of p53 alleviates inhibition by MDM2. *Cell*. 1997;91(3):325–334.
124. Kasthuber ER, Lowe SW. Putting p53 in context. *Cell*. 2017;170(6):1062–1078.
125. Boettcher S, et al. A dominant-negative effect drives selection of TP53 missense mutations in myeloid malignancies. *Science*. 2019;365(6453):599–604.
126. Wong TN, et al. Role of TP53 mutations in the origin and evolution of therapy-related acute myeloid leukaemia. *Nature*. 2015;518(7540):552–555.
127. Coombs CC, et al. Therapy-related clonal hematopoiesis in patients with non-hematologic cancers is common and associated with adverse clinical outcomes. *Cell Stem Cell*. 2017;21(3):374–382.
128. Bolton KL, et al. Cancer therapy shapes the fitness landscape of clonal hematopoiesis. *Nat Genet*. 2020;52(11):1219–1226.
129. Arends CM, et al. Dynamics of clonal hematopoiesis under DNA-damaging treatment in patients with ovarian cancer. *Leukemia*. 2024;38(6):1378–1389.
130. Lindsley RC, et al. Prognostic mutations in myelodysplastic syndrome after stem-cell transplantation. *N Engl J Med*. 2017;376(6):536–547.
131. Rossi M, et al. Induction of PPM1D following DNA-damaging treatments through a conserved p53 response element coincides with a shift in the use of transcription initiation sites. *Nucleic Acids Res*. 2008;36(22):7168–7180.
132. Lu X, et al. PPM1D dephosphorylates Chk1 and p53 and abrogates cell cycle checkpoints. *Genes Dev*. 2005;19(10):1162–1174.
133. Burocziava M, et al. Ppm1d truncating mutations promote the development of genotoxic stress-induced AML. *Leukemia*. 2023;37(11):2209–2220.
134. Chen CW, et al. SRCAP mutations drive clonal hematopoiesis through epigenetic and DNA repair dysregulation. *Cell Stem Cell*. 2023;30(11):1503–1519.
135. Chen S, et al. Mutant p53 drives clonal hematopoiesis through modulating epigenetic pathway. *Nat Commun*. 2019;10(1):5649.
136. Loizou E, et al. A gain-of-function p53-mutant oncogene promotes cell fate plasticity and myeloid leukemia through the pluripotency factor FOXH1. *Cancer Discov*. 2019;9(7):962–979.
137. Desai P, et al. Somatic mutations precede acute myeloid leukemia years before diagnosis. *Nat Med*. 2018;24(7):1015–1023.
138. Vlasschaert C, et al. A practical approach to curate clonal hematopoiesis of indeterminate potential in human genetic data sets. *Blood*. 2023;141(18):2214–2223.
139. Young AL, et al. Clonal hematopoiesis and risk of acute myeloid leukemia. *Haematologica*. 2019;104(12):2410–2417.
140. NIH. SEER* Explorer: An interactive website for SEER cancer statistics. https://seer.cancer.gov/statistics-network/explorer/application.html?site=1&data_type=1&graph_type=2&compareBy=sex&chk_sex_3=3&chk_sex_2=2&rate_type=2&race=1&age_range=1&hdn_stage=101&advopt_precision=1&advopt_show_ci=on&hdn_view=0&advopt_show_apc=on&advopt_display=2#resultsRegion0. Accessed August 19, 2024.
141. Haferlach T, et al. Landscape of genetic lesions in 944 patients with myelodysplastic syndromes. *Leukemia*. 2014;28(2):241–247.

142. Papaemmanuil E, et al. Genomic classification and prognosis in acute myeloid leukemia. *N Engl J Med*. 2016;374(23):2209–2221.
143. Grinfeld J, et al. Classification and personalized prognosis in myeloproliferative neoplasms. *N Engl J Med*. 2018;379(15):1416–1430.
144. Gu M, et al. Multiparameter prediction of myeloid neoplasia risk. *Nat Genet*. 2023;55(9):1523–1530.
145. Watson CJ, et al. The evolutionary dynamics and fitness landscape of clonal hematopoiesis. *Science*. 2020;367(6485):1449–1454.
146. Khoury JD, et al. The 5th edition of the World Health Organization classification of haematolymphoid tumours: myeloid and histiocytic/dendritic neoplasms. *Leukemia*. 2022;36(7):1703–1719.
147. James C, et al. A unique clonal JAK2 mutation leading to constitutive signalling causes polycythemia vera. *Nature*. 2005;434(7037):1144–1148.
148. Weeks LD, et al. Prediction of risk for myeloid malignancy in clonal hematopoiesis. *NEJM Evid*. 2023;2(5).
149. Swaminathan M, et al. Hematologic malignancies and Li-Fraumeni syndrome. *Cold Spring Harb Mol Case Stud*. 2019;5(1):a003210.
150. Bernard E, et al. Implications of TP53 allelic state for genome stability, clinical presentation and outcomes in myelodysplastic syndromes. *Nat Med*. 2020;26(10):1549–1556.
151. Rodriguez-Meira A, et al. Single-cell multi-omics identifies chronic inflammation as a driver of TP53-mutant leukemic evolution. *Nat Genet*. 2023;55(9):1531–1541.
152. Gibson CJ, et al. Clonal hematopoiesis associated with adverse outcomes after autologous stem-cell transplantation for lymphoma. *J Clin Oncol*. 2017;35(14):1598–1605.
153. Quiros PM, et al. NPM1 gene mutations can be confidently identified in blood DNA months before de novo AML onset. *Blood Adv*. 2022;6(7):2409–2413.
154. Cancer Genome Atlas Research Network, et al. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N Engl J Med*. 2013;368(22):2059–2074.
155. Vassiliou GS, et al. Mutant nucleophosmin and cooperating pathways drive leukemia initiation and progression in mice. *Nat Genet*. 2011;43(5):470–475.
156. Borrow J, et al. Terminal deoxynucleotidyl transferase promotes acute myeloid leukemia by priming FLT3-ITD replication slippage. *Blood*. 2019;134(25):2281–2290.
157. Borrow J, et al. Molecular roulette: nucleophosmin mutations in AML are orchestrated through N-nucleotide addition by TdT. *Blood*. 2019;134(25):2291–2303.
158. Miller CA, et al. Failure to detect mutations in U2AF1 due to changes in the GRCh38 reference sequence. *J Mol Diagn*. 2022;24(3):219–223.
159. Caiado F, et al. Aging drives Tet2^{+/−} clonal hematopoiesis via IL-1 signaling. *Blood*. 2023;141(8):886–903.
160. Uddin MM, et al. Longitudinal profiling of clonal hematopoiesis provides insight into clonal dynamics. *Immun Ageing*. 2022;19(1):23.
161. Mack T, et al. Cost-effective and scalable clonal hematopoiesis assay provides insight into clonal dynamics. *J Mol Diagn*. 2024;26(7):563–573.