Evolution of myeloproliferative neoplasms from normal blood stem cells

Sahand Hormoz,[1-3] Vijay G. Sankaran[3-5] and Ann Mullally[6,7]

[1]Department of Systems Biology, Harvard Medical School, Boston, MA; [2]Department of Data Science, Dana-Farber Cancer Institute, Boston, MA; [3]Broad Institute of MIT and Harvard, Cambridge, MA; [4]Department of Pediatric Oncology, Dana-Farber Cancer Institute, Boston, MA; [5]Division of Hematology/Oncology, Boston Children's Hospital, Harvard Medical School, Boston, MA; [6]Division of Hematology, Department of Medicine, Stanford University School of Medicine, Stanford, CA and [7]Hematology Division, VA Palo Alto Health Care System, Palo Alto, CA, USA

Correspondence: S. Hormoz
sahand_hormoz@hms.harvard.edu
A. Mullally
amullal@stanford.edu

Received: July 11, 2024.
Accepted: November 19, 2024.
Early view: December 5, 2024.

https://doi.org/10.3324/haematol.2023.283951

Abstract

Over the course of the last decade, genomic studies in the context of normal human hematopoiesis have provided new insights into the early pathogenesis of myeloproliferative neoplasms (MPN). A preclinical phase of MPN, termed clonal hematopoiesis was identified and subsequent lineage tracing studies revealed a multi-decade long time interval from acquisition of an MPN phenotypic driver mutation in a hematopoietic stem cell to the development of overt MPN. Multiple germline variants associated with MPN risk have been identified through genome-wide association studies and in some cases functional interrogation of the impact of the variant has uncovered new insights into hematopoietic stem cell biology and MPN development. Increasingly sophisticated methods to study clonal contributions to human hematopoiesis and measure hematopoietic stem cell fitness have helped to discern the biology underlying the tremendous clinical heterogeneity observed in MPN. Despite these advances, significant knowledge gaps remain, particularly with respect to germline genetic contributors to both MPN pathogenesis and phenotypic diversity, as well as limitations in the ability to prospectively quantify rates of clonal expansion in individual MPN patients. Ultimately, we envisage a personalized approach to MPN care in the future, in which an individualized genetic assessment can predict MPN trajectory and this information will be used to inform and guide therapy. MPN is particularly amenable to precision medicine strategies and our increased understanding of the evolution of MPN from normal blood stem cells provides a unique opportunity for early therapeutic intervention approaches and potentially MPN prevention strategies.

Introduction

Myeloproliferative neoplasms (MPN) are chronic blood cancers, generally defined by their chronicity and clinical heterogeneity. This review focuses specifically on Philadelphia chromosome-negative MPN, which include polycythemia vera (PV), essential thrombocythemia (ET), and primary myelofibrosis (PMF). Philadelphia chromosome-negative MPN are driven by somatic mutations in key phenotypic driver genes such as *JAK2*, *CALR*, and *MPL*, which arise in the hematopoietic stem cell (HSC) compartment and result in diverse clinical manifestations observed across different MPN subtypes. MPN can evolve over time, with some patients progressing further to myelofibrosis or acute myeloid leukemia. Recent advances in genomic studies have provided new insights into the preclinical phase of MPN development and the mechanisms underlying disease progression.

Although some of the clinical heterogeneity of MPN is explained by the different causative somatic driver mutations (e.g., *JAK2 versus CALR*), this is not the full story. In the last 10 years, we have gained a wealth of knowledge from genomic studies of hematopoiesis in healthy individuals which has informed our understanding of MPN. Recent studies have uncovered a preclinical stage of MPN known as clonal hematopoiesis (CH), with lineage tracing studies showing that it can take several decades for an HSC harboring an MPN driver mutation to progress to full-blown MPN. Large genome-wide association studies (GWAS) have identified new MPN risk loci that are starting to reveal the biological mechanisms underlying MPN heritability as well as clinical heterogeneity. Excitingly, with increasingly so-

phisticated genomic approaches, we can measure clonal diversity in human HSC by exploiting "natural barcodes" (e.g., mutations) and quantify output of individual HSC. These advances have enabled detailed study of the evolution from normal polyclonal hematopoiesis to CH to overt MPN. In this review, we summarize key recent studies in the field and highlight the relevance of these findings in counseling patients and in developing new approaches focused on earlier therapeutic intervention and, potentially, MPN prevention.

Inherited risk of myeloproliferative neoplasms

For many years, studies on the pathogenesis of MPN have predominantly focused on the role and impact of acquired mutations that occur within hematopoietic stem cells,[1,2] including *JAK2*, *CALR*, and *MPL* mutations, which are clearly the important drivers underlying MPN development. However, the observation of familial aggregation of these disorders has long suggested an inherited component to the risk of developing MPN. Recent studies have substantiated this suspicion and advanced our understanding of the genetic basis of MPN, revealing new biological insights into their development.

Initial evidence for an inherited contribution to MPN came from the observation of their aggregation within families.[3,4] These familial clusters hinted at a genetic predisposition to the disease, but the specific mechanisms and the extent of the inherited risk remained unclear. This changed with the advent of large population-based studies, which provided a clearer picture of the genetic landscape underlying MPN. These studies demonstrated that first-degree relatives of individuals with MPN have an approximately 5-fold or greater increased risk of developing the disease compared to population controls.[5-7] These findings underscore the importance of inherited factors in the etiology of MPN and paved the way for subsequent genetic investigations.

One of the earliest and most significant discoveries in this area was the identification of the *JAK2* 46/1 haplotype as a key factor conferring an increased risk of MPN.[8-11] The *JAK2* 46/1 haplotype is associated with a roughly 2-fold increase in risk, accounting for a significant fraction of the genetic predisposition to MPN. This discovery was particularly important because it linked MPN to specific genetic variations, allowing for more targeted studies on the molecular mechanisms driving these chronic blood cancers.

Of note, while the identification of the *JAK2* 46/1 haplotype has been valuable, the precise mechanism by which an increased risk for acquiring MPN with this haplotype arises has remained unknown. Importantly, the risk alleles appear to increase the risk for a variety of MPN driver mutations beyond *JAK2* mutations.[11] While the variants do impact replication timing at this locus, this might simply be correlative, and the mechanisms through which this variation is acting have not been defined. This has partially been attributable to the significant extent of linkage between polymorphisms at this locus. Moreover, it is likely that there are multiple independent causal variants in this locus that occur in *JAK2* and adjacent genes, so further work is needed to define causal mechanisms.

