INTEGRATIVE SINGLE CELL PROFILING AND POPULATION GENOMIC STUDIES TO CHARACTERIZE THE ROLE OF SOMATIC MUTATIONS IN HEMATOLOGIC MALIGNANCIES

by

Maria Sirenko

A Dissertation

Presented to the Faculty of the Louis V. Gerstner, Jr.

Graduate School of Biomedical Sciences,

Memorial Sloan Kettering Cancer Center

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

New York, NY

May 2023

_____ _____

Elli Papaemmanuil, PhD Date

Dissertation Mentor

Dedication

My dissertation is dedicated to my parents, Natalia and Andrei Sirenko, my Fiancé, Mark Unter, and my friends, for your unconditional love and support through my journey.

Abstract

Somatic mutations cause cancer and determine patient clinical presentation and outcomes. Systematic profiling of somatic mutations in cancer has led to 1) the adoption of mutation profiling to inform patient diagnosis and risk stratification and 2) the development and clinical translation of targeted therapeutic agents. However, the majority of patients with cancer harbor multiple genomic alterations(Greenman et al. 2007; Forbes et al. 2017). Understanding how serially acquired mutations cooperate to drive oncogenesis and determine clinical presentation is a major challenge in the field.

This thesis elucidates the role of acquired somatic mutations in two complementary and clinically relevant contexts: 1) Myelodysplastic Syndromes (MDS) patients who have newly discovered mutations in *UBA1*; and 2) high risk (*IDH1/2*-mutant) Acute Myeloid Leukemia (AML).

The incorporation of genetic information in MDS has enabled improved patient stratification and definition of molecularly-defined subgroups(Bernard et al. 2022, 2020). However, a subset of patients do not have any identified molecular alterations. Further, some patients have concomitant inflammatory presentations which complicates diagnosis and clinical management. Somatic mutations in *UBA1* have recently been linked to VEXAS (Vacuoles, E1 enzyme, X-linked, Autoinflammatory, Somatic) syndrome, an adult-onset autoinflammatory disorder. Approximately 40% of VEXAS patients are also diagnosed with MDS. However, the prevalence of *UBA1* mutations and associated clinical presentation in MDS patients are unknown. To understand the overlap between MDS with no genetic alterations in known myeloid pathogenesis genes and investigate a potential

etiology for inflammation-associated MDS, a large representative diagnostic and treatment-naive MDS cohort (n= 3,323) ascertained through the International Working Group for MDS was profiled by ddPCR (n=375) or by NGS panel (n=2027) for mutations in *UBA1*. 54 MDS patients with *UBA1* mutations were identified and correlation was performed with cytogenetic and sequencing data from 121 genes involved in myeloid pathogenesis to identify patterns of co-mutation. Clinical associations of *UBA1* mutations were evaluated and clinical history was reviewed for inflammatory conditions in 44 patients. *UBA1* mutation in isolation or in combination with other mutations implicated in myeloid pathogenesis accounted for a significant proportion (7%) of patients diagnosed as MDS but without established disease classification. This study informs clinical guidelines for the screening of *UBA1* mutations in MDS patients.

AML is an aggressive blood cancer with exceedingly poor patient outcomes(Kantarjian et al. 2021). Systematic sequencing studies in myeloid neoplasms have identified genes recurrently mutated in AML("Genomic and Epigenomic Landscapes of Adult De Novo Acute Myeloid Leukemia" 2013; Papaemmanuil et al. 2016), raising opportunities for the development of targeted therapeutics (e.g. *IDH1/2* and *FLT3*) and molecularly guided clinical care, whereby genetic information informs diagnosis, risk stratification and, in some cases, determines therapeutic options. To understand how serially acquired mutations shape disease presentation in the context of *IDH1/2*-mutant AML, 15 diagnostic patient samples were profiled using multimodal single cell profiling approaches inclusive of targeted DNA sequencing,single cell RNA-sequencing and cell surface marker profiling. Integrated single cell gene expression profiling and genotyping enabled the assignment of single cells to phylogenetic subclones and the delivery of clone-specific gene expression profiles. Patient samples were also profiled after treatment with *IDH1/2* targeted inhibitors to identify clonal responses to therapy and identify transcriptomic shifts

from diagnosis to remission and in a subset of cases, relapse. This study delivers detailed mapping of *IDH1/2*-mutant AML and a framework for dissecting the impact of sequentially acquired somatic mutations on disease phenotype at diagnosis and under therapy.

This thesis integrates large scale population genomic approaches and single cell profiling techniques to delineate the role of somatic mutations in two clinically relevant contexts: *UBA1*-mutant MDS and *IDH1/2*-mutant AML. This contributes to the understanding of how co-occurring mutations drive disease pathogenesis and will guide the development of new diagnostic and therapeutic strategies.

Acknowledgements

I would like to thank the many mentors I have had in my life that have been instrumental to my growth as an individual and in my career.

Thank you to my high school teachers, Connie Illian and Michael Gilmore who fostered a love of writing and science that I carry forward.

For my early research experiences in high school and undergrad, I thank Lisa Miller and Warren Zipfel. Lisa, I am grateful for the high level of independence you granted me, even as a high school and early undergraduate student. Warren, thank you for your mentorship during my time at Cornell and your support applying to graduate school and fellowships. Avtar Singh, thank you for your mentorship while in Zipfel lab and beyond. I am most thankful for your advice while I was applying to graduate school and fellowships. Thank you for reading my NSF GRFP application many, many, many times.

My tenure at MSKCC is a time I remember very fondly. MSKCC is an incredible place to learn and do science. I am so grateful for the colleagues, mentors, and friends I have made along the way. I hope our paths cross again soon.

Thank you to Dan Heller for mentoring me in the Gerstner Sloan Kettering Summer Undergraduate Research Program (SURP). The experience in your lab changed the trajectory of my career dramatically. I often think back upon your advice that the best place for an engineer to be is somewhere where you can find lots of problems to solve. You

mentioned the impact of walking down the hall, running into an oncologist and asking them what we can be doing better. I strive to do this every day.

Thank you to Gerstner Sloan Kettering Graduate School for both SURP and the PhD program. The amazing experience I had in SURP is the reason I decided to stay for my PhD. Thank you to Kenneth Marians, Michael Overholtzer and Linda Burnley for creating this amazing and truly one of a kind program. Thank you for always putting the students first and supporting me in taking courses, traveling for conferences, and giving me the opportunity to teach qcBio. Thank you to David McDonagh for being the world's best registrar. Thank you to Tom Magaldi, Stacy De La Cruz, Julie Masen, and Raphaelle Chassagne, Ivan Gerena and Maria Torres for making GSK a wonderful institution. Mike, on the first day of SURP, we took a group photo in the lobby of Zuckerman and you said to remember this moment because these friendships and collaborations will last a lifetime. You were right.

Thank you to Lauren Dennison Maslin and Matthew Boudreau, dear friends from SURP. I am grateful for your friendship and look forward to continuing to celebrate our successes.

I would also like to thank my classmates and friends in GSK for the camaraderie and always being willing to share your knowledge. In particular, thank you to Brittany Woods for your mentorship in the early days of my PhD and teaching me a lot about AML biology and a bit about mice. Mollie Chipman, I am so glad we went on this journey together and I am so grateful for your friendship.

Thank you to my thesis committee. Ross Levine and Omar Abdel-Wahab, thank you for your insights and guidance on both my projects and career development. Ross, thank you for welcoming me to your lab for experimental support and feedback in lab meetings. Andrew Intlekofer, thank you for serving on my Thesis Proposal Exam Committee and as the Chair on my Thesis Defense Committee. Thank you to Stavroula Kousteni for serving as my External Examiner for my thesis defense.

Thank you to Dan Landau for all your guidance and enthusiasm in the *IDH1/2* AML project. Thank you for welcoming me into your lab, where I spent the majority of the first few years. Ronan Chaligné, thank you for teaching me all things experimental single cell and hours of career advice. Thank you to Robert Myers for making the seemingly endless genotyping of transcriptomes PCRs pass by quickly with many laughs.

Thank you to Iannis Aifantis and the Aifantis lab, with whom the *IDH1/2* project has become a wonderful collaboration. Soobeom Lee and Zhengxi Sun, it has been a great pleasure to work together. I look forward to joining the lab as a Postdoctoral Fellow!

Thank you to Richard Koche for teaching qcBio with me for 4 years. Thank you to Kayleigh Rutherford, Georgios Asimomitis and Angel Tang for your help teaching as well. It has been immense fun to teach this course together and learn from you.

For career development and training, I would like to thank Ushma Neill, Thalyana Stathis and Yaihara Fortis. Thalyana Stathis, thank you for organizing amazing career panels and events that helped me realize all the things I can do with a PhD. Yaihara and Ushma, thank you for moving mountains for me ahead of the NSF GRFP deadline. Ushma, you

make MSKCC and GSK an amazing place. We are so lucky to have you and I am so grateful for your support. Ushma, I still can not believe you were checking the NSF GRFP results at 4am and found out even before I did!

Thank you to my benchmates: Erika Gedvilaite, Caitlin Stewart, Prachi Kothari, and Grittney Tam. Thank you for sharing with me your expertise on library prep and sequencing and many years of friendship beyond that.

Thank you to the Integrated Genomics Operation (IGO) team, including Marisa Dunigan, Neeman Mohibullah, Andrea Farina, Kety Huberman, and Agnes Viale, for your expertise, kindness, patience, technical support, and development of numerous bespoke assays. This thesis was made possible by data generated by IGO.

Thank you to the Center for Hematologic Malignancies project coordinators, including Chris Famulare and Minal Patel. Designing the patient cohorts for the *IDH1/2* AML project was not a simple task, but I regularly appreciate that the upfront effort was worth it to support a well-designed study.

Thank you to my dear friends and hematologists, Charlotte Bierley, Kamal Menghrajani, Sheng Cai and Kelly Bolton. First, thank you for answering all my clinical questions, of which there have been many. Your insights were crucial to driving this work forward. Most importantly, thank you for your boundless encouragement over the years. Charlotte, coffee and tea breaks with you were one of the highlights of my PhD. I am so grateful for our discussions regarding everything from cell type annotation to academic careers and beyond. I hope we have the chance to work together again soon. Kamal, thank you for always picking up the phone and your enthusiasm celebrating the sunny days with me.

Yay! Sheng, processing samples for single cell experiments went from a dreaded necessity to a day I looked forward to catching up with you and chatting about life, career, etc. Remember, no big experiments on a Monday. Kelly, thank you for the career advice and unceasing support.

Being a part of the Computational Oncology Service at MSKCC has been a source of regular inspiration and encouragement. I felt at home here and will miss learning from this impressive community of computational oncologists! I left every WIP seminar energized. Thank you to Sohrab Shah, Ben Greenbaum, Ed Reznik, Cynthia Berry and Nicole Rusk.

Thank you to Elli Papaemmanuil and the entire Papaemmanuil Lab. I joined the lab inspired by the possibility to take on translational research projects that would have real-world impact on cancer patients. I am grateful for the huge amount of independence I had in the lab and the freedom to lead a very challenging project in *IDH1/2* AML. I am also grateful for the support to screen MDS patients for *UBA1*. A highlight of this project is being able to identify MDS patients who can now receive a proper diagnosis of MDS and be treated accordingly. Both of these projects required immense support from the lab to succeed. I worked very closely with several lab members. Elsa Bernard, it has been a great pleasure to work together on the *UBA1* project and learn about MDS genomics and risk stratification from you. I am amazed by your contributions to the MDS field and am so excited for your next steps! Georgios Asimomitis, thank you for many stimulating single cell meetings and consulting on various analyses. Kayleigh Rutherford, I am so grateful for many years of your friendship and encouragement. I would say this has been a Great Success! Teng Gao, I am glad we had the chance to work together on the mCA CH project, I learned a lot from you in this time, especially dplyr. Noushin Farnoud, I am glad I learned how to annotate oncogenic variants from the expert. Gaelle Ouggourni, thank you for

teaching me how to bsub. Thank you to the engineers, Juan Medina, Juanes Arango-Ossa, and Max Levine for helping keep all data organized and pipelines running. Genesis Pineda, thank you for being wonderful and for organizing *everything*, especially my defense. Dylan Domenico, I am glad we shared an office and many Pomodoros and laughs together. Genesis and Dylan, I leave the party planning committee in your hands. Elli Papaemmanuil, thank you for your mentorship and for building a team of amazing scientists and engineers that I have learned a lot from. I am particularly grateful for your mentorship in scientific thinking, study design, writing, and presentations. I leave your lab as an independent and confident scientist!

Thank you to Laika for bringing endless joy as we read papers, coded analyses, and wrote this thesis together.

Thank you to my soon-to-be husband, Mark Unter, for taking this journey with me. I am thankful for your unwavering support and could not imagine this experience without you by my side.

Thank you to my parents, Andrei and Natalia Sirenko, for paving the way for me to have every opportunity. So far, I've achieved every goal I set, and it's in no small part thanks to your sacrifices and hard work. I hope I've made you proud!

Lastly, thank you to the patients and their families without whom none of this work would be possible.