The identification of the *JAK2* 46/1 haplotype was just the beginning. Subsequent GWAS have identified over 24 distinct loci associated with moderate (typically <1.5 to 2-fold odds), but significant impacts on the increased risk of MPN.[12-14] These studies have not only expanded the list of genetic factors associated with MPN but have also shed light on new biological pathways involved in their pathogenesis. Interestingly, many of the loci identified are correlated with blood cell phenotypes, such as leukocyte counts, and some functional data have suggested that these variants might impact the degree of hematopoietic cell self-renewal.[14] Indeed functional and correlative analyses of risk alleles at these loci have suggested that these variants tend to augment the extent of self-renewal of HSC, as exemplified by studies perturbing a disease-associated enhancer of the *GFI1B* gene,[14] suggesting a common and recurrent mechanism through which genetic variants can predispose individuals to MPN. Germline variants in the telomerase reverse transcriptase (*TERT*) locus have been identified as MPN risk alleles in several GWAS.[12-14] The most recent study by Bao *et al*. found that the top two independent *TERT* variants for increased telomere length were also leading variants for MPN risk, supporting a positive risk effect of increased telomere length on MPN acquisition.[14] This interpretation is consistent with a recent report showing that germline mutations in *POT1* associated with long telomere length conferred predisposition to a familial clonal hematopoiesis syndrome in which *JAK2* V617F was frequently the CH-associated driver mutation.[15] The authors proposed a model in which telomere-independent factors determine the risk of *JAK2* mutagenesis and long telomere length (as a consequence of mutant *POT1*) sustained the longevity of *JAK2*-mutant clones through enhanced replication competence. Beyond identifying risk loci, GWAS have revealed important insights into the biology of MPN. For instance, the association of certain loci with blood cell phenotypes and stem cell renewal dynamics has highlighted the complex interplay between genetic predisposition and disease manifestation. Whether these impacts of germline variants occur in the context of predisposition to MPN prior to driver mutation acquisition or whether these variants synergize to augment clonal expansion after acquisition remains poorly understood. It is possible that both of these impacts might have critical roles in disease pathogenesis. These findings suggest that the risk of developing MPN is influenced by factors that regulate stem cell behavior, including the balance between self-renewal and differentiation.

Familial cases have also been valuable. Studies of such

cases have identified rare duplications on chromosome 14 which involve a number of genes.[16,17] While initial studies had suggested that the *ATG2B* and *GSKIP* genes in this region contribute to the phenotype and indeed they do seem to impact HSC function,[18] consistent with results from GWAS, other studies have suggested that other genes in this region might be implicated.[19] Interestingly, the *TCL1A* gene that is in this region has also been implicated through common variant association studies for germline genetic modifiers of clonal expansion in CH.[20,21]

More recently, rare variant association studies have identified genes such as *CTR9*, where partial loss of function can augment transcription elongation and lead to greater HSC self-renewal.[22] This discovery underscores the role of transcriptional regulation in stem cell behavior and its impact on the development of MPN. Moreover, these studies highlight the importance of considering both common and rare genetic variants in understanding the genetic architecture of MPN. By better defining these and other mechanisms underlying inherited risk, unique opportunities for MPN prevention might be enabled.

The study of genetic modifiers has further enriched our understanding of MPN. For example, recent research has demonstrated how variants across the genome associated with different blood counts, including monocyte and platelet measurements, might influence the likelihood of an individual presenting with an MPN and the likelihood that *JAK2* mutant clones can expand.[23] However, we must exercise caution when interpreting precise causality in such studies, as confounders are possible and might influence multiple phenotypes. Nonetheless, these studies are important as they suggest that genetic variations can modify the clinical presentation of MPN, potentially affecting the rate of clonal expansion and disease progression. More studies are needed to examine how many of the variants identified to date might impact such phenotypes. Such findings emphasize the complexity of MPN pathogenesis and the role of genetic background in shaping disease outcomes.

The journey from observing familial aggregation to identifying specific genetic risk factors for MPN has been a transformative one. It has not only revealed the inherited contributions to the risk of acquiring MPN but has also opened new avenues for understanding the fundamental underpinnings of these myeloid malignancies. The identification of risk loci through GWAS and the insights gained from studies on rare variants and genetic modifiers have significantly advanced our knowledge of MPN pathogenesis. As research continues, it is likely that additional regions of interest will be identified, further elucidating the genetic landscape of MPN. These advances hold promise for the development of more targeted therapies and personalized approaches to managing MPN, ultimately improving outcomes for individuals affected by these complex and heterogeneous chronic blood cancers.

Clonal evolution of healthy hematopoietic stem cells

Even in healthy individuals, without a diagnosis of CH or MPN, HSC exhibit non-trivial clonal dynamics. HSC generate about a hundred billion new blood cells every day to replace dying cells.[24,25] To generate new blood cells, HSC divide asymmetrically to make a progenitor cell and another HSC. It is also possible for an HSC to divide symmetrically to make two HSC. This first became evident through engraftment experiments. Expansion of HSC after transplantation is direct evidence that HSC can undergo symmetric division and replicate themselves.[26] However, whether HSC undergo symmetric divisions and the rates at which these symmetric divisions occur in steady-state hematopoiesis in humans remained open questions. In a seminal study, Catlin *et al.*[27] measured the changing ratio of maternal and paternal X-chromosome gene expression in blood cells from females and inferred that HSC divide symmetrically once every 40 weeks. To reach this conclusion, the authors of the paper assumed that the replication rate of HSC expressing a maternal X-chromosome was different from that of HSC expressing a paternal chromosome due to germline mutations. The extent of this variability was taken from direct observations of differences in replication rates of HSC in different cat species partially attributed to X-chromosome polymorphisms.

Differences of replication rates across different HSC clones do not have to be caused by germline mutations alone. They can also be caused by somatic mutations accrued by the HSC or by the stochastic nature of cell division. The most direct evidence for the role of somatic mutations came from CH studies in which bulk sequencing of peripheral blood of healthy individuals was performed.[28,29] These studies showed that certain somatic mutations occur at high frequencies across healthy individuals and that these frequencies increase with age. We discuss below efforts to quantify the fitness advantage provided by a particular somatic mutation.

Finally, the clonal structure of HSC can also change because of the stochastic nature of cell divisions alone. As long as the probability of an HSC dividing symmetrically into two HSC is equal to the probability of an HSC dividing into two progenitor cells or dying then the total number of HSC in the population remains constant. Over time, just by chance, the number of descendants of a particular HSC can grow in number whereas the number of descendants of another HSC can decrease or diminish altogether. This neutral dynamics has been documented in other types of somatic stem cells, such as stem cells in the epidermis and the intestinal crypt.[30-32] Lee-Six *et al.* reconstructed the phylogenetic tree of individual HSC obtained from a 59-year-old man using the patterns of somatic mutations accrued in each cell.[33] The coalescent structure of this tree was consistent with neutral dynamics. In addition,

the authors inferred that HSC divide every 2 to 20 months and estimated the number of HSC that actively make white blood cells at any given time to be around 50,000 to 200,000 cells.

In a subsequent paper by Mitchell *et al.*[34] the same analysis was performed on ten individuals who ranged from newly born to 81 years old. In this larger cohort, hematopoiesis in the younger individuals, those under the age of 65 years, exhibited a highly polyclonal structure with tens of thousands of HSC clones contributing to blood production, consistent with neutral dynamics of HSC clones. In contrast, in individuals aged over 75 years blood production was maintained by tens of clones. These clones had accrued a fitness advantage over the other clones and had expanded over decades to dominate the stem cell pool. Strikingly, in only roughly 20% of the clones could the expansion be attributed to a known driver mutation such as a CH mutation. Therefore, even in healthy individuals a small subset of HSC clones eventually dominates the stem cell pool because of a fitness advantage. In the vast majority of these cases the driver mutation for this fitness advantage is unknown and cannot be identified by compiling commonly occurring somatic mutations across the population.[35-37] What causes some HSC clones to expand remains an intriguing unresolved problem. Perhaps epigenetic factors are at play, or the driver mutations are so rare that they are not shared across multiple people and therefore remain unidentified. Taken together, these observations paint a picture in which the clonal structure of HSC is dominated by selection. However, the fitness advantage of the clones that are selected for is relatively small and the clonal expansion requires decades to unfold, only manifesting later in life. In the minority of clones in which the fitness advantage is due to a known driver mutation associated with hematologic disorders, this expansion is referred to as CH. A natural question is what is the fitness advantage provided by a specific mutation and how does it vary across different individuals?

Lineage tracing studies in myeloproliferative neoplasms

One approach to determine when a driver mutation occurs and the fitness of the resulting mutant clone is to reconstruct the phylogenetic tree of individual HSC. Recently, Van Egeren *et al.*[38] applied tools from phylogenetics intended for species to individual cancer cells to reconstruct the history of disease expansion in individual patients. Building on prior work in the context of normal hematopoiesis,[33] the authors obtained HSC from the bone marrow of MPN patients, cultured individual stem cells for 8 weeks to increase their numbers, and performed whole genome sequencing on the resulting colonies. They then used the patterns of somatic mutations accrued in the genomes of the individual stem cells to reconstruct their phylogenetic tree. This study also looked at the prevalence of the *JAK2* mutation in other hematopoietic compartments besides the stem cells. The variant allele fraction of the *JAK2* mutation varied significantly across hematopoietic compartments. For instance, granulocytes - dominant contributors to the VAF signal in bulk sequencing of peripheral blood - exhibited a different variant allele fraction compared to red blood progenitors. Moreover, the *JAK2* mutation was even detectable in non-myeloid lineages.

Strikingly, they discovered that the *JAK2* V617F mutation occurred decades before diagnosis - at age 9±2 in a 34-year-old patient and age 19±3 in a 63-year-old patient. The population of *JAK2*-mutant HSC grew exponentially by 63±15% and 44±13% every year in the two patients, respectively. Similarly, Williams *et al.*[39] reconstructed phylogenetic trees of HSC from 12 MPN patients and also observed that the *JAK2* mutation occurred in childhood or even *in utero* in some patients and that the mutant clone grew at a rate that ranged from 3% to 190% in individual patients. *In utero* acquisition has also been reported for the *CALR* mutation, another key MPN phenotypic driver. Sousos *et al.* recently described a case of monozygotic twins who presented at the ages of 37 and 38 years old with *CALR*-mutant myelofibrosis.[40] DNA sequencing analyses identified a type 1 *CALR* mutation in myeloid cells and no mutation detectable in T cells or dermal fibroblasts, indicating a somatic origin for the *CALR* mutation. Whole-genome sequencing lineage tracing studies found a shared *CALR*-mutant clone present in both twins consistent with somatic acquisition of the *CALR* mutation in one twin, followed by twin-to-twin transplacental transmission, and subsequent separate postnatal clonal evolution in each twin (e.g., one twin acquired a *TET2* mutation). Both twins developed myelofibrosis with similar MPN latency. Taken together, these lineage tracing studies have shown that the driver MPN mutation occurs in an HSC and that the expansion of the mutated clone may take decades.

To better understand how the expansion history of mutated HSC can be inferred from a phylogenetic tree, let us examine such a tree (Figure 1). The length of the branch from the root of the tree to the most recent common ancestor of all the cancer stem cells is approximately proportional to the age at which the cancer mutation first occurs. For example, the shorter the branch, the earlier the mutation occurred in life. Similarly, the coalescent structure of the tree contains information about how fast the population expanded. The faster the population expands, the less time the population spends at the sufficiently small size where most cells are closely related, resulting in a shorter period of time along the phylogenetic tree where multiple branches coalesce together. Therefore, the statistics of the coalescence of the branches contain information about the fitness of the mutated HSC.[41,42]