Table of Contents

List of Figures ..... **xvi**

List of Tables ..... **xvii**

List of Abbreviations ..... **xviii**

Chapter 1: Introduction ***1***

Population genomics studies on large hematologic malignancies patient cohorts identify recurrently mutated genes, patterns of co-mutation and clinical associations **1**

Single cell technologies enable understanding of the biological consequences of somatic mutations in primary patient samples **2**

Chapter 2: Prevalence, co-mutation patterns, and clinical presentation of UBA1 mutations in Myelodysplastic Syndromes ***5***

Abstract **5**

Introduction **6**

Methods **7**

Results **10**

UBA1 p.M41T/V/L mutations in unclassifiable MDS or MDS without disease-defining mutations 10

UBA1 mutations in a representative MDS cohort 17

Clinical presentation of inflammatory and/or VEXAS-associated symptoms in *UBA1*-mutant MDS 28

Frameshift VUS in non-canonical *UBA1* transcripts in patients with inflammation and IRD 29

Association of *UBA1* VAF with inflammatory clinical presentation 32

Transformation to AML and overall survival 34

Discussion **36**

Chapter 3: High resolution mapping of IDH1/2-mutant Acute Myeloid Leukemia ***38***

Abstract **38**

Introduction **39**

Results **42**

Genomic landscape of *IDH1/2*-mutant AML 42

Single cell profiling of *IDH1/2*-mutant AML and CH 46

IDH1/2-Mutant AML at diagnosis is characterized by a stem-like and inflammatory phenotype 49

Integrative single cell profiling delineates clonal responses to combination of IDH1/2 inhibitors and intensive chemotherapy 55

Defining clone-specific phenotypes in diagnosis AML 62

Deriving mutation-associated transcriptomic signatures 70

Relapse after IDHi Monotherapy Occurs in the Context of Stemness and Inflammation and downregulation of MHC Class II 71

IDH2-mutant Clonal Hematopoiesis Exhibits Dysregulation of *IDH2*-associated Pathways 78

Discussion **81**

Methods **84**

Chapter 4: Perspectives and Future Directions ***98***

Summary **98**

Advances in single cell genomics technologies will enable clone-specific phenotyping within more complex clonal architectures **99**

The in-depth characterization of primary patient samples identifies patient-specific patterns of treatment response and resistance in *IDH1/2*-mutant AML **101**

Single-cell analyses of *IDH*-mutant CH and AML to understand high risk of *IDH* CH transformation **102**

Inflammation plays a key role in disease initiation, maintenance, and progression across the spectrum of CH, MDS and AML. **104**

Co-mutation analysis highlights *DNMT3A* and *TET2* as key interacting partners with *UBA1* mutations, highlighting the need for further investigation in VEXAS/MDS **105**

Integrative clone-specific gene expression analysis may identify therapeutic vulnerabilities in VEXAS/MDS **107**

Novel genomic drivers of hematologic malignancies and autoinflammatory diseases **109**

Bibliography ***111***

List of Figures

Figure 2-1 Schematic of study design showing Cohort A (top) profiled by ddPCR and Cohort B (bottom) profiled by NGS for UBA1 mutations. 11
Figure 2-2 Co-mutation patterns and clinical characteristics of UBA1 p.M41T/V/L mutant MDS patients in Cohort A 15
Figure 2-3 UBA1 mutations detected by ddPCR and NGS in Cohorts A and B .. 21
Figure 2-4 WHO 2016 classification and IPSS-M Risk Category of UBA1-mutant patients in total cohort (Cohort A+B) 22
Figure 2-5 Oncoprint for total cohort of UBA1-mutant patients (n=54) including results of retrospective review of clinical history for inflammatory features (bottom). Patients are ordered by decreasing UBA1 VAF per group (p.M41, non-p.M41 pathogenic, VUS). 25
Figure 2-6 Analysis of VAF of UBA1 and co-occurring mutations 27
Figure 2-7 Review of female patient with UBA1 frameshift VUS 31
Figure 2-8 Association of UBA1 VAF with clinical presentation of inflammatory symptoms 33
Figure 2-9 Survival analysis of UBA1 p.M41T/V/L mutant and wildtype MDS patients 35
Figure 3-1 Co-mutation landscape in IDH1/2-mutant AML informs study cohort 43
Figure 3-2 Marker expression informs cell type annotation 48
Figure 3-3 Stem-like and inflammatory phenotype of IDH1/2-mutant AML at diagnosis shaped by co-mutations in NPM1 and SRSF2 50
Figure 3-4 Gene expression programs specific to IDH1/2-mutant AML 54
Figure 3-5 Patient treatment overview and identification of somatic mutations and copy number alterations from single cell transcriptomic data 56
Figure 3-6 Tracking clonal and cellular evolution using integrated single cell gene expression and genotyping under combination of IDHi and intensive chemotherapy 58
Figure 3-7 Distinct phenotypes of a IDH1 mutant parent clone and NRAS co-mutant subclone in AML patient at diagnosis 63
Figure 3-8 Distinct phenotypes of a IDH2 mutant clone and SRSF2 co-mutant subclone 67
Figure 3-9 Genetic and transcriptomic patterns of relapse after IDHi monotherapy 72
Figure 3-10 Transcriptomic and cellular composition shifts in serial samples under IDHi monotherapy 77
Figure 3-11 Joint single cell gene expression profiling and genotyping in IDH2 mutant CH reveals upregulated inflammation and hypoxia pathways 79
Figure 4-1 Schematics of putative recurrent clonal structures in MDS: UBA1 co-mutation with A) DNMT3A and B,C) TET2. 106

List of Tables

Table 2-1 Cohort A Characteristics for UBA1 p.M41T/V/L -mutant and wildtype patients 12
Table 2-2 Characteristics of Cohort B (n=2,027). 18
Table 2-3 Characteristics of n=54 UBA1-mutant MDS cases. Table is split by UBA1 mutation: p.M41, non-p.M41 pathogenic, and VUS 20
Table 3-1 Supplementary Table 1: Cohort Characteristics 95

List of Abbreviations

AML: Acute Myeloid Leukemia
CH: Clonal hematopoiesis
CCUS: Clonal cytopenia of unknown significance
CNA: Copy number aberration
ddPCR: Digital droplet polymerase chain reaction
GMP: Granulocyte-monocyte progenitor
GTEx: Genotype-tissue expression database
HSC: Hematopoietic stem cell
HSPC: Hematopoietic stem and progenitor cell
IDH1/2: Isocitrate Dehydrogenase 1/2
INDEL: Insertions and deletions
IPSS-M: International Prognostic Scoring System – Molecular
IRD: Inflammatory and Rheumatologic Disease
MDS: Myelodysplastic Syndromes
MEP: Megakaryocyte-erythrocyte progenitor
NGS: Next Generation Sequencing
RTK-Ras: Receptor tyrosine kinase-Ras
scRNA-seq: Single cell ribonucleic acid sequencing
SNV: Single Nucleotide Variant
UBA1: Ubiquitin like modifier activating enzyme 1
VAF: Variant Allele Fraction
VEXAS: vacuoles, E1 enzyme, X-linked, autoinflammatory, somatic
VUS: Variant of Unknown Significance
WHO: World Health Organization

Chapter 1: Introduction

Population genomics studies on large hematologic malignancies patient cohorts identify recurrently mutated genes, patterns of co-mutation and clinical associations

Somatic mutations accumulate in normal tissues as a function of age, as well as with environmental exposures(Martincorena and Campbell 2015; Martincorena et al. 2015, 2018; Yizhak et al. 2019; Yokoyama et al. 2019). Within the hematopoietic system, the acquisition of mutations with age in genes that promote cell fitness result in clonal expansions, a phenomenon termed clonal hematopoiesis(Genovese et al. 2015). While this clonal expansion predisposes for leukemia, it is not sufficient for overt disease(Desai et al. 2018; Abelson et al. 2018; Young et al. 2019). Transformation is underwritten by the acquisition of secondary mutations. To identify the recurrently mutated genes in AML and other myeloid neoplasms that underlie disease initiation and progression, large patient cohorts have been systematically sequenced ("Genomic and Epigenomic Landscapes of Adult De Novo Acute Myeloid Leukemia" 2013; Haferlach et al. 2014; Papaemmanuil et al. 2016; Bernard et al. 2022).

The discovery of these mutations has informed clinical decision support of hematologic malignancies in three ways: 1) understanding correlations between gene mutations and clinical phenotype and outcomes informs diagnostic classification(Khoury et al. 2022; Arber et al. 2022); 2) the definition of prognostic biomarkers enables risk-adapted clinical management and treatment decisions(Tazi et al. 2022; Bernard et al. 2022); 3) targeted therapeutics have been developed (Stein et al. 2017; DiNardo et al. 2018; de Botton et al. 2021; Stone et al. 2017; Perl et al. 2019; Issa et al. 2022; Erba et al. 2022). However, the

full potential of molecularly guided medicine is challenged by the fact that most patients have multiple mutations in driver genes and the frequency of most mutated genes in cancer is rare. Patients with AML have a median of 3 mutations in one of ~100 myeloid driver genes. The systematic sequencing of large and well clinically annotated patient cohorts has enabled definition of specific patterns co-mutation and order of mutation acquisition (Papaemmanuil et al. 2016). For example, mutations in epigenetic modifiers such as *IDH1/2* or *DNMT3A* are frequently observed as early or clonal mutations, followed by the acquisition of secondary mutations in chromatin or spliceosome genes, and conversely, mutations in the receptor tyrosine kinase (RTK)-*RAS* signaling pathway genes (such as *NRAS*), are frequently observed as subclonal, late events. These patterns of co-mutation point to potential biological synergies that are selected for in disease initiation and progression and nominate these interactions for further characterization within primary patient samples.

Single cell technologies enable understanding of the biological consequences of somatic mutations in primary patient samples

Bulk genomic profiling has provided putative patterns of mutation acquisition based on large cohorts and single cell genomic profiling has validated these recurrent clonal architectures within individual patient samples (Miles et al. 2020; Morita et al. 2020).

However, it has historically been challenging to study the biological consequences of these somatic mutations as they sit within individual patients' complex clonal architecture. Conventional tools for isolating cell populations, such as flow cytometry, are not able to isolate cells from primary patient samples on the basis of specific somatic mutations. Bulk

phenotyping methods, such as RNA-sequencing, thus provide only an average signal across all cells within the sample, obscuring the contributions from particular phylogenetic subclones. Fortunately, the advent of genomic technologies described below has laid the foundation to enable the dissection of clone-specific signals from primary patient samples.

Single cell gene expression profiling studies of normal hematopoiesis (Velten et al. 2017; Karamitros et al. 2018; Popescu et al. 2019; Pellin et al. 2019) and AML (van Galen et al. 2019) have delivered a detailed map of how transcriptional regulators define the cellular hierarchy in normal hematopoiesis and how this is dysregulated in AML. These studies serve as a foundation and reference for studies of specific AML genetic subtypes.

Furthermore, technologies aimed at simultaneous detection of copy number states and somatic mutations with single cell phenotyping are emerging. This includes platforms which enable single cell targeted DNA profiling and surface protein expression (Miles et al. 2020) and single cell mutation genotyping and gene expression(van Galen et al. 2019; Nam et al. 2019; Petti et al. 2019; Velten et al. 2021). Additionally, to further resolve tumor clonal architecture, single cell copy number states can be inferred from single cell gene expression (Gao et al. 2022; Tickle et al., n.d.; Fan et al. 2018).

These resources, including the learnings of genetic co-mutation patterns from population genomics studies, single cell maps of normal and AML hematopoiesis, and new methods which enable linking single cell genotypes and phenotypes, now enable the investigation of clonal architecture and the role of somatic mutations in driving clone-specific phenotypes in primary patient samples.

This thesis addresses the challenges posed by genetic heterogeneity in the context of myeloid neoplasms. In Chapter 2, patients with Myelodysplastic Syndrome were systematically screened for mutations in UBA1, which has recently been linked to a severe autoinflammatory disease associated with high prevalence of hematologic malignancies including MDS. In Chapter 3, integrative single cell profiling was employed to characterize the phenotype of a high-risk AML subtype defined by mutations in IDH1/2 and their frequent co-mutation partners.

Chapter 2: Prevalence, co-mutation patterns, and clinical presentation of *UBA1* mutations in Myelodysplastic Syndromes

Abstract

Mutations in *UBA1*, which are linked to VEXAS syndrome, have been reported in patients diagnosed with Myelodysplastic Syndromes (MDS). However, the prevalence and clinical implications of *UBA1* mutations in MDS are unknown. We analyzed the prevalence and characteristics of *UBA1* mutations in a representative cohort of patients with MDS. First, we used ddPCR to screen a selected cohort of 375 male patients lacking disease-defining mutation or established disease classification by WHO guidelines and identified 28 patients (7%) with *UBA1* p.M41T/V/L mutations. To understand the prevalence in MDS more broadly and enable detection of mutations beyond the p.Met41 hotspot, we used targeted next-generation sequencing to profile 2027 representative patients and identified an additional 28 variants in 27 patients, which we classified as likely/pathogenic (n = 12) and unknown significance (n=15). We focused on the 40 patients with likely/pathogenic variants, representing 2% of our cohort. All patients were male and most (63%) were classified by WHO 2016 as MDS-MLD/SLD (25/40). The most frequent co-occurring mutations were in *TET2* (n=12), *DNMT3A* (n=10), *ASXL1* (n=3), and *SF3B1* (n=3). 5 patients had loss of ChrY. We performed a retrospective clinical review for inflammation or inflammatory-rheumatological symptoms in 24/40 patients with pathogenic variants and found that 14/24 cases had inflammation-associated clinical presentation including inflammatory-rheumatological disease diagnosis (n=9), inflammatory symptoms (n=11), chondritis (n=6), thrombo-embolic disease (n=5), non-infection fever (n=5), and vacuoles

(n=4). Within the subset of patients with no MDS driver mutations, *UBA1*-mutated patients had worse overall survival. Our study highlights that *UBA1* mutations are enriched in, but not limited to, the subset of MDS cases without established MDS driver mutations. The adverse risk and distinct clinical presentation of *UBA1*-mutant MDS argues for systematic screening, particularly in patients with inflammatory disease presentation.

Introduction

Patients with Myelodysplastic Syndromes (MDS) have heterogeneous clinical presentation, variable response to therapy and differential outcomes. In recent years, the consideration of genetic alterations alongside established clinical and morphologic features in guidelines for disease classification and risk stratification have enabled treatment decisions tailored to each patient's tumor and led to improved patient outcomes. However ~6-10% of MDS patients have no established disease defining mutations (Haferlach et al. 2014; Bernard et al. 2022). Furthermore, approximately 10-30% of MDS patients present with inflammatory symptoms, imposing further challenges in diagnosis and clinical management, as they frequently associate with high-risk disease features and comorbidities (Braun and Fenaux 2013; Zhao, Boy, et al. 2021). The etiology of this mixed presentation is not well understood and we lack clinical management plans tailored to the underlying pathophysiology (de Hollanda et al. 2011; Mekinian et al. 2016; Montoro et al. 2018; Oganesyan et al. 2021).

One such overlap between MDS and inflammatory clinical presentation is VEXAS (vacuoles, E1 enzyme, X-linked, autoinflammatory, somatic) syndrome, a systemic autoinflammatory disease(Beck et al. 2020). VEXAS is linked to mutations in the *UBA1*

gene, which encodes the major E1-activating enzyme required for ubiquitylation. While high variant allele fraction mutations at the p.Met41 hotspot were the first to be linked to VEXAS, additional variants have since been characterized as pathogenic(Bourbon et al. 2021; Poulter et al. 2021; Collins et al. 2022; Sakuma et al. 2023; Beck et al. 2023; Stiburkova et al. 2023). *UBA1* mutations have been reported in patients that harbor common mutations associated with myeloid neoplasia including *DNMT3A*, *TET2*, and *SF3B1* in the context of VEXAS(Beck et al. 2020), clonal hematopoiesis(Fernandez et al. 2022; Patel et al. 2022), and MDS(Zhao, Schell, et al. 2021; Georgin-Lavialle et al. 2022). The incidence of MDS in VEXAS patients is high (25-55%) (Beck et al. 2020; Bourbon et al. 2021; Georgin-Lavialle et al. 2022), however, the incidence of *UBA1* mutations and VEXAS syndrome have not been systematically screened in a representative MDS population.