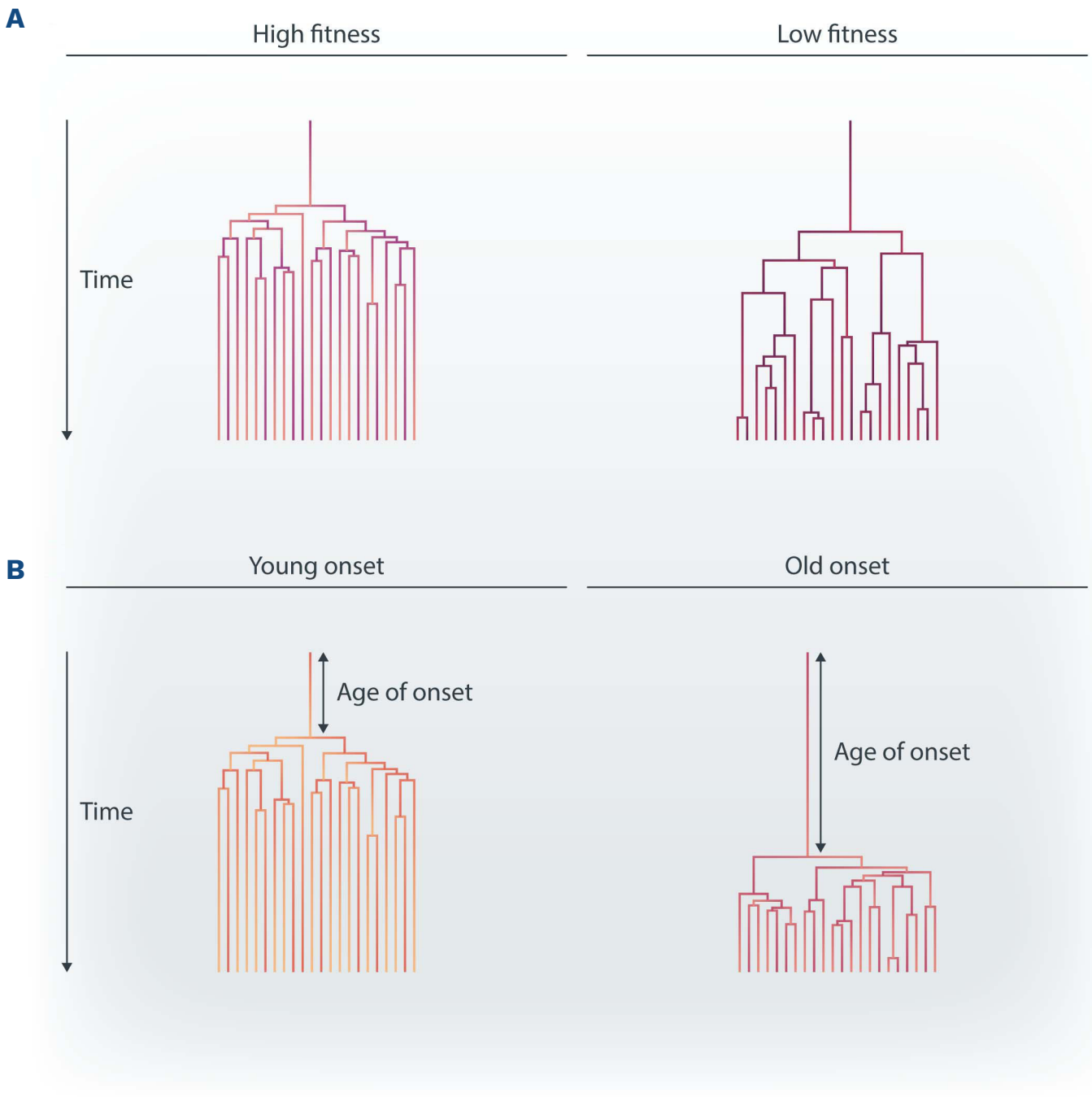

Figure 1. Schematic of phylogenetic trees of blood stem cells that have undergone clonal expansion. The length of the top branch of the tree is proportional to the age at which the fitness advantage (for example through a driver mutation) was obtained by this clone of cells. The extent of the fitness advantage is captured in the coalescent structure of the trees. (A) Higher fitness results in a larger number of coalescent events in a shorter period of time because the mutant clone grew faster and therefore spent less time as a small clone. (B) A later age of onset results in a longer top branch of the tree because the cell in which the driver mutation occurred had more time to accrue somatic mutations.

In mouse models of MPN, cells can be engineered to record their lineage histories within their own DNA. This innovation allows for the construction of phylogenetic trees of HSC without relying on naturally occurring somatic mutations. Recently, two such mouse models, called CARLIN[43] and DARLIN,[44] have been developed. These models utilize Cas9 to generate edits within synthetic target sites engineered into the genome. The target sites are transcribed into mRNA and can be read using single-cell RNA sequencing techniques alongside each cell's full transcriptome. These edits uniquely mark each HSC in the mouse, enabling researchers to track their subsequent clonal expansion. Usart *et al.*[45] recently developed the first lineage-recording mouse model of MPN. Such models hold great promise for elucidating the dynamics of clonal expansion of *JAK2*-mutated HSC.

Quantifying clonal expansion in *JAK2*-mutant clonal hematopoiesis

A surprising observation about the fitness of the *JAK2*-mutated clone in MPN patients inferred from the phylogenetic trees of individual HSC was the large variability across different patients. Could this be related to the large variability in rate of progression of MPN across patients? To answer this question, we need to measure fitness of the *JAK2* mutation in a large cohort of patients. Here we review some of these efforts in both CH and MPN. These studies fall into two broad categories: those that infer the fitness of the mutation from measurements at a single timepoint, and those that directly estimate the rate of growth of the mutant clone from longitudinal data.

Two studies have developed clever approaches to infer fitness from measurements done at a single timepoint. Watson *et al.*[46] used the prevalence of common CH driver mutations at a given variant allele frequency in a large cohort of healthy individuals to infer the fitness of the mutations. The idea is that a mutation with a high allele frequency either has a low fitness but occurred early in life or has a high fitness and occurred later in life. Therefore, looking at the prevalence of a mutation at high allele frequencies in a population and accounting for the age of the individuals in that population can inform us about the fitness of the mutation. Their approach makes some simplifying assumptions, the most important of which is that the fitness of the mutation is the same across all individuals. They inferred a fitness advantage of roughly 15% per year for the *JAK2* V617F mutation.

In another study, Weinstock *et al.*[20] used a similar expected relationship between fitness and age of occurrence of the driver mutation. The authors reasoned that if two mutations are observed at the same allele fraction, the fitter mutation would have occurred more recently because the resulting clone would have required less time to expand. More recent mutations have more time to accrue passenger somatic mutations. Therefore, the number of passenger somatic mutations can be used to infer fitness of a driver mutation in each subject. The authors measured the number of passenger somatic mutations at the same allelic fraction as the driver mutation in roughly 5,000 individuals with CH mutations in the TOPMed cohort. Because of the limited sequencing depth, the authors could only infer the relative fitness of driver mutations compared with a reference mutation (non-R882 *DNMT3A* mutations in this case). *JAK2* V617F was observed to be one of the fastest growing mutations.

The most direct way to infer fitness is by measuring the size of the mutant clones across decades. A recent longitudinal study by Fabre *et al.*[47] detected the *JAK2* mutation in the peripheral blood of a subset of the subjects in a cohort of approximately 400 Sardinian healthy individuals whose blood was sampled up to five times over 13 years. This study observed that the rate of increase of the fraction of peripheral blood cells with mutated *JAK2* showed the highest variability across individuals compared with mutations in all the other genes that are commonly mutated in blood disorders, such as *DNMT3A* and *TET2*. It should be noted that a longitudinal study by Robertson *et al.*,[48] which tracked the *JAK2* mutation variant allele fraction in three individuals from the Lothian Birth Cohorts, found less variability in the expansion rate of *JAK2*-mutated clones compared to those with other CH mutations. Taken together, the variability in the fitness of the *JAK2*-mutant clone across patients may not be a generic property of all CH driver mutations but unique to the *JAK2* mutation.

All of the above studies were focused on CH. However, a study[49] in 2012 followed the progression to MPN in 48 individuals in whom the *JAK2* mutation was accidentally detected as a part of the Copenhagen study that sequenced the blood of 50,000 healthy volunteers. The fraction of *JAK2*-mutated blood stem cells grew at different rates across the 48 individuals and was higher in the subset of volunteers who were later diagnosed with MPN. Taken together, these studies corroborate the findings from reconstructed phylogenetic trees of HSC: the fitness of *JAK2*-mutated HSC is highly variable across different individuals and potentially connected to disease progression.

Each approach to estimating clonal fitness has distinct strengths and limitations. The prevalence-based approach (Watson *et al.*[46]) provides population-level insights but assumes constant fitness across individuals and time. The passenger mutation burden approach (Weinstock *et al.*[20]) can directly estimate relative fitness by counting somatic mutations, but it requires deep sequencing and assumes constant growth rates. Additionally, it becomes challenging to interpret the somatic mutation burden when multiple clones are present. Longitudinal studies (Fabre *et al.*[47]) track clonal growth over time in peripheral blood, providing direct measurements of fitness but require longitudinal samples and are limited to the specific timepoints sampled. The phylogenetic tree analysis approach (Van Egeren *et al.*,[38] Williams *et al.*[39]) uses coalescent events to infer fitness, primarily capturing early clonal expansion when populations are small. This method relies on extensive single-cell DNA sequencing, making it resource-intensive. Given these constraints, combining insights from multiple methods may provide the most comprehensive view of clonal fitness across the different stages of disease.

Unique opportunities exist with the development of improved tools for lineage tracing in humans. For example, while current whole-genome sequencing approaches to track somatic mutations in individual colonies or clonal populations can be expensive and limited in terms of the number of samples that can be examined, inferential methods using deep whole-genome sequencing of bulk samples could enable improved insights into the clonal architecture of these diseases, using approaches similar to those described above for assessing fitness across a single timepoint in individuals. In addition, with advances in single-cell genomics, there are opportunities to gain valuable information on both cell state, including gene expression or accessible chromatin, with information on either the presence of specific driver mutations underlying the clonal expansions, markers that can track clonal populations and allow phylogenetic inference such as mitochondrial DNA mutations, or combinations of these different markers.[50-52] Further development of these and other approaches could enable critical new insights into the factors driving different behaviors among clones with similar driver mutations, including germline variation, order of mutations,[53] and altered cell states.[54-56]

Clinical implications and future directions

The genomic studies outlined in this review have reshaped our understanding of MPN pathogenesis. As findings from these studies have emerged, physicians and researchers are figuring out the clinical implications of this new knowledge. Prior work, such as that by Abelson *et al.*,[57] demonstrated gene-specific risk of leukemic progression in CH. Two recent studies have provided a framework to help predict the risk of progression from CH to an overt myeloid neoplasm, including MPN.[58,59] Both studies analyzed sequenced exomes of healthy United Kingdom Biobank participants to identify individuals with CH and studied the association between specific CH-associated mutations as well as other laboratory parameters and the emergence of overt myeloid neoplasia. Weeks *et al*. developed the Clonal Hematopoiesis Risk Score (*http://www.chrsapp.com/*) as a tool to distinguish a minority of individuals with CH who are at high risk of progression to myeloid neoplasia from the majority who have minimal risk of progression.[58] Gu *et al*. developed Myeloid Neoplasm (MN)-predict and quantified hazard ratios for each predictive variable (e.g., *JAK2* V617F mutation) with risk of a specific myeloid neoplasm (i.e., MPN, myelodysplastic syndrome or acute myeloid leukemia).[59] Weeks *et al*. identified *JAK2* V617F as a high-risk mutation for progression to myeloid neoplasia (hazard ratio = 32 in univariate analysis) while Gu *et al*. found that *JAK2* V617F was associated with increased risk of developing MPN specifically.[58,59] Interestingly, Gu *et al*. also noted that the platelet count was significantly higher in individuals who subsequently developed MPN in the 5-year period prior to MPN diagnosis as compared to that in controls, even though the platelet count often fell within the high normal range in this “pre-MPN” phase. Currently, these predictive models are useful for clinicians in counseling individuals who are identified to have a *JAK2* V617F mutation but who do not meet diagnostic criteria for any of the MPN subtypes (i.e., ET, PV or PMF).

An enhanced knowledge of MPN pathogenesis also raises questions regarding therapeutic implications. As we begin to dissect the genetic and biological factors that influence the clonal expansion rate of MPN stem cells within an individual, we can start to consider stratification of risk and ultimately link this to therapeutic intervention. If it were possible to identify those individuals/patients in whom MPN stem cell expansion is occurring at a particularly rapid rate, a rationale in support of early therapeutic intervention in these individuals/patients can be made (i.e., a precision medicine approach). However, designing clinical trials in individuals with CH raises challenges around clinical trial design, defining appropriate endpoints and ethical considerations. Preclinical animal modeling has the potential to help to provide data and a rationale in support of intervention at the CH stage as it relates to MPN. Multiple genetic mouse models of MPN have been developed and published.[60] To date, these models have typically employed hematopoietic-specific transgenic Cre mice (e.g., Mx1-Cre) that bring about simultaneous expression of *JAK2* V617F or mutant *CALR* in all hematopoietic cells. To accurately model CH, expression of the genetic driver mutation should be induced in a small fraction of HSC only. Bone marrow transplantation using a low ratio of mutant to wild-type donor bone marrow cells has been the most common approach used to achieve this,[61] although a limitation of which is the need to irradiate recipient mice and in doing so perturb the bone marrow microenvironment. To avoid the need for irradiation, adoptive transfer was employed to study the effects of *Tet2*-mutant hematopoietic cells on the heart.[62] As noted above, recently, an abstract describing a genetic lineage tracing approach (i.e., CARLIN) to track individual *JAK2*-mutant HSC *in vivo* was reported.[45] A challenge of genetic lineage tracing as an approach for preclinical modeling is the long duration it takes to breed the relevant alleles before the experimental question can be addressed, e.g., what happens to the trajectories of individual *JAK2*-mutant HSC in CH when a therapeutic selective pressure is applied. In MPN, there is a long history of preclinical modeling in mice supporting the development of new therapies, including for JAK inhibitors, pegylated interferons and several others.[63-67]