In this study, we deliver a comprehensive landscape of *UBA1* mutations in MDS. First, we define the prevalence of *UBA1* mutations in a representative cohort of clinically and molecularly annotated MDS patients and then characterize the clinical and molecular presentation of *UBA1*-mutant MDS. This informs diagnostic classification guidelines for patients with VEXAS, MDS, or overlap syndromes.

Methods

Cohort Selection

Molecular and clinical data were reviewed from a representative diagnostic and treatment-naive MDS cohort (n= 3,323) ascertained through the International Working Group for

MDS (Bernard et al. 2022) **(Figure 2-1A)**. We prioritized 375 cases (Cohort A) for genotyping of *UBA1* hotspot p.Met41 by digital droplet PCR based on: 1) male gender; and 2) no identified driver mutations or few mutations, such as DTA (*DNMT3A*, *TET2*, *ASXL1*) genes; or 3) MDS unclassifiable (eg. MDS-U) or 4) subsets of MDS/MPN overlap syndromes (MDS/MPN-U, MDS/MPN-RS-T, aCML) per 2016 WHO Classification **(Table 2-1)**.

For the expanded cohort (Cohort B), targeted DNA sequencing of the entire UBA1 locus was performed on 2027 representative MDS samples **(Figure 2-1 and Table 2-2)**.

Digital droplet PCR for UBA1 p.Met41 mutations

Three ddPCR assays were designed for the *UBA1* p.M41 locus: p.Met41Thr (c.122T→C), p.Met41Val (c.121A→G), p.Met41Leu (c.121A→C). 375 samples were profiled by each assay. A droplet cutoff of 3 positive droplets was used to identify *UBA1* p.M41 mutant samples. As no *UBA1*-mutant patient sample controls were available, synthetic DNA fragments (IDT gBlocks) with the mutant allele sequence were used for assay validation and as positive controls.

Targeted DNA sequencing of *UBA1*

2027 MDS samples were sequenced with a targeted panel for *UBA1* and uniformly analyzed. Capture based NGS sequencing in paired-end 100bp configuration was performed for all *UBA1* exons from all known transcripts, including splice sites, to 100X coverage. Sample quality control and bioinformatics processing was performed as previously described (Bernard et al. 2022, 2020). In brief, alignment was performed by BWA, quality control by fastQC, Picard, and MultiQC, variants were called using

Mutect(Cibulskis et al. 2013) (SNV and indel), Strelka(Saunders et al. 2012) (SNV and Indel), CaVEMan(Jones et al. 2016) (SNV) and Pindel(Ye et al. 2009) (Indel). Variants were annotated for protein change and using VAGrENT and Ensembl-VEP, with the most deleterious protein change across all transcripts retained. Variants were filtered by the following criteria: on target, VAF>2%, passed by Mutect and Strelka or called by CaVEMan with fewer than 2 flags. Common germline polymorphisms were removed by referencing the gnomAD database. Variants passing these filters were annotated as "likely pathogenic" if previously reported and validated in the literature(Beck et al. 2020; Poulter et al. 2021; Collins et al. 2022; Bourbon et al. 2021) and "variant of unknown significance" (VUS) otherwise. For patients profiled by both ddPCR and NGS assays, all known ddPCR variants were recovered in the NGS variants, including one <2% VAF variant which was manually recovered from filtered variants.

Statistical Analysis

Distributions of continuous variables between two subgroups were compared using the Wilcoxon rank sum test. For overall survival analysis, time-to-event data was measured from time of sample collection to the time of the death from any cause or last follow-up using the survminer R package version 0.4.9 and survival R package version 3.3-1. Overall survival probabilities were estimated using the Kaplan-Meier method and comparisons between subgroups were conducted using the log-rank test. All analyses were conducted using R version 4.0.5.

Retrospective review of inflammatory symptoms

A retrospective chart review was conducted for 44/54 identified *UBA1*-mutant MDS patients across 14 centers who contributed the samples. To standardize reporting of

clinical features by these independent abstractors, a questionnaire (115 questions) was circulated to tabulate inflammatory rheumatic disease (IRD) diagnoses, inflammatory symptoms, treatment for IRD, hematologic diseases, bone marrow counts at time of MDS and/or IRD, and update on survival or transformation to AML at last follow up. Responses to the questionnaire were uniformly curated and patients were classified as having either 1) No inflammatory or IRD presentation or 2) low-grade presentation or 3) high-grade presentation. Patients were graded as "low-grade" if they did not have a formal diagnosis of IRD or multi-organ inflammatory presentation. Patients whose sole potentially inflammation-associated clinical presentation was non-specific, such as fever of unknown origin or weight loss, were classified as low-grade. For high grade, patients had a diagnosis of IRD and/or presented with inflammation localized to multiple organ systems, such as chondritis and lung or cartilage inflammation, or arthralgia.

Results

UBA1 p.M41T/V/L mutations in unclassifiable MDS or MDS without disease-defining mutations

First, we report the results of *UBA1* ddPCR profiling of Cohort A **(Table 2-1).** We identified 30 *UBA1* p.M41T/V/L mutations (median number of mutant droplets: 1301; range: 4-12166) in 7% (n=28 of 375 patients) of the cohort profiled with ddPCR. This included 15 patients with p.M41T (c.122T>C), 7 with p.M41V (c.121A>G), 4 with p.M41L (c.121A>C), and 2 patients with more than one mutation (1 with p.M41V/T, and 1 with p.M41L/V) **(Figure 2-1)**.

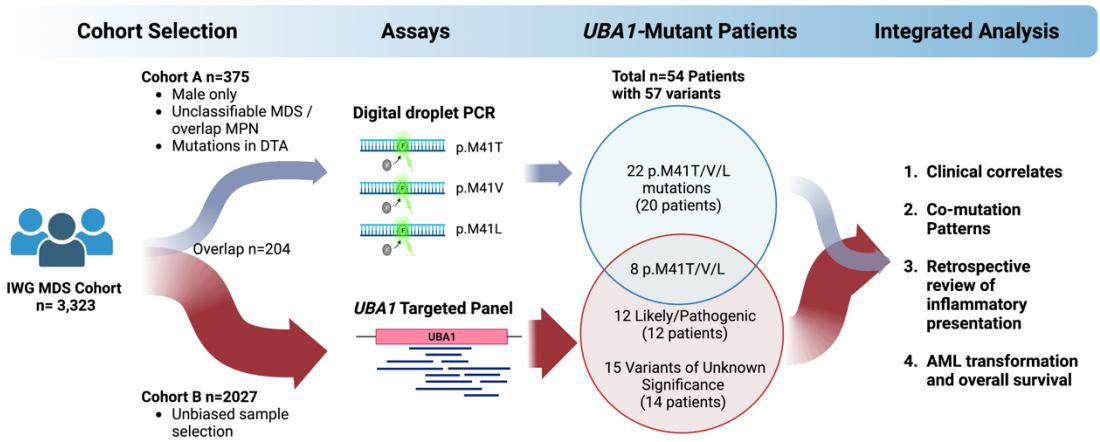


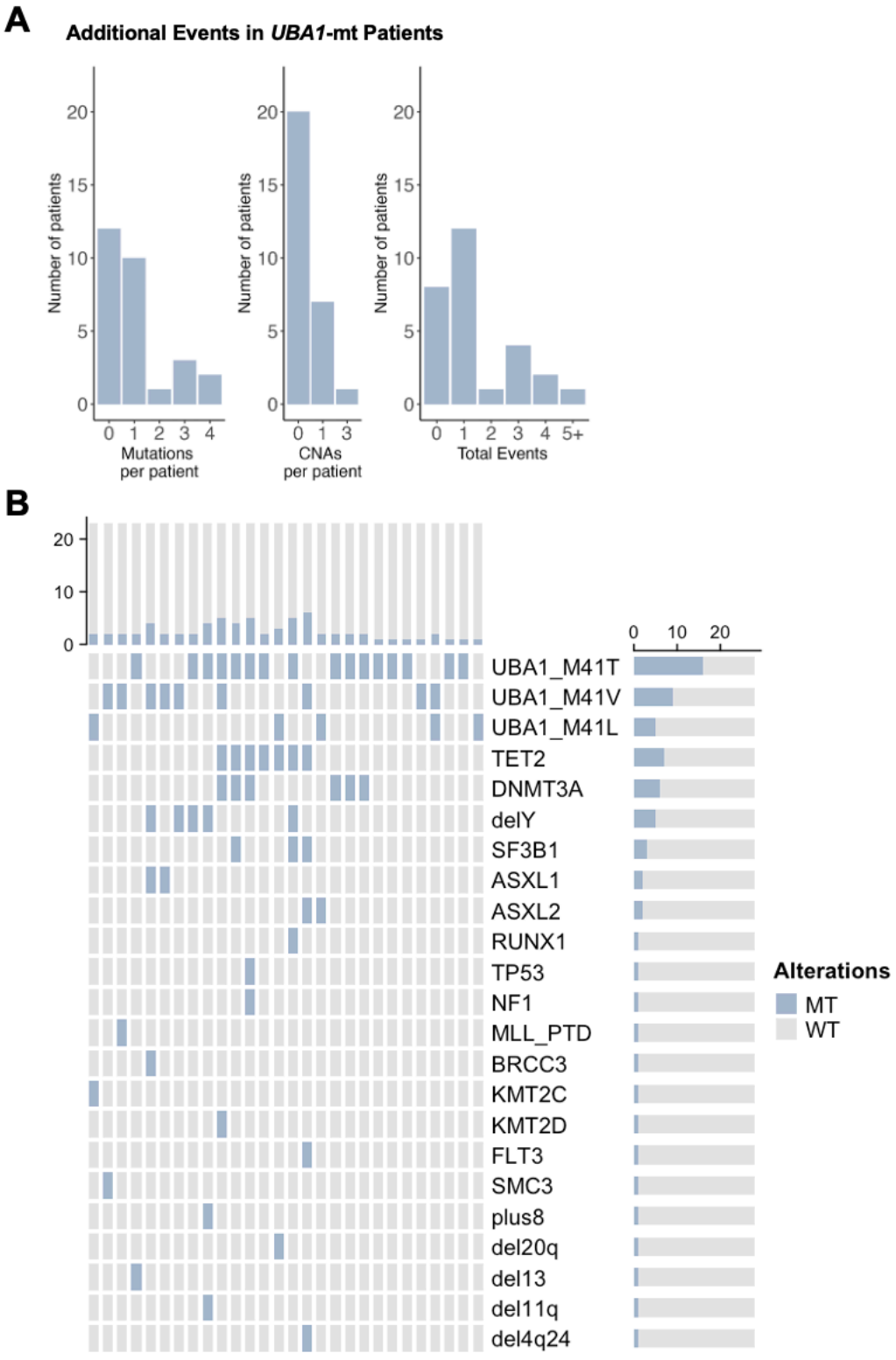
Figure 2-1 Schematic of study design showing Cohort A (top) profiled by ddPCR and Cohort B (bottom) profiled by NGS for UBA1 mutations.

Table 2-1 Cohort A Characteristics for UBA1 p.M41T/V/L -mutant and wildtype patients

Characteristic	0, N = 347[1]	1, N = 28[1]
AGE_AT_SAMPLE_TIME	71 (62, 77)	72 (65, 77)
Unknown	7	0
WHO_2016_SIMPLIFY_2		
aCML	22 (6.6%)	1 (3.7%)
CMML	20 (6.0%)	1 (3.7%)
MDS-EB1/2	47 (14%)	2 (7.4%)
MDS-RS-SLD/MLD	7 (2.1%)	0 (0%)
MDS-SLD/MLD	161 (48%)	18 (67%)
MDS-U	40 (12%)	2 (7.4%)
MDS/MPN-RS-T	13 (3.9%)	1 (3.7%)
MDS/MPN-U	20 (6.0%)	1 (3.7%)
NA	5 (1.5%)	0 (0%)
other	0 (0%)	1 (3.7%)
Unknown	12	1
BM_BLAST	2.0 (1.0, 4.0)	2.0 (1.0, 4.0)
Unknown	15	0
WBC	4 (3, 7)	4 (3, 7)
Unknown	36	2
ANC	2.2 (1.0, 4.0)	2.9 (1.8, 4.3)
Unknown	22	1
MONOCYTES	0.32 (0.12, 0.63)	0.30 (0.06, 0.50)
Unknown	63	5
HB	9.80 (8.70, 11.40)	10.90 (9.10, 11.35)
Unknown	14	1
PLT	108 (58, 201)	125 (84, 174)
Unknown	16	1

[1] Median (IQR); n (%)

In 71% (20 of 28) of *UBA1*-mutant cases, the karyotype was normal **(Figure 2-2A)**. Loss of the Y chromosome was observed in 5 cases. Three cases had MDS-associated cytogenetic abnormalities: del(11q), del(13) and focal del(20q) **(Figure 2-2B)**. The patient with del(11q) had two additional events: loss of the Y chromosome as well as gain(8).


A
Additional Events in UBA1-mt Patients
Number of patients
Mutations per patient
CNAs per patient
Total Events
B
UBA1_M41T
UBA1_M41V
UBA1_M41L
TET2
DNMT3A
delY
SF3B1
ASXL1
ASXL2
RUNX1
TP53
NF1
MLL_PTD
BRCC3
KMT2C
KMT2D
FLT3
SMC3
plus8
del20q
del13
del11q
del4q24
Alterations
MT
WT

(continued) i

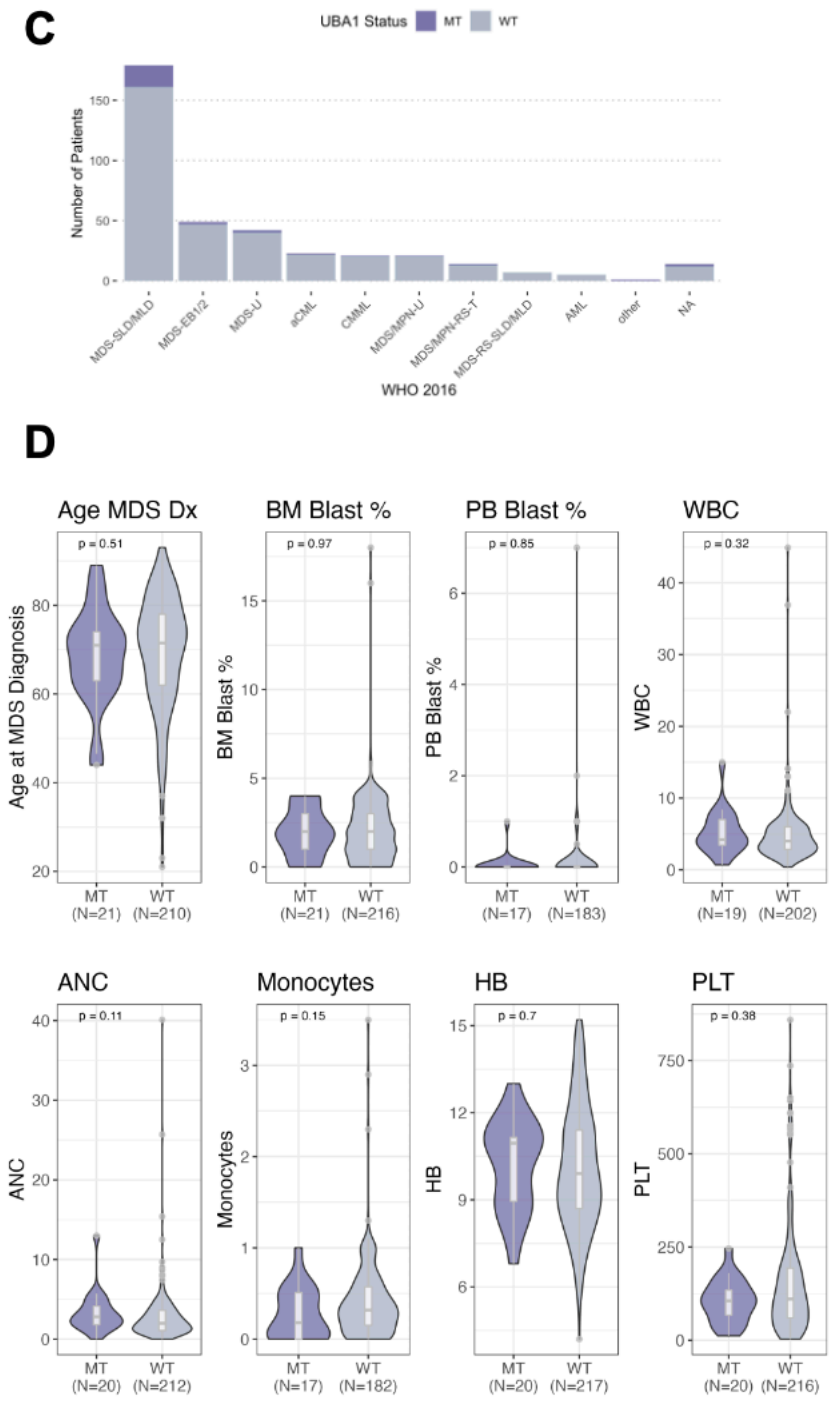


Figure 2-2 Co-mutation patterns and clinical characteristics of UBA1 p.M41T/V/L mutant MDS patients in Cohort A

A) Histogram of co-occurring mutations (SNVs and insertions/deletions), CNAs, and total events (mutations+CNAs) in *UBA1*-mutant patients
B) Oncoprint showing all co-occurring events in *UBA1* mutant patients
C) WHO 2016 classification for patients profiled in Cohort A colored by *UBA1* p.M41T/V/L mutation status.
D) Clinical characteristics of *UBA1* p.M41T/V/L mutant and wildtype patients in Cohort A.

In 43% of *UBA1*-mutant cases (12 of 28), *UBA1* mutations were the sole gene mutations**(Figure 2-2A)**. For a further 43% we observed co-mutation with DTA genes (*DNMT3A*, *TET2*, *ASXL1*) with median variant allele fraction (VAF) of 0.21 (0.02-0.56) **(Figure 2-2B)**. *SF3B1* co-mutations with VAF>0.3 were observed in 3 cases, which were classified per WHO 2016 as MDS-U, aCML, and MDS/MPN-RS-T.

67% (18/27) *UBA1*-mutant cases were classified per 2016 WHO as MDS-SLD/MLD, and two cases were MDS with excess blasts 1, and one case was CMML-1 with 9% blasts **(Figure 2-2C)**. Both MDS excess blasts cases had 5% bone marrow blasts and *UBA1* p.M41T mutations: at 2% VAF with del(11q) and gain(8) while the other had *UBA1* p.M41T at 78% and no other genomic events. The CMML case had *UBA1* p.M41T at 0.5 VAF and no other mutations in myeloid pathogenesis genes. The median age at diagnosis in cases with *UBA1* mutations was 72 years (range 44-89) and 71 years (range 21-93) for cases with no UBA1 p.M41 mutation **(Figure 2-2D)**. No significant difference was found when comparing clinical parameters (age, blast percentages, blood counts) between *UBA1*-mutant and -wildtype MDS **(Figure 2-2D)**.

These data demonstrate that mutations in *UBA1* found in isolation or in combination with other mutations in myeloid genes account for a significant proportion (7%) of patients diagnosed and treated as MDS but without established disease classification or disease-defining co-mutation patterns. This highlights the need to understand the clinical associations of UBA1 mutations in patients diagnosed as MDS in order to inform diagnostic classification guidelines.

UBA1 mutations in a representative MDS cohort

In order to 1) determine whether mutations in *UBA1* were indeed enriched in patients without disease classification, 2) detect recently described mutations beyond the p.M41 locus (such as p.S56F, p.A478S, p.S621C)(Bourbon et al. 2021; Poulter et al. 2021; Collins et al. 2022; Sakuma et al. 2023; Beck et al. 2023; Stiburkova et al. 2023), and 3) enable unbiased discovery of novel variants in *UBA1*, we deployed a targeted DNA sequencing strategy for the entire *UBA1* locus across an expanded cohort of 2,027 MDS samples (Cohort B) **(Figure 2-1)**. Cohort B characteristics are in **Table 2-2**.

Table 2-2 Characteristics of Cohort B (n=2,027).

Characteristic	N = 2,027[1]
AGE_AT_SAMPLE_TIME	71 (63, 78)
Unknown	6
WHO_2016_SIMPLIFY_2	
aCML	17 (0.8%)
CMML	271 (14%)
MDS-del5q	88 (4.4%)
MDS-EB1/2	505 (25%)
MDS-RS-SLD/MLD	328 (16%)
MDS-SLD/MLD	571 (28%)
MDS-U	42 (2.1%)
MDS/MPN-RS-T	31 (1.5%)
MDS/MPN-U	34 (1.7%)
NA	113 (5.6%)
other	5 (0.2%)
Unknown	22
SEX	
F	789 (39%)
M	1,233 (61%)
Unknown	5
BM_BLAST	3 (1, 8)
Unknown	32
WBC	4.3 (2.8, 6.9)
Unknown	199
ANC	2.07 (1.04, 3.74)
Unknown	110
MONOCYTES	0.40 (0.20, 0.80)
Unknown	425
HB	9.80 (8.60, 11.30)
Unknown	41
PLT	129 (66, 236)
Unknown	56

[1] Median (IQR); n (%)

By NGS in Cohort B, we identified 35 *UBA1* mutations in 34 (1.7%) patients, of which 20 (0.97%) had likely pathogenic variants and 14 had variants of unknown significance (VUS) **(Figure 2-3A, Table 2-3)**. The 20 likely pathogenic cases comprised 13 p.M41T/V/L (VAF range 0.013-0.94), as well as 7 patients with non-p.M41 mutations where 6/7 had VAF>0.35: 2 with p.S56F (VAF 0.87 and 0.93), 2 p.A478S (VAF 0.72 and 0.80), 2 p.S621C (VAF 0.80 and 0.82) and 1 p.Y55H (VAF 0.025) **(Figure 2-3B)**. Three of the VUS were truncating frameshift variants in alternative transcripts: p.K14fs*51 (VAF 0.93; ENST00000442035), p.S37fs*1 (VAF 0.502; ENST00000451702), and a complex event with p.K3N and K3fs*99 (VAF 0.615 and 0.583; ENST00000451702). 204 cases were profiled by both ddPCR and NGS and of those, 8/8 p.M41T/V/L mutations were identified by both assays. This included one mutation with 1% VAF that would likely have been missed by NGS profiling alone due to a VAF cutoff of 2%. We observed a very high correlation between VAF calculated from NGS or ddPCR (R^2=0.999 p<0.0001) **(Figure 2-3C)**. Despite profiling a large cross-section of MDS patients only 1% (20/2027) of MDS patients were positive for known pathogenic mutations in *UBA1*. This confirmed our original hypothesis that mutations in *UBA1* were most likely to be found in male MDS patients without disease-driving molecular alterations (prevalence 7%).

Table 2-3 Characteristics of n=54 UBA1-mutant MDS cases. Table is split by UBA1 mutation: p.M41, non-p.M41 pathogenic, and VUS

Characteristic	p.M41T/V/L, N = 33[1]	Other Pathogenic, N = 7[1]	VUS, N = 14[1]
AGE_AT_SAMPLE_TIME	73 (65, 79)	69 (64, 73)	72 (55, 77)
WHO_2016_SIMPLIFY_2			
aCML	1 (3.1%)	0 (0%)	0 (0%)
CMML	2 (6.2%)	0 (0%)	1 (7.1%)
MDS-EB1/2	4 (12%)	0 (0%)	5 (36%)
MDS-RS-SLD/MLD	0 (0%)	1 (14%)	1 (7.1%)
MDS-SLD/MLD	20 (62%)	5 (71%)	7 (50%)
MDS-U	2 (6.2%)	1 (14%)	0 (0%)
MDS/MPN-RS-T	1 (3.1%)	0 (0%)	0 (0%)
MDS/MPN-U	1 (3.1%)	0 (0%)	0 (0%)
other	1 (3.1%)	0 (0%)	0 (0%)
Unknown	1	0	0
IPSSM			
Very-Low	7 (23%)	2 (33%)	1 (9.1%)
Low	16 (52%)	2 (33%)	3 (27%)
Moderate-Low	4 (13%)	0 (0%)	1 (9.1%)
Moderate-High	2 (6.5%)	2 (33%)	2 (18%)
High	2 (6.5%)	0 (0%)	2 (18%)
Very-High	0 (0%)	0 (0%)	2 (18%)
Unknown	2	1	3
SEX			
F	0 (0%)	0 (0%)	3 (21%)
M	33 (100%)	7 (100%)	11 (79%)
BM_BLAST	2.00 (1.00, 4.00)	1.00 (1.00, 1.00)	4.00 (3.00, 8.00)
Unknown	0	0	1
WBC	4.30 (3.42, 7.61)	2.45 (2.04, 4.95)	3.80 (2.59, 5.23)
Unknown	2	1	2
ANC	3.00 (1.80, 4.66)	1.10 (0.94, 3.23)	1.58 (1.00, 2.84)
Unknown	1	0	2
MONOCYTES	0.31 (0.10, 0.52)	0.36 (0.20, 0.65)	0.10 (0.10, 0.44)
Unknown	5	2	3
HB	10.45 (8.75, 11.15)	9.90 (8.48, 11.60)	10.35 (8.17, 11.07)
Unknown	1	0	0
PLT	118 (86, 174)	57 (25, 65)	50 (31, 70)
Unknown	1	0	1

[1] Median (IQR); n (%)

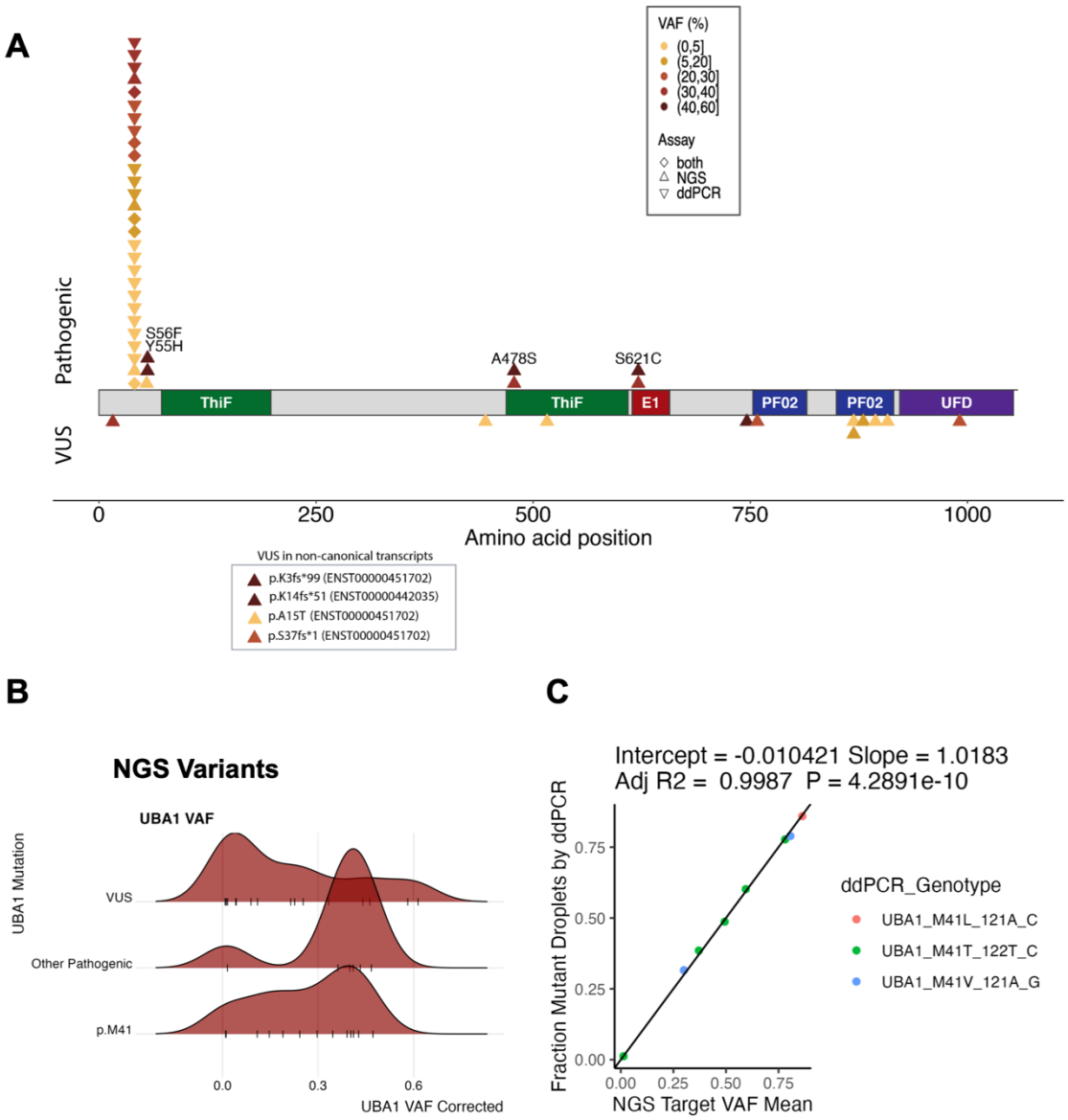


Figure 2-3 UBA1 mutations detected by ddPCR and NGS in Cohorts A and B

A) Lollipop plot showing likely-pathogenic mutations (top) and VUS (bottom) in *UBA1* detected by ddPCR and/or NGS. Variants are colored by VAF and point shape indicated whether the variant was detected by ddPCR, NGS, or both. VUS in the non-canonical transcripts are reported in the box below the axis.
B) VAF of *UBA1* mutations detected by NGS in Cohort B.
C) Correlation of *UBA1* NGS VAF and ddPCR mutant droplet fraction for patients (n=8) with *UBA1* p.M41T/V/L variants detected by both assays.