In humans, therapeutic interventions could take the form of clinical trials in individuals with *JAK2*-mutant CH at high risk of progression to MPN and/or in patients with ET and PV deemed to be at high risk of progression to myelofibrosis and/or acute myeloid leukemia. The types of therapeutic interventions could include approved MPN treatments, such as ropeginterferon, for which sustained reductions in *JAK2* V617F allele burden have been reported in treated PV patients.[68] Novel therapies with the potential for clonal selectivity would also be good candidates for testing in clinical trials. Historically, the MPN field has lacked treatments with clonal selectivity but with the development of rationally designed targeted therapies for MPN phenotypic driver mutations, there is optimism that this may change.[69,70] Currently, early-phase clinical trials testing a variety of mutant *CALR*-directed therapies are active. These include trials testing a monoclonal antibody (NCT06034002[71]), a T-cell-redirecting bispecific antibody (NCT06150157[72]), in addition to mutant CALR-directed vaccine (NCT05025488 – open; NCT05444530 – closed to accrual). Preclinical studies of a *JAK2* V617F-selective inhibitor[73] were recently reported, with phase I clinical trials of this drug (NCT06213818) and of a separate type II JAK2 inhibitor (NCT06343805)[74] both now open and recruiting. Individuals with *JAK2*-mutant CH and younger patients with ET and PV, who typically have an MPN phenotypic driver mutation as the sole genetic abnormality, have the potential to benefit from early therapeutic intervention, but the interventions must preferentially target MPN stem cells over normal HSC and be well tolerated for this benefit to be realized. Currently younger patients with ET and PV are often treated with pegylated interferons and

molecular responses have been reported.[68,75] Challenges in introducing pegylated interferon in the preclinical phase of MPN include the potential need for multi-year therapy and treatment tolerability.

Conclusion

It is an exciting time in MPN biology, with advances in genomics and single-cell technologies driving novel discoveries, with high potential for swift clinical translation. Notwithstanding the immediate unmet clinical need for new therapies for patients with established and advanced-phase MPN, a unique opportunity exists to exploit our rapidly evolving understanding of the early stages of MPN pathogenesis for therapeutic benefit. In the coming years, we hope that further scientific innovation in this area will advance personalized medicine approaches focused on early therapeutic intervention to positively alter the natural history of MPN in individual patients.

Disclosures

In the last 24 months, AM has received research funding from Relay and Morphic and consulted for Morphic, Protagonist, Incyte, Nuvalent, PharmaEssentia and Cellarity. VGS serves as an advisor to Ensoma. SH has no conflicts of interest to disclose.

Contributions

All authors wrote the paper together.

Funding

SH acknowledges funding from NHI NHLBI (R01HL158269 and R01HL158192), the Leukemia & Lymphoma Society (LLS Discovery Grant ID: 8041-24), MPN Research Foundation, and the Starr Cancer Consortium. AM acknowledges funding from NIH NHLBI (R01HL131835, R01 HL167139), Department of Defense Congressionally Directed Medical Research Programs (W81XWH2110909), the Leukemia & Lymphoma Society (LLS Discovery Grant ID: 8041-24) and the Starr Cancer Consortium. VGS acknowledges funding from the NIH (R01CA265726, R33CA278393, and R01DK103794), the Mather's Foundation, the New York Stem Cell Foundation, the Alex's Lemonade Stand Foundation, and the Howard Hughes Medical Institute.

References

1. Jamieson CHM, Gotlib J, Durocher JA, et al. The JAK2 V617F mutation occurs in hematopoietic stem cells in polycythemia vera and predisposes toward erythroid differentiation. Proc National Acad Sci. 2006;103(16):6224-6229.
2. Mead AJ, Mullally A. Myeloproliferative neoplasm stem cells. Blood. 2017;129(12):1607-1616.
3. Tashi T, Swierczek S, Prchal JT. Familial MPN predisposition. Curr Hematol Malig Rep. 2017;12(5):442-447.
4. Kralovics R, Stockton DW, Prchal JT. Clonal hematopoiesis in familial polycythemia vera suggests the involvement of multiple mutational events in the early pathogenesis of the disease. Blood. 2003;102(10):3793-3796.
5. Sud A, Chattopadhyay S, Thomsen H, et al. Familial risks of acute myeloid leukemia, myelodysplastic syndromes, and myeloproliferative neoplasms. Blood. 2018;132(9):973-976.
6. Sud A, Chattopadhyay S, Thomsen H, et al. Analysis of 153 115 patients with hematological malignancies refines the spectrum of familial risk. Blood. 2019;134(12):960-969.
7. Andersen MA, Bjerrum OW, Ranjan A, et al. Myeloproliferative neoplasms in Danish twins. Acta Haematol. 2018;139(3):195-198.
8. Jones AV, Chase A, Silver RT, et al. JAK2 haplotype is a major risk factor for the development of myeloproliferative neoplasms. Nat Genet. 2009;41(4):446-449.
9. Olcaydu D, Harutyunyan A, Jäger R, et al. A common JAK2 haplotype confers susceptibility to myeloproliferative neoplasms. Nat Genet. 2009;41(4):450-454.
10. Kilpivaara O, Mukherjee S, Schram AM, et al. A germline JAK2 SNP is associated with predisposition to the development of JAK2V617F-positive myeloproliferative neoplasms. Nat Genet. 2009;41(4):455-459.
11. Jones AV, Campbell PJ, Beer PA, et al. The JAK2 46/1 haplotype predisposes to MPL-mutated myeloproliferative neoplasms. Blood. 2010;115(22):4517-4523.
12. Hinds DA, Barnholt KE, Mesa RA, et al. Germ line variants predispose to both JAK2 V617F clonal hematopoiesis and myeloproliferative neoplasms. Blood. 2016;128(8):1121-1128.
13. Tapper W, Jones AV, Kralovics R, et al. Genetic variation at MECOM, TERT, JAK2 and HBS1L-MYB predisposes to myeloproliferative neoplasms. Nat Commun. 2015;6(1):6691.
14. Bao EL, Nandakumar SK, Liao X, et al. Inherited myeloproliferative neoplasm risk affects haematopoietic stem cells. Nature. 2020;586(7831):769-775.
15. DeBoy EA, Tassia MG, Schratz KE, et al. Familial clonal hematopoiesis in a long telomere syndrome. N Engl J Med. 2023;388(26):2422-2433.
16. Saliba J, Saint-Martin C, Stefano AD, et al. Germline duplication of ATG2B and GSKIP predisposes to familial myeloid malignancies. Nat Genet. 2015;47(10):1131-1140.
17. Pegliasco J, Hirsch P, Marzac C, et al. Germline ATG2B/GSKIP-containing 14q32 duplication predisposes to early clonal hematopoiesis leading to myeloid neoplasms. Leukemia. 2022;36(1):126-137.
18. Sakai S, Hasegawa A, Ishimura R, et al. Loss of Atg2b and Gskip impairs the maintenance of the hematopoietic stem cell pool size. Mol Cell Biol. 2022;42(1):e0002421.
19. Babushok DV, Stanley NL, Morrissette JJD, et al. Germline duplication of ATG2B and GSKIP genes is not required for the familial myeloid malignancy syndrome associated with the duplication of chromosome 14q32. Leukemia. 2018;32(12):2720-2723.
20. Weinstock JS, Gopakumar J, Burugula BB, et al. Aberrant activation of TCL1A promotes stem cell expansion in clonal haematopoiesis. Nature. 2023;616(7958):755-763.
21. Wen S, Kuri-Morales P, Hu F, et al. Comparative analysis of 136,401 Admixed Americans and 419,228 Europeans reveals ancestry-specific genetic determinants of clonal