Among the 40 patients with pathogenic *UBA1* variants from either ddPCR and/or NGS (Cohort A and B), the WHO 2016 classification (available for n=38 of 40) was MDS-SLD/MLD (25), MDS-EB1 (4), MDS-U (3), CMML (2), MDS-RS-MLD (1), aCML (1), MDS/MPN-RS-T (1), MDS/MPN-U (1)**(Figure 2-4A and Table 2-3)**. The most frequent co-occurring events were in *DNMT3A* (n = 12), *TET2* (n=10), *ASXL1* (n=3), *SF3B1* (n=3) and loss of the Y chromosome (n=5) **(Figure 2-5)**.

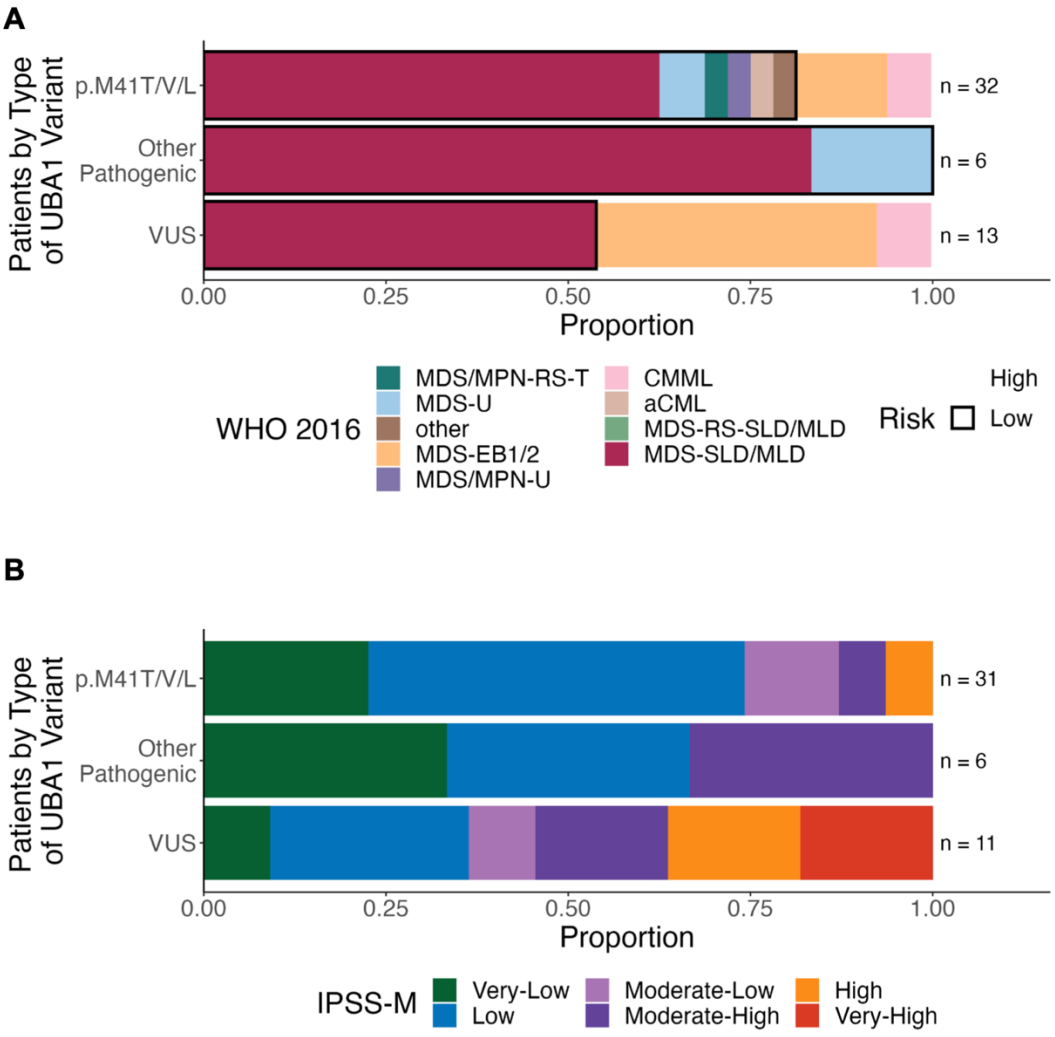


Figure 2-4 WHO 2016 classification and IPSS-M Risk Category of UBA1-mutant patients in total cohort (Cohort A+B)

A) Stacked barplot of WHO 2016 class for total cohort of *UBA1*-mutant patients (available for n=51/54).

B) Stacked barplot of IPSS-M Risk Category for total cohort of *UBA1*-mutant patients (available for n=48/54).

Next, we describe the 14 patients with VUS. 3/14 patients were female (variants: p.D16V, a splice variant c.2275-1G>A, and p.K3fs*99 in ENST00000451702). By WHO 2016 classification, MDS-SLD/MLD (n=6), MDS-EB1/2 (n=4), MDS-RS-MLD (n=1), and CMML-1 (n=1) **(Figure 2-4A)**. These patients had a higher number of mutations in established myeloid driver genes, with 50% (7 of 14) having 3 or more co-mutations, than the cases with pathogenic *UBA1* mutations (18%; 7 of 40) and were more likely to harbor alterations associated with adverse-prognosis such as TP53 (n=2), -7/7q (n=3), *ASXL1* (n=4), *BCOR* (n=2), *EZH2* (n=2), *RUNX1* (n=2), and *U2AF1* (n=1) (Bernard et al. 2022) **(Figure 2-5)**.

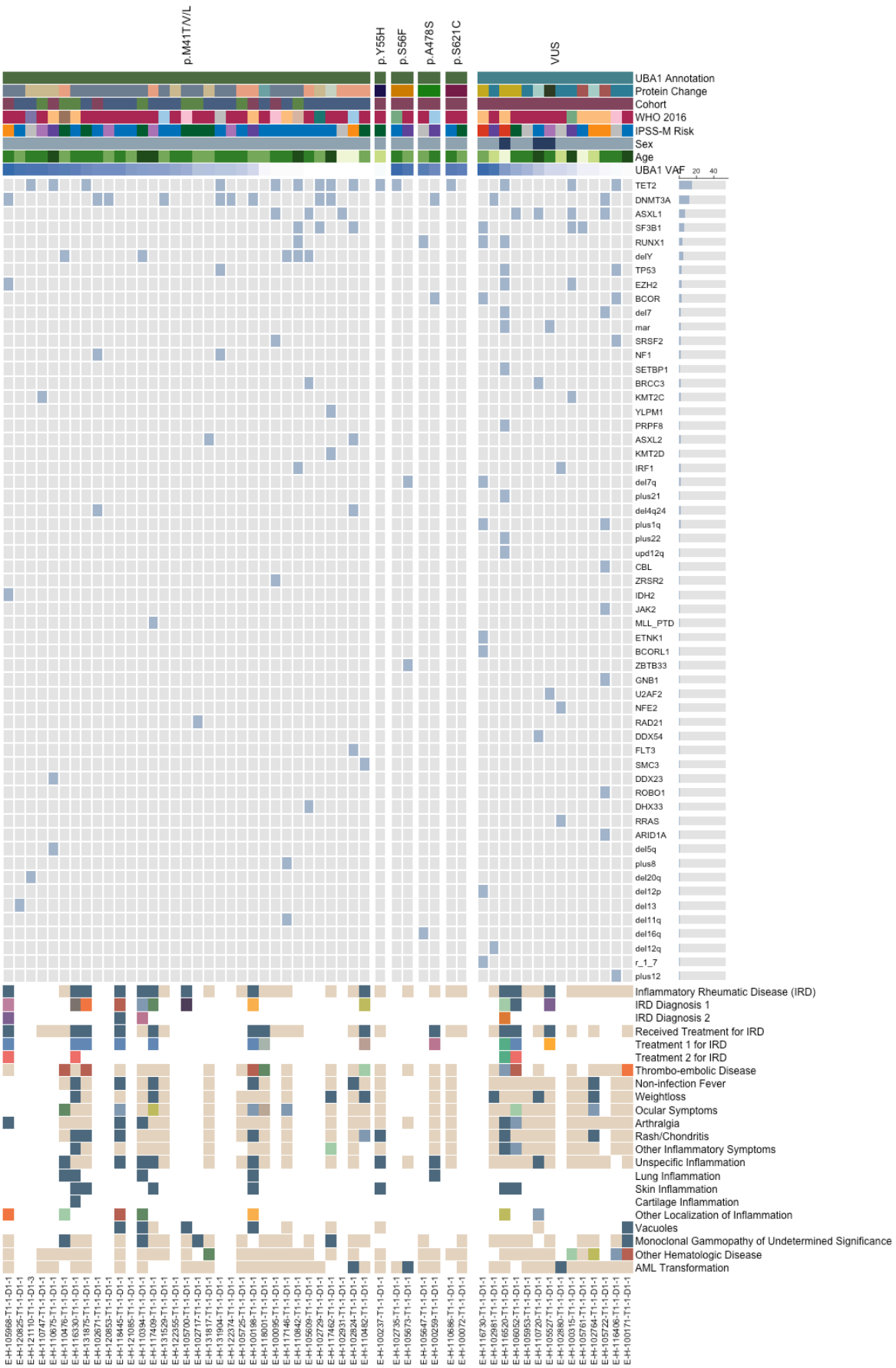
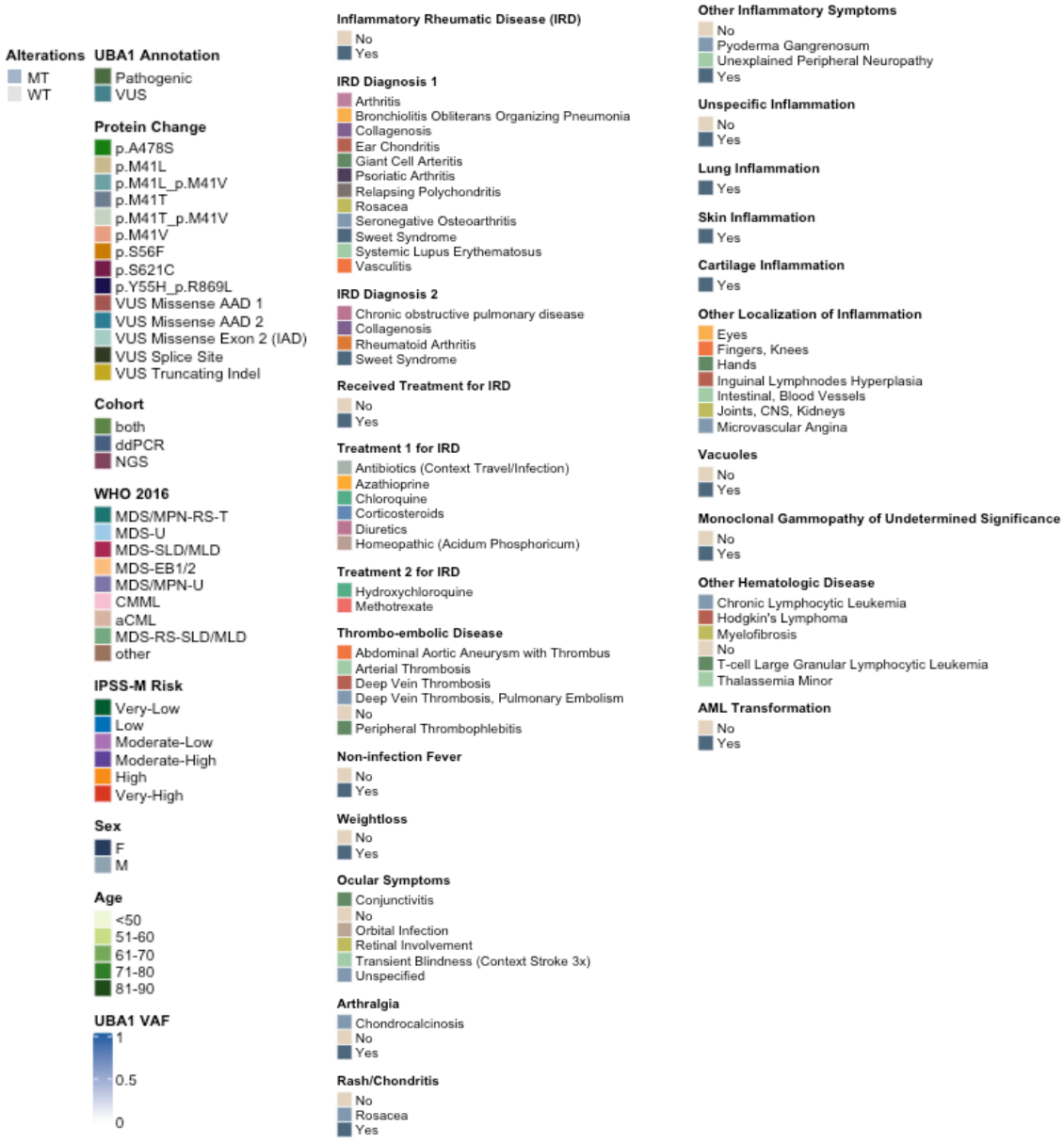

p.M41T/V/L
p.Y55H
p.S56F
p.A478S
p.S621C
VUS
UBA1 Annotation
Protein Change
Cohort
WHO 2016
IPSS-M Risk
Sex
Age
UBA1 VAF
20
40
TET2
DNMT3A
ASXL1
SF3B1
RUNX1
delY
TP53
EZH2
BCOR
del7
mar
SRSF2
NF1
SETBP1
BRCC3
KMT2C
YLPM1
PRPF8
ASXL2
KMT2D
IRF1
del7q
plus21
del4q24
plus1q
plus22
upd12q
CBL
ZRSR2
IDH2
JAK2
MLL_PTD
ETNK1
BCORL1
ZBTB33
GNB1
U2AF2
NFE2
RAD21
DDX54
FLT3
SMC3
DDX23
ROBO1
DHX33
RRAS
ARID1A
del5q
plus8
del20q
del12p
del13
del11q
del16q
del12q
r_1_7
plus12
Inflammatory Rheumatic Disease (IRD)
IRD Diagnosis 1
IRD Diagnosis 2
Received Treatment for IRD
Treatment 1 for IRD
Treatment 2 for IRD
Thrombo-embolic Disease
Non-infection Fever
Weightloss
Ocular Symptoms
Arthralgia
Rash/Chondritis
Other Inflammatory Symptoms
Unspecific Inflammation
Lung Inflammation
Skin Inflammation
Cartilage Inflammation
Other Localization of Inflammation
Vacuoles
Monoclonal Gammopathy of Undetermined Significance
Other Hematologic Disease
AML Transformation

Figure 2-5 Oncoprint for total cohort of UBA1-mutant patients (n=54) including results of retrospective review of clinical history for inflammatory features (bottom). Patients are ordered by decreasing UBA1 VAF per group (p.M41, non-p.M41 pathogenic, VUS).

(continued) ii



As this study leveraged the cohort from the IPSS-M study (Bernard et al. 2022), the IPSS-M risk category was readily available for 48 patients **(Figure 2-4B)**. Among *UBA1*-mutant patients with pathogenic variants, the majority were IPSS-M Very-Low/Low Risk (73% 27/37). In contrast, patients with VUS were more likely to be Moderate-Low or higher risk (64% 7/11) compared to those with pathogenic variants (27%).

To understand the clonal relationship between *UBA1* mutations and other frequently mutated genes, we compared VAFs. In 8 patients with pathogenic *UBA1* > 2% VAF and co-occurring DNMT3A mutations, *DNMT3A* and *UBA1* VAF exhibited strong correlation (slope = 0.88, r = 0.87, p = 0.0005), suggesting that these mutations occur in the same cells and co-mutation leads to clonal expansion **(Figure 2-6A)**. Conversely, co-mutations in *RUNX1* tended to be subclonal to *UBA1* (3/4), and *TET2* co-mutation were either subclonal (n=4) or clonal (n=8) to *UBA1* **(Figure 2-6B)**.

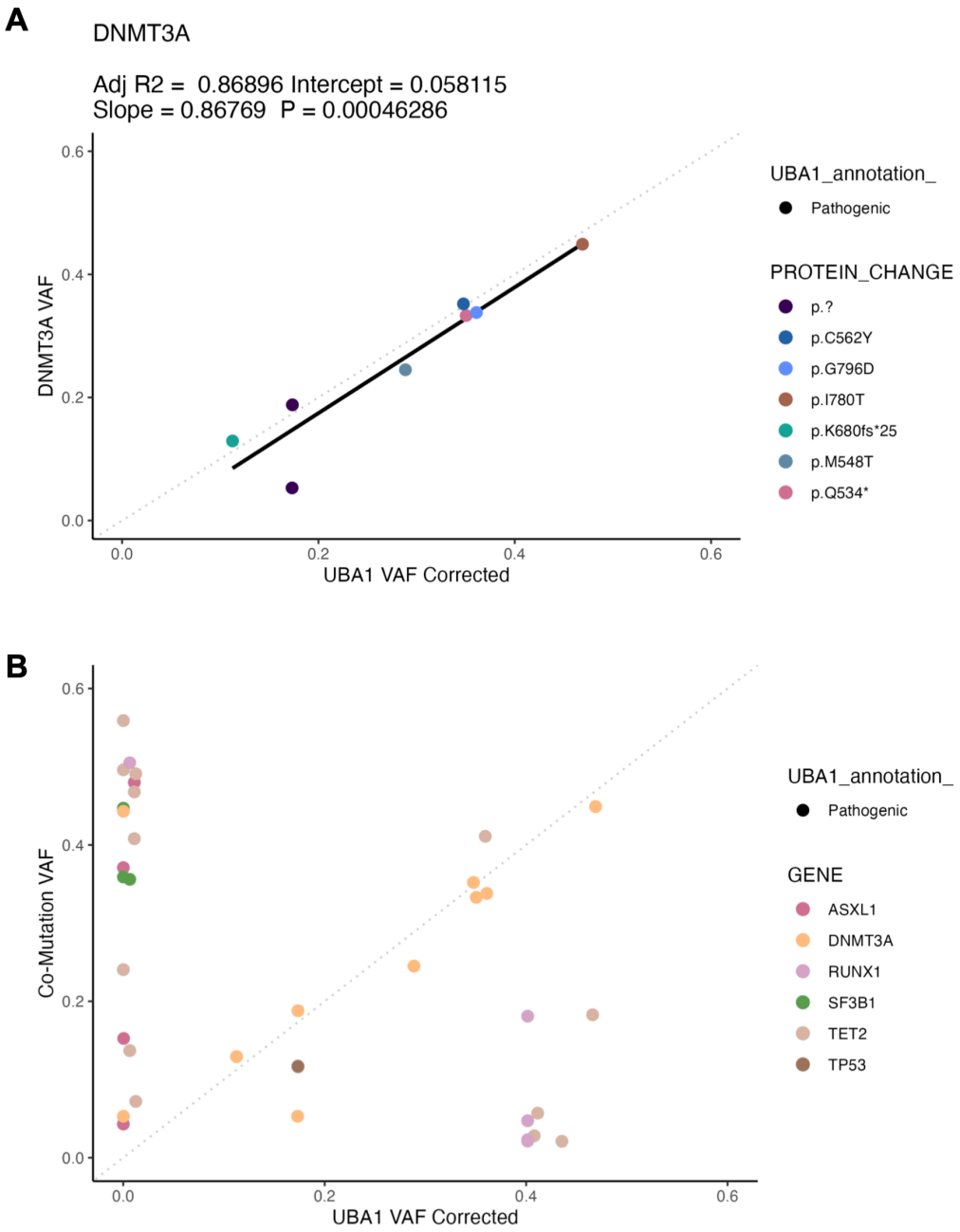


Figure 2-6 Analysis of VAF of UBA1 and co-occurring mutations

A) VAF of pathogenic *UBA1* and co-occurring *DNMT3A* mutations in 8 patients. Grey dashed line represents *DNMT3A* VAF = *UBA1* VAF. Black solid line represents a linear model fit to the data.
B) VAF of pathogenic *UBA1* and all additional co-occurring mutations. Gray dashed line represents co-mutation VAF = *UBA1* VAF.

Clinical presentation of inflammatory and/or VEXAS-associated symptoms in *UBA1*-mutant MDS

A retrospective review of IRD and inflammatory symptoms was conducted for 44/54 (81%) cases **(Figure 2-5)**. First, we describe the clinical presentation of the total 24/33 cases with pathogenic *UBA1* p.M41T/V/L variants with available data. 9/18 (50%) with available data had IRD including relapsing polychondritis, sweet syndrome, vasculitis, bronchiolitis obliterans with organizing pneumonia, ear chondritis, and rosacea. 3 patients had more than one IRD. 8/8 received treatment for IRD including steroids (n=6), methotrexate (n=2). Four of 19 patients had MGUS, 4 had bone marrow myeloid/erythroid vacuoles. We acknowledge that we are likely under-estimating the true incidence of vacuoles since this feature is not always reported on pathology reports and it was not possible to return to original images in all cases. Other symptoms included thromboembolic disease (5/16), non-infection fever (5/14), weight loss (4/14), ocular symptoms (5/15), arthralgia/joint stiffness (3/16), rash/chondritis (5/15), other inflammation (6/14) most frequently localized to skin (4) and lung (4). The time from IRD diagnosis to MDS diagnosis was on average 102 days (range 16.5 years prior to MDS diagnosis to 1.4 years after).

Within the cases with likely-pathogenic mutations outside of the p.M41T/V/L locus (n=7), none of the 5/7 of patients with available data had an IRD diagnosis, but two patients had inflammatory symptoms including i) rash in a patient with *UBA1* p.Y55H variant (VAF 0.025) and VUS variant p.R869L at 0.02 VAF with co-occuring mutations in *TET2* (p.Q138L 0.491 VAF and p.C1289Y at 0.072 VAF) and ii) lung inflammation in a patient with *UBA1* p.A478S (VAF 0.722) and *DNMT3A* and *BCOR* co-mutations.

Within the VUS cohort, 7/12 cases with available clinical history had IRD (n=3) or other symptoms (n=4) **(Figure 2-5)**. The IRD diagnoses included sweet syndrome, systemic lupus erythematosus, rheumatoid arthritis and collagenosis for which all three patients received treatment, including steroids, methotrexate, azathioprine and hydroxychloroquine. Four patients had skin involvement from nonspecific inflammation (n=2), rash/chondritis (n=2) and pyoderma gangrenosum (n=1). One additional patient with *UBA1* p.I894S VAF 0.022 but no other myeloid mutations had MGUS, vacuoles, aortic aneurysm with thrombus, and Hodgkin's Lymphoma. *UBA1* p.I894S and p.I894F have been separately reported in two symptomatic patients, but in those patients, the VAF was high (56% and 37% VAF, respectively) and both patients had additional *UBA1* mutations at p.N606I and p.Y55H, respectively (Sakuma et al. 2023). Together, these reports build evidence that link the *UBA1* p.I894 locus to inflammatory presentation, but further functional characterization of *UBA1* VUS will be necessary to determine if they are disease-causative.

Frameshift VUS in non-canonical *UBA1* transcripts in patients with inflammation and IRD

Interestingly, patient E-H-116520-T1-1-D1-1 with the complex truncating event *UBA1* p.K3N and K3fs*99 (VAF 0.615 and 0.583) in transcript ENST00000451702 was a female and had significant inflammatory presentation. The transcript ENST00000451702 is lowly expressed compared to the canonical ENST00000377351 **(Figure 2-7A)** in the normal tissue gene expression database (GTEx). This patient was diagnosed with Systemic Lupus Erythematosus and Rheumatoid Arthritis 16 and 13 years, respectively, prior to MDS diagnosis at age 41. She also had deep vein thrombosis, pulmonary embolism,

rash/chondritis, arthralgia, and additional inflammation localized to the kidney and central nervous system. She was refractory to many lines of therapy including chloroquine and rituximab, cyclosporine, cyclophosphamide, and mycophenolic acid. Furthermore, by both clinical karyotyping and review of copy number state by NGS revealed 3 copies of Chr X **(Figure 2-7B)**. While VEXAS has been reported in women with monosomy ChrX (Stubbins et al. 2022; Barba et al. 2021), this case represents duplication of the ChrX allele carrying the UBA1 mutation. Interestingly, she underwent allo-HSCT and experienced remission of both IRD and MDS. Her relapse of MDS was associated with concomitant relapse of IRD.

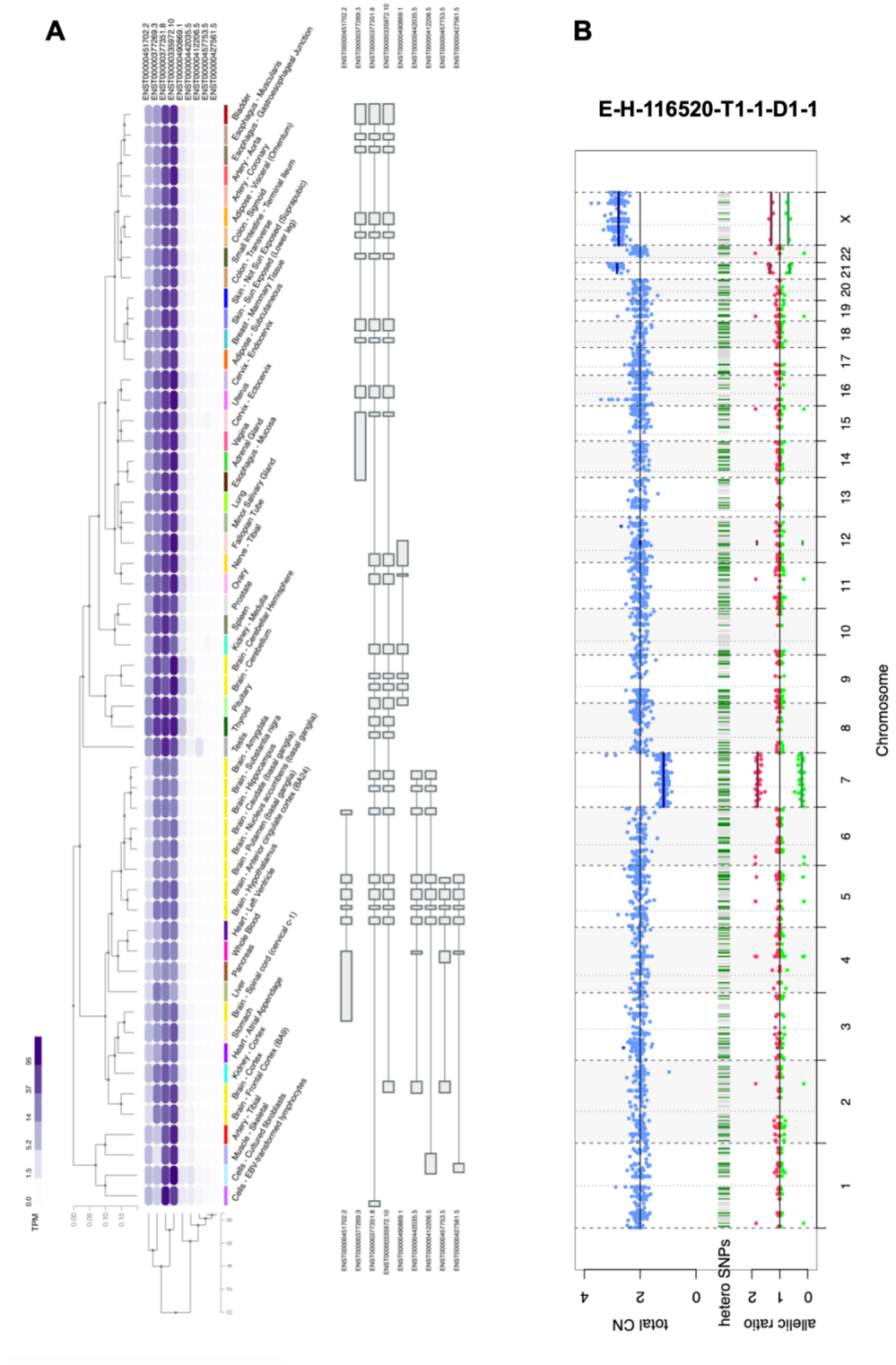


Figure 2-7 Review of female patient with UBA1 frameshift VUS

A) UBA1 isoform expression in the GTEx database.
B) Copy number state per chromosome inferred from NGS panel sequencing of female patient (H116520) with *UBA1* VUS in non-canonical transcript.

Another frameshift and truncating VUS, *UBA1* p.S37fs*1 (VAF 0.502) also in transcript ENST00000451702, was found in a male patient who had neutrophilic dermatosis/Sweet's Syndrome, diagnosed 4 months prior to MDS diagnosis, as well as skin inflammation and pyoderma gangrenosum, and chondrocalcinosis.