haematopoiesis. medRxiv. 2024 Feb 7. doi: 10.1101/2024.02.07.24302442 [preprint, not peer-reviewed].
22. Zhao J, Cato LD, Arora UP, et al. Inherited blood cancer predisposition through altered transcription elongation. Cell. 2024;187(3):642-658.e19.
23. Guo J, Walter K, Quiros PM, et al. Inherited polygenic effects on common hematological traits influence clonal selection on JAK2V617F and the development of myeloproliferative neoplasms. Nat Genet. 2024;56(2):273-280.
24. Orkin SH, Zon LI. Hematopoiesis: an evolving paradigm for stem cell biology. Cell. 2008;132(4):631-644.
25. Mikkola HKA, Orkin SH. The journey of developing hematopoietic stem cells. Development. 2006;133(19):3733-3744.
26. Notta F, Doulatov S, Laurenti E, Poeppl A, Jurisica I, Dick JE. Isolation of single human hematopoietic stem cells capable of long-term multilineage engraftment. Science. 2011;333(6039):218-221.
27. Catlin SN, Busque L, Gale RE, Guttorp P, Abkowitz JL. The replication rate of human hematopoietic stem cells in vivo. Blood. 2011;117(17):4460-4466.
28. Jaiswal S, Fontanillas P, Flannick J, et al. Age-related clonal hematopoiesis associated with adverse outcomes. N Engl J Med. 2014;371(26):2488-2498.
29. Jaiswal S, Ebert BL. Clonal hematopoiesis in human aging and disease. Science. 2019;366(6465):eaan4673.
30. Simons BD, Clevers H. Strategies for homeostatic stem cell self-renewal in adult tissues. Cell. 2011;145(6):851-862.
31. Snippert HJ, Flier LG van der, Sato T, et al. Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Lgr5 stem cells. Cell. 2010;143(1):134-144.
32. Clayton E, Doupé DP, Klein AM, Winton DJ, Simons BD, Jones PH. A single type of progenitor cell maintains normal epidermis. Nature. 2007;446(7132):185-189.
33. 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.
34. Mitchell E, Chapman MS, Williams N, et al. Clonal dynamics of haematopoiesis across the human lifespan. Nature. 2022;606(7913):343-350.
35. Genovese G, Kähler AK, Handsaker RE, et al. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. N Engl J Med. 2014;371(26):2477-2487.
36. Zink F, Stacey SN, Norddahl GL, et al. Clonal hematopoiesis, with and without candidate driver mutations, is common in the elderly. Blood. 2017;130(6):742-752.
37. Bernstein N, Chapman MS, Nyamondo K, et al. Analysis of somatic mutations in whole blood from 200,618 individuals identifies pervasive positive selection and novel drivers of clonal hematopoiesis. Nat Genet. 2024;56(6):1147-1155.
38. Egeren DV, Escabi J, Nguyen M, et al. Reconstructing the lineage histories and differentiation trajectories of individual cancer cells in JAK2-mutant myeloproliferative neoplasms. bioRxiv 2020 Aug 24. doi: 10.1101/2020.08.24.265058 [preprint, not peer-reviewed].
39. Williams N, Lee J, Mitchell E, et al. Life histories of myeloproliferative neoplasms inferred from phylogenies. Nature. 2022;602(7895):162-168.
40. Sousos N, Leathlobhair MN, Karali CS, et al. In utero origin of myelofibrosis presenting in adult monozygotic twins. Nat Med. 2022;28(6):1207-1211.
41. Lan S, Palacios JA, Karcher M, Minin VN, Shahbaba B. An efficient Bayesian inference framework for coalescent-based nonparametric phylodynamics. Bioinformatics. 2015;31(20):3282-3289.
42. Karcher MD, Palacios JA, Lan S, Minin VN. phylodyn: an R package for phylodynamic simulation and inference. Mol Ecol Resour. 2017;17(1):96-100.
43. Bowling S, Sritharan D, Osorio FG, et al. An engineered CRISPR-Cas9 mouse line for simultaneous readout of lineage histories and gene expression profiles in single cells. Cell. 2020;181(6):1410-1422.e27.
44. Li L, Bowling S, McGeary SE, et al. A mouse model with high clonal barcode diversity for joint lineage, transcriptomic, and epigenomic profiling in single cells. Cell. 2023;186(23):5183-5199.e22.
45. Usart M, Almeida T, Roux J, et al. Cellular barcoding of JAK2-V617F hematopoietic stem cells reveals no substantial preferences in the contribution of individual stem cell clones to erythroid versus megakaryocytic lineages. Blood. 2023;142(Supplement 1):859.
46. Watson CJ, Papula AL, Poon GYP, et al. The evolutionary dynamics and fitness landscape of clonal hematopoiesis. Science. 2020;367(6485):1449-1454.
47. Fabre MA, Almeida JG de, Fiorillo E, et al. The longitudinal dynamics and natural history of clonal haematopoiesis. Nature. 2022;606(7913):335-342.
48. Robertson NA, Latorre-Crespo E, Terradas-Terradas M, et al. Longitudinal dynamics of clonal hematopoiesis identifies gene-specific fitness effects. Nat Med. 2022;28(7):1439-1446.
49. Nielsen C, Bojesen SE, Nordestgaard BG, Kofoed KF, Birgens HS. JAK2V617F somatic mutation in the general population: myeloproliferative neoplasm development and progression rate. Haematologica. 2014;99(9):1448-1455.
50. Weng C, Yu F, Yang D, et al. Deciphering cell states and genealogies of human haematopoiesis. Nature. 2024;627(8003):389-398.
51. Izzo F, Myers RM, Ganesan S, et al. Mapping genotypes to chromatin accessibility profiles in single cells. Nature. 2024;629(8014):1149-1157.
52. Sankaran VG, Weissman JS, Zon LI. Cellular barcoding to decipher clonal dynamics in disease. Science. 2022;378(6616):eabm5874.
53. Ortmann CA, Kent DG, Nangalia J, et al. Effect of mutation order on myeloproliferative neoplasms. N Engl J Med. 2015;372(7):601-612.
54. SanMiguel JM, Eudy E, Loberg MA, et al. Distinct tumor necrosis factor alpha receptors dictate stem cell fitness versus lineage output in Dnmt3a-mutant clonal hematopoiesis. Cancer Discov. 2022;12(12):2763-2773.
55. Cai Z, Kotzin JJ, Ramdas B, 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.e5.
56. Rodriguez-Meira A, Norfo R, Wen S, et al. Single-cell multi-omics identifies chronic inflammation as a driver of TP53-mutant leukemic evolution. Nat Genet. 2023;55(9):1531-1541.
57. Abelson S, Collord G, Ng SWK, et al. Prediction of acute myeloid leukaemia risk in healthy individuals. Nature. 2018;559(7714):400-404.
58. Weeks LD, Niroula A, Neuberg D, et al. Prediction of risk for myeloid malignancy in clonal hematopoiesis. NEJM Evid. 2023;2(5):10.
59. Gu M, Kovilakam SC, Dunn WG, et al. Multiparameter prediction of myeloid neoplasia risk. Nat Genet. 2023;55(9):1523-1530.