The clinical presentation of these patients nominates these *UBA1* VUS for further functional validation to determine their potential pathogenicity. *UBA1* isoform expression in MDS and VEXAS patients is under investigation.

Association of *UBA1* VAF with inflammatory clinical presentation

Because not all patients with *UBA1* p.M41T/V/L mutations had a clinical presentation suggestive of VEXAS syndrome, we next asked whether there was an association between *UBA1* VAF and presentation with inflammatory symptoms or IRD. 8/12 (67%) patients with *UBA1* p.M41 VAF >5% exhibited VEXAS-associated symptoms **(Figure 2-8A)**. Within the 7 patients with non-p.M41 pathogenic variants, 3 had no inflammatory features and two had low grade, limited to rash in one patient and lung inflammation in the other. Amongst the VUS, two patients, which were described in the previous section with high VAF frameshift truncating variants had high-grade, and 10 had None or Low grade.

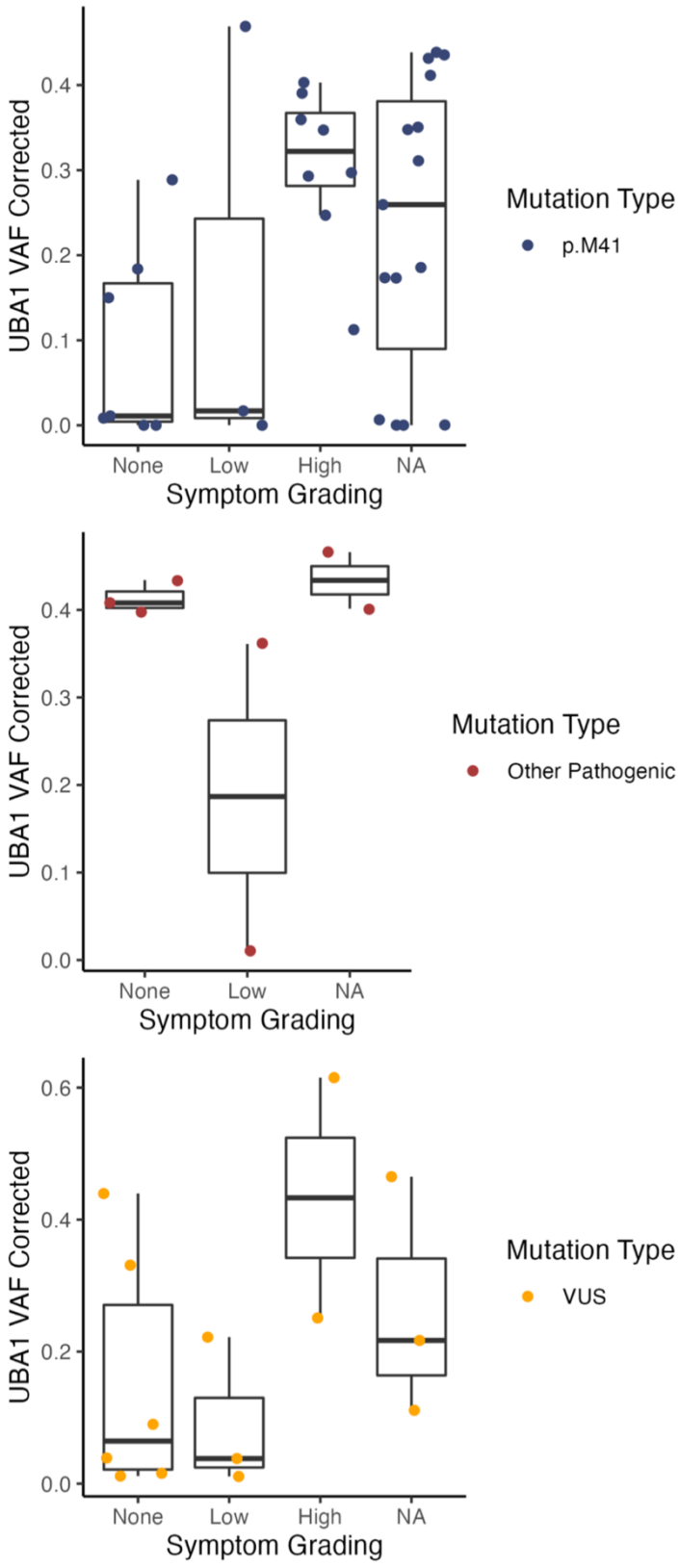


Figure 2-8 Association of UBA1 VAF with clinical presentation of inflammatory symptoms

Transformation to AML and overall survival

Despite growing numbers of identified VEXAS/MDS patients, to date only one UBA1-mutant MDS patient has been reported to have transformed to AML (Battipaglia G 2023). Thus, we next examined whether any *UBA1*-mutant patients within our cohort transformed to AML (median follow up 2.2 years; range: 0-13.7 years). Three patients in the total cohort transformed to AML, however, they did not have high VAF UBA1 p.Met41 mutations. One patient had *UBA1* p.M41V, but at a very small VAF (0.0002 by ddPCR) and co-mutation in *TET2*, *SF3B1*, *FLT3*, *ASXL2*, del4q24 **(Figure 2-5)**. The second had a *UBA1* p.S56F mutation (VAF 0.868) and a deletion of chromosome 7q. The third had VUS *UBA1* p.R869L (VAF 0.222) and co-mutations in *IRF1*, *NFE2* and *RRAS*. Among the other patients that died or were censored after 1 year (n=37), none transformed to AML.

Finally, we asked whether *UBA1* mutations conferred adverse risk in MDS. To address this, we focused on the ddPCR cohort as this is the MDS patient population most likely to have *UBA1* mutations and predominantly comprises MDS cases considered to be low risk (Bernard et al. 2022). There was no difference in overall survival between UBA1-mutant (n=27) and wildtype (n=312) patients (log-rank test p = 0.7) **(Figure 2-9A)**, however, when subsetting for patients with no driver mutations detected by targeted DNA sequencing (UBA1-mutant n=7; wildtype n=45), UBA1-mutant patients had worse overall survival (median 5.2 vs 7.3 years; log-rank test p = 0.03) **(Figure 2-9B)**.

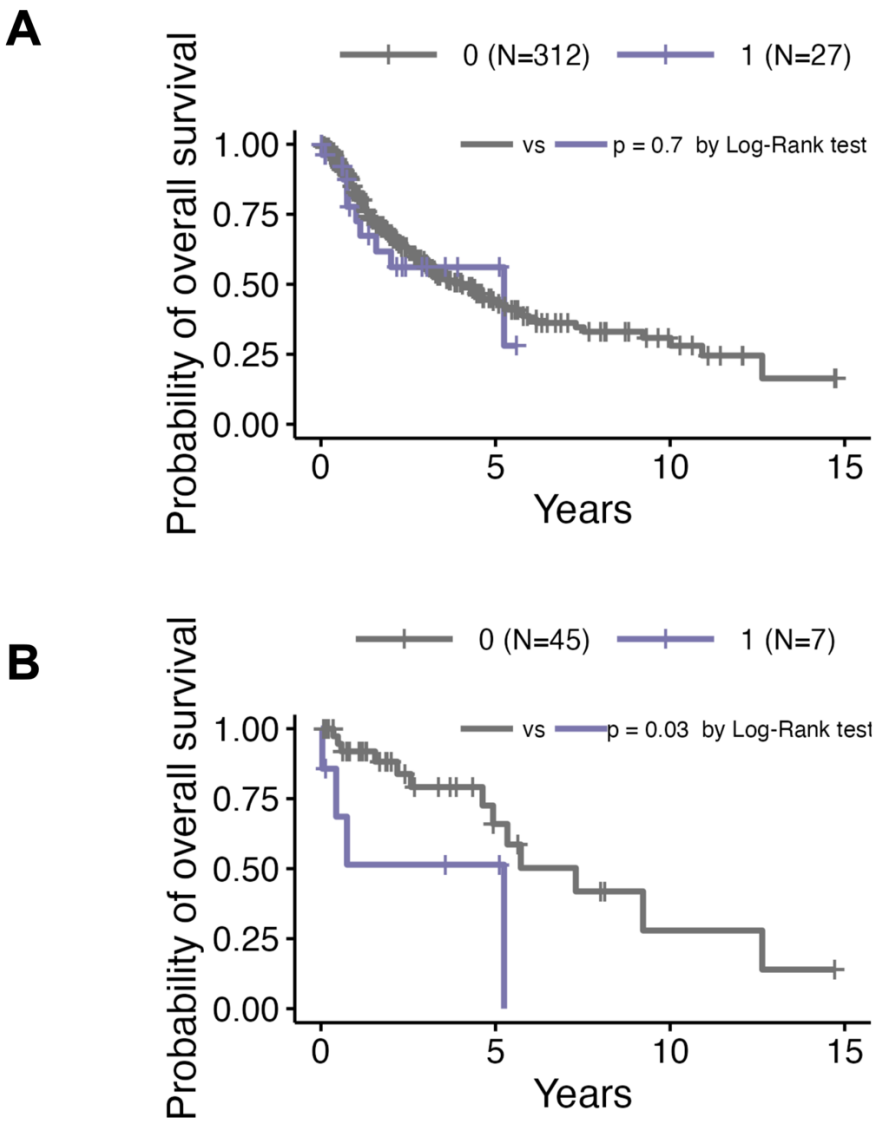


Figure 2-9 Survival analysis of UBA1 p.M41T/V/L mutant and wildtype MDS patients

A) Kaplan-Meier curve for overall survival in Cohort A (ddPCR) of MDS patients with or without mutations in *UBA1*. Number of patients per group and p-value are indicated on the plot.
B) Kaplan-Meier curve for overall survival in the subset of patients in Cohort A with no myeloid driver mutations, with or without mutations in *UBA1*. Number of patients per group and p-value are indicated on the plot.

Discussion

In this study, we deliver a landscape of *UBA1* alterations in the largest diagnostic and representative MDS cohort to date, unravel the critical genetic co-operators of *UBA1* mutations driving clonal evolution, and inform disease classification for MDS/VEXAS overlap syndromes. Findings from this study enable basic, translational and clinical research.

First, we inform clinical diagnostic assay development and implementation. We identify that the patients most likely to have *UBA1*-mutations are those without established MDS disease classification and few, if any, co-occurring mutations. Our analysis, together with previous studies, suggest that while p.M41 mutations are the most common *UBA1* alterations in MDS, other variants have been identified in patients with VEXAS-like clinical presentation (Poulter et al. 2021; Bourbon et al. 2021; Beck et al. 2023). As clinical manifestations of inflammation or IRD predominantly occur in patients with *UBA1* VAF >5%, it is not necessary to employ ultra-sensitive assays for variant detection. Future longitudinal studies tracking *UBA1* clonal evolution will be needed to assess the selective potential of low-VAF *UBA1*-mutated clones.

Next, we identify biological pathway dependencies through our analysis of patterns of co-mutation and order of mutation acquisition. The distinct patterns of clonal architecture in *DNMT3A*- or *TET2*-mutant cases suggest different mechanisms of cooperation in these frequently occurring pairwise genotypes. The interaction

between these mutations, which are also often found in clonal hematopoiesis and MDS, may help elucidate risk factors for developing MDS.

Through a retrospective review of *UBA1*-mutant patient clinical histories, we are able to integrate *UBA1* mutations and co-occurring MDS driver mutations with a complete picture of clinical manifestations of inflammatory symptoms or IRD. We find a diverse clinical presentation among patients that cannot be explained by *UBA1* mutation alone. Within larger cohorts of *UBA1*-mutant patients, it may be possible to fully understand how co-occurring mutations in MDS driver genes shape clinical presentation and outcomes. Finally, within the MDS patient population with no myeloid driver mutations, *UBA1* mutations confer adverse risk on survival. Thus, this data shows that patients with MDS but no clear canonical driver mutations should be screened for UBA1 mutations, particularly in the presence of inflammatory features. More broadly, systematic assessment of *UBA1* mutations in MDS patients will enable proper clinical management.

Chapter 3: High resolution mapping of *IDH1/2*-mutant Acute Myeloid Leukemia

Abstract

Mutations in Isocitrate dehydrogenase 1/2 (*IDH1/2*) are early initiating events in Acute Myeloid Leukemia (AML), and are associated with high-risk disease. The complex clonal architecture and cellular heterogeneity in IDH1/2-mutant AML underlies the heterogeneous clinical presentation and outcomes. Utilizing integrated single cell genotyping and gene expression, we demonstrate a stem-like and inflammatory phenotype of IDH1/2-mutant AML and identify transcriptomic programs associated with phylogenetic clones harboring mutations in *NPM1*, *NRAS*, and *SRSF2*. Furthermore, these clones had diverse responses to treatment with combination IDH1/2 inhibitors and chemotherapy including elimination, reconstitution of myeloid differentiation, or retention within progenitor populations. In patients that relapse after IDH1/2 inhibitor monotherapy, we identify upregulated stemness, inflammation, mitochondrial metabolism, and anti-apoptotic factors, as well as downregulated MHC Class II antigen presentation at relapse. Applying this framework to *IDH2*-mutant clonal hematopoiesis reveals an upregulation of *IDH2*-associated pathways, including inflammation, early at the pre-leukemic stage. In conclusion, we provide a detailed phenotyping of *IDH1/2*-mutant AML and a framework for dissecting contributions of recurrently mutated genes in AML at diagnosis and under therapy, with implications for precision medicine.

Introduction

Acute Myeloid Leukemia (AML) is an aggressive blood cancer with rapid onset, limited therapeutic options, and 25% 5-year survival (Kantarjian et al. 2021). AML is characterized by the stepwise acquisition of oncogenic mutations, leading to aberrant hematopoiesis and the accumulation of immature cells of the myeloid lineage. From population genomics studies, we understand the genomic landscape of AML is defined by specific and ordered patterns of co-mutation(Papaemmanuil et al. 2016). These patterns likely reflect functionally relevant interactions between frequent co-mutators which shape disease biology and clinical presentation.

The Isocitrate Dehydrogenase 1/2 (*IDH1/2*) genes are mutated in about 20% of AML cases and represent a high-risk disease subset that often predicts relapse (Papaemmanuil et al. 2016; Paschka et al. 2010). These mutations are often acquired early and thus considered initiating events in AML. This is further corroborated by studies in clonal hematopoiesis where the presence of *IDH1/2* mutations has been associated with high risk of AML transformation(Desai et al. 2018; Abelson et al. 2018; Young et al. 2019). Mutations in *IDH1* (p.R132) and *IDH2* (p.R140 and p.R172) confer neomorphic activity to IDH, leading to the increased production of the oncometabolite 2-hydroxyglutarate (R-2-HG). R-2-HG overproduction competitively inhibits alpha-ketoglutarate (α-KG) dependent enzymes (Ward et al. 2010; Clark, Yen, and Mellinghoff 2016), including the TET family of demethylases (Moran-Crusio et al. 2011) resulting in a globally hypermethylated phenotype and blocked differentiation (Figueroa et al. 2010; Xu et al. 2011; C. Lu et al. 2012; Lev M. Kats et al. 2014).