60. Jacquelin S, Kramer F, Mullally A, Lane SW. Murine models of myelofibrosis. Cancers (Basel). 2020;12(9):2381.
61. Hormaechea-Agulla D, Matatall KA, Le DT, et al. Chronic infection drives Dnmt3a-loss-of-function clonal hematopoiesis via IFNγ signaling. Cell Stem Cell. 2021;28(8):1428-1442.e6.
62. Wang Y, Sano S, Yura Y, et al. Tet2-mediated clonal hematopoiesis in non-conditioned mice accelerates age-associated cardiac dysfunction. JCI Insight. 2024;9(21):e187715.
63. Mullally A, Lane SW, Ball B, et al. Physiological Jak2V617F expression causes a lethal myeloproliferative neoplasm with differential effects on hematopoietic stem and progenitor cells. Cancer Cell. 2010;17(6):584-596.
64. Mullally A, Bruedigam C, Poveromo L, et al. Depletion of Jak2V617F myeloproliferative neoplasm-propagating stem cells by interferon-α in a murine model of polycythemia vera. Blood. 2013;121(18):3692-3702.
65. Dagher T, Maslah N, Edmond V, et al. JAK2V617F myeloproliferative neoplasm eradication by a novel interferon/arsenic therapy involves PML. J Exp Med. 2020;218(2):e20201268.
66. Kleppe M, Koche R, Zou L, et al. Dual targeting of oncogenic activation and inflammatory signaling increases therapeutic efficacy in myeloproliferative neoplasms. Cancer Cell. 2018;33(1):29-43.e7.
67. Casu C, Oikonomidou PR, Chen H, et al. Minihepcidin peptides as disease modifiers in mice affected by β-thalassemia and polycythemia vera. Blood. 2016;128(2):265-276.
68. Kiladjian J-J, Klade C, Georgiev P, et al. Long-term outcomes of polycythemia vera patients treated with ropeginterferon alfa-2b. Leukemia. 2022;36(5):1408-1411.
69. Kramer F, Mullally A. Antibody targeting of mutant calreticulin in myeloproliferative neoplasms. J Cell Mol Med. 2024;28(5):e17896.
70. Nair PC, Piehler J, Tvorogov D, et al. Next-generation JAK2 Inhibitors for the treatment of myeloproliferative neoplasms: lessons from structure-based drug discovery approaches. Blood Cancer Discov. 2023;4(5):352-364.
71. Reis E, Buonpane R, Celik H, et al. Discovery of INCA033989, a monoclonal antibody that selectively antagonizes mutant calreticulin oncogenic function in myeloproliferative neoplasms (MPNs). Blood. 2022;140(Supplement 1):14-15.
72. Kuchnio A, Samakai E, Hug E, et al. Discovery of JNJ-88549968, a novel, first-in-class CALRmutxCD3 T-cell redirecting antibody for the treatment of myeloproliferative neoplasms. Blood. 2023;142(Supplement 1):1777.
73. Stubbs MC, Celik H, Ai Y, et al. Preclinical evaluation of INCB160058 - a novel and potentially disease-modifying therapy for JAK2V617F mutant myeloproliferative neoplasms. Blood. 2023;142(Supplement 1):860.
74. Rai S, Stetka J, Usart M, et al. The second generation type II JAK2 inhibitor, AJ1-10502, demonstrates enhanced selectivity, improved therapeutic efficacy and reduced mutant cell fraction compared to type I JAK2 inhibitors in models of myeloproliferative neoplasms (MPNs). Blood. 2022;140(Supplement 1):6722-6723.
75. Quintás-Cardama A, Kantarjian H, Manshouri T, et al. Pegylated interferon alfa-2a yields high rates of hematologic and molecular response in patients with advanced essential thrombocythemia and polycythemia vera. J Clin Oncol. 2009;27(32):5418-5424.