In AML, *IDH1/2* mutations are mutually exclusive with TET2(Figueroa et al. 2010), and are frequently found with co-mutations in *DNMT3A*, *NPM1*, and *SRSF2*. These specific pairwise and ordered interactions indicate functional redundancy (*TET2*) or cooperativity that likely underwrites the transition from pre-leukemic cells to overt myeloid malignancies. As *IDH1/2* mutations are early events in leukemogenesis and often persist at relapse, they are attractive drug targets(Ok et al. 2019; Shlush et al. 2014). Thus, first-in-class targeted therapies for *IDH1/2* mutant AML have been clinically investigated and/or FDA approved for relapsed/refractory disease (Stein et al. 2017; DiNardo et al. 2018; de Botton et al. 2021) and newly diagnosed AML (Pollyea et al. 2019; Roboz et al. 2020), as well as in rational combination with intensive chemotherapy (Stein et al. 2021) or hypomethylating agents (DiNardo et al. 2021; Montesinos et al. 2022). These studies along with preclinical models, suggest that the inhibitors reduce 2-HG production, and thus induce cellular differentiation of leukemic blasts, with normalization of the immature-to-mature cell population ratio (Amatangelo et al. 2017; Fang Wang et al. 2013; Yen et al. 2017). However, the clinical implementation of IDH1/2 targeted agents has revealed highly variable patient responses which have implicated resistance-associated mutations, such as in the spliceosome or Receptor Tyrosine Kinase (RTK)-*RAS* pathway (Amatangelo et al. 2017; DiNardo et al. 2018; Choe et al. 2020). This highlights the importance of better understanding how serially acquired mutations shape disease biology at diagnosis and ultimately influence treatment response. Despite detailed phenotyping and modeling studies of *IDH1/2*-mutant AML, the biological consequences of these mutations within the context of cooperating mutations in primary disease is not well understood. A major complication of bulk phenotyping studies is the clonal and genetic heterogeneity in primary disease.

The application of high throughput genomic profiling at single cell resolution to AML patient samples has defined the cellular hierarchy in AML (van Galen et al. 2019; Lasry et al. 2022). However, these studies have been conducted across genomically heterogeneous patient populations, making it difficult to extract insights regarding the role of specific driver mutations and their co-mutated genes in disease biology. The further development of methods which link single cell genotype (van Galen et al. 2019; Nam et al. 2019; Rodriguez-Meira et al. 2019; Velten et al. 2021; Petti et al. 2019) and allelic state (Tickle et al., n.d.; Fan et al. 2018; Gao et al. 2022) to single cell transcriptomic profiles and cell lineage provides the opportunity to design studies to understand the transcriptomic consequences of each mutation across leukemic cell lineages and clonal phylogenies.

In this study, we characterized how *IDH1/2* mutations and their most frequently occurring co-mutations dysregulated hematopoiesis in a genomically representative cohort of 15 *IDH1/2*-mutant AML patients. We leveraged newly developed single cell techniques to assign leukemic cells to phylogenetic subclones in order to delineate the role of serially acquired co-mutations in shaping cell type composition and transcriptional programs. Furthermore, we track genetic, clonal and transcriptional changes at three stages of the AML disease process: diagnosis, remission, and relapse under IDH1/2 inhibitor therapy, to identify transcriptional programs altered at response and relapse. Finally, we identified mIDH-related biological processes including inflammation and hypoxia pathways, that are already dysregulated at the pre-leukemic state of CH. Together, this work provides a high-resolution map of *IDH1/2*-mutant AML and is highly relevant for validating potential new targets or biomarkers.

Results

Genomic landscape of *IDH1/2*-mutant AML

IDH1/2 mutations represent early genomic events in AML, but are not sufficient for the development of overt leukemia (Sasaki, Knobbe, Munger, et al. 2012; Chen et al. 2013; Lev M. Kats et al. 2014). To identify recurrently co-mutated genes in *IDH1/2*-mutant AML, we performed a meta-analysis of 3,653 AML patients with comparable molecular profiling data (Tazi et al. 2022; Papaemmanuil et al. 2016). Of these, 284 (8%) and 604 (17%) patients had mutations in *IDH1* or *IDH2*, respectively. *IDH1/2* mutations are not typically observed in fusion driven AML or AML with TP53 mutations and or complex karyotype (Odds Ratio (OR) *IDH1* 0.32; *IDH2* p.R140 0.29)**(Figure 3-1A)**. We confirm the mutual exclusivity between mutations in *IDH1*, *IDH2* and *TET2* (OR *IDH1*:*IDH2* p.R140 0.24; *IDH1*:*IDH2* p.R172 0.13), (OR *IDH1* 0.2; *IDH2* p.R140 0.29; *IDH2* p.R172 0.05), due to their epistatic relationship (Figueroa et al. 2010).

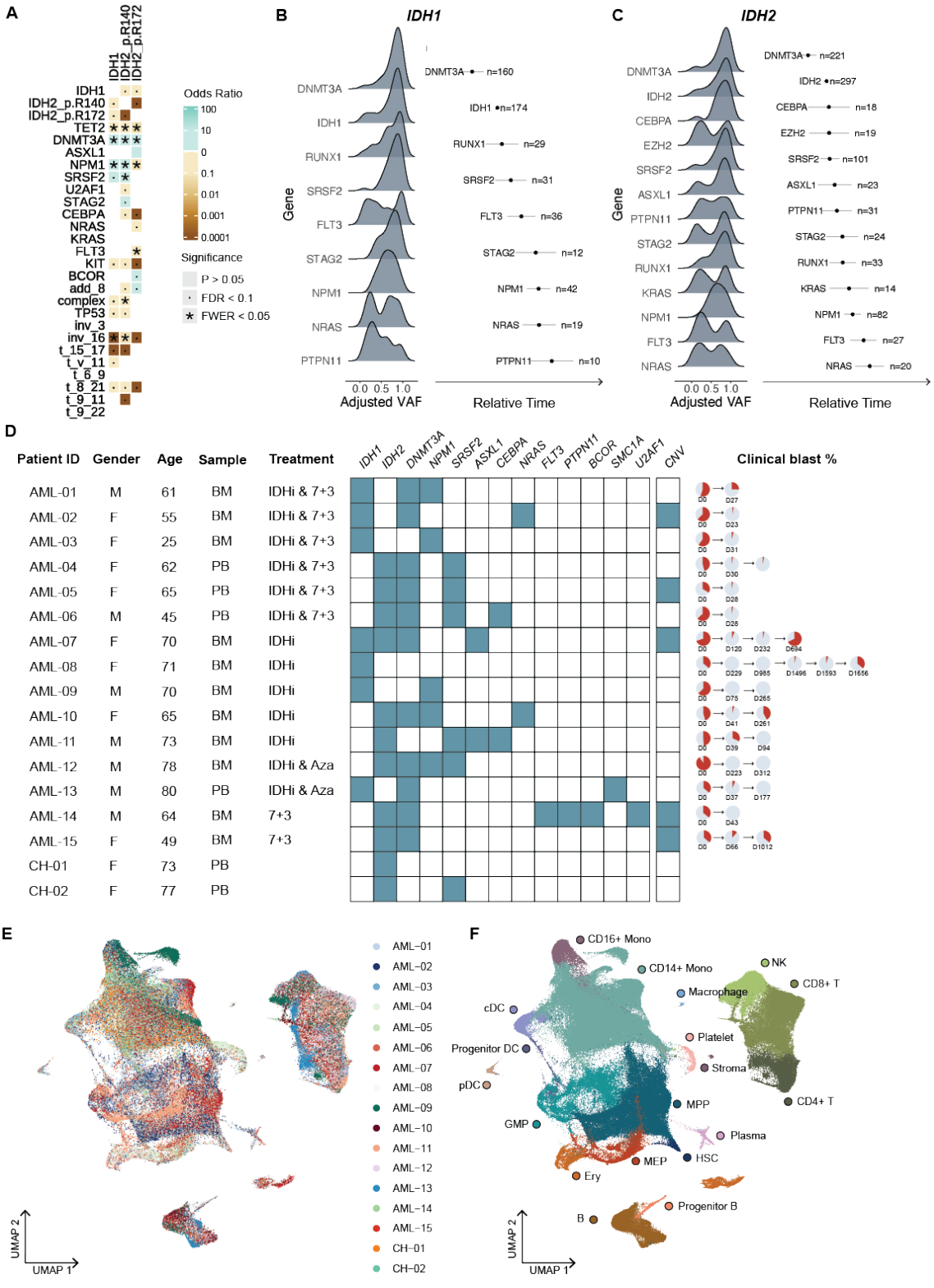


Figure 3-1 Co-mutation landscape in IDH1/2-mutant AML informs study cohort

A) Pairwise associations among gene mutations and cytogenetic alterations in 3,653 AML diagnosis AML samples, subset for *IDH1*- (n = 284) or *IDH2*-mutant (n = 604) cases for visualization. The color of each tile reflects the odds ratio for each pair whereby brown indicates mutual exclusivity (observed, relative to the expected co-mutation based on each alteration's gene frequency) and green indicates pairs that are co-mutated (found together more frequently than would be expected by each genes individual frequency). Colored tiles indicate significant relationships, tile annotation by star (*) indicates family wise error rate (FWER<0.05) or point (.) indicates false discovery rate (FDR < 0.1).
B) VAF distribution and relative order of mutation acquisition in *IDH1*- or
C) *IDH2*-mutant AML based on pairwise precedence amongst genes shown in Bradley-Terry plot. X axis reflects relative order of acquisition. Points indicate the point estimate and black lines indicate 95% confidence intervals of gene ordering in time.
D) Summary of AML and CH patients. For each patient, genetic alterations detected by targeted DNA sequencing or cytogenetics(red) are shown in the oncoprint. Pie charts represent the blast percentage reported clinically from bone marrow aspirate at the time of sample collection, indicated in days from diagnosis above the pie chart.
E) Uniform Manifold Approximation and Projection (UMAP) visualization of 259,610 single cells from integrated from 17 primary samples from unaffected (n=12), AML (n=15) and CH (n= 2) donors at all profiled timepoints. Points are colored by donor identifier.
F) UMAP colored by marker-based cell type annotation and classification of single cells to major hematopoietic cell types.

Both *IDH1* and *IDH2* are significantly co-mutated with *DNMT3A* (OR range 2.10-4.59). Enrichment for NPM1 mutations is observed in the context of *IDH1* (OR 2.35) and *IDH2* p.R140 (2.81) but not with *IDH2* p.R172 (OR 0.06). *IDH2* p.R140 co-mutation with *SRSF2* was the strongest association within *IDH1/2*-mutant AML (OR 5.45, FWER p-value 1.32e-24), corroborating their cooperatively in leukemogenesis (Yoshimi et al. 2019). However, the acquisition of NPM1 and SRSF2 mutations represent different phylogenetic trajectories in *IDH1/2*-mutated AML, as evidenced by the mutual exclusivity between *NPM1* and *SRSF2* mutations (OR 0.51, FWER p-value 3.87e-04). In *IDH1/2* AML with co-mutations in myelodysplasia related genes, *IDH2* p.R172 mutations were associated with *ASXL1* (OR 1.94) whereas *IDH2* p.R140 was associated with *STAG2* (OR 2.00). Last,

NRAS mutations were mutually exclusive to the *IDH2* p.R172 hotspot, and although mutations in *NRAS* are frequently seen in *IDH1/2* mutated AML (13%) there is no evidence of a pairwise interaction between these two genes. Taken together this analysis identifies composite genotypes that are independently selected for during AML transformation.

Next, we inferred the preferential order of mutation acquisition using the variant allele fractions (VAF) of co-mutations in reference to *IDH1* (**Figure 3-1B**) or *IDH2* (**Figure 3-1C**). DNMT3A. *IDH1* and *IDH2* mutations are acquired early along with *DNMT3A,* substantiating that these are early and disease initiating events. Mutations in chromatin-spliceosome genes (eg. *SRSF2*, *STAG2*, *ASXL1*) or *NPM1* are secondary events, whereas mutations in RTK-*RAS* pathway (eg. *NRAS*, *FLT3*) are typically acquired late (Miles et al. 2020). These patterns of clonal architecture nominate biologically relevant interactions for characterization in primary patient samples.

To study the role of co-mutations in shaping disease presentation and treatment response to IDH-directed inhibitor (IDHi) monotherapy or in combination with intensive chemotherapy or hypomethylating agents, patient samples with representative genotypes as determined by prospective clinical sequencing were nominated for single cell gene expression profiling. The cohort totals 43 cryopreserved bone marrow aspirate or peripheral blood samples from 15 patients with a diagnostic sample and additional longitudinal timepoints (range: 23-1593 days, median 3 samples per patient) (**Figure 3-1D**). Of these 15 patients, 5 patients were treated with IDHi monotherapy, 6 patients were treated with IDHi in combination with intensive chemotherapy (NCT02632708), 2 patients with IDHi and hypomethylating agent azacytidine, and 2 patients with intensive chemotherapy alone. Detailed cohort characteristics are provided in **Figure 3-1B**, Methods, and Supp Table 1.

In this cohort, we had 6 AML patients with an *IDH1* mutation, 8 patients with *IDH2* mutations, as well as one patient with mutations in both *IDH1* and *IDH2* (*IDH1* 25% VAF; *IDH2* 12% VAF) (**Figure 3-1D**, Supp Table 1). *IDH1/2* VAF was consistent with *IDH1/2* mutations being acquired as early disease initiating events (Supp Table 1). Representative genotypes inclusive of *DNMT3A* mutations (n=9 patients), *NPM1* (n=5), and *SRSF2* (n=5) (Figure 1D, Supp Table) were mapped. **(Figure 3-1A).** Five patients had copy number aberrations reported by clinical karyotype and fluorescence *in situ* hybridization (FISH) (Supp Table 1). Finally, to study how *IDH2* mutations affect hematopoiesis at a pre-leukemic state, we profiled peripheral blood mononuclear cells from two donors with *IDH2* mutant CH(Bolton et al. 2020). In CH-01, *IDH2* mutation (VAF 16%) was the sole abnormality whereas in CH-02 we observed an *IDH2* mutation (VAF 17%) and an *SRSF2* (VAF 17%) mutation, both likely present in the same clone.

Single cell profiling of *IDH1/2*-mutant AML and CH

Bone marrow or peripheral blood mononuclear cells from 15 AML and 2 CH donors were profiled by single cell RNA sequencing (scRNA-seq) (**Figure 3-1E**). To enable comparison against non-malignant hematopoiesis, we integrated scRNA-seq from 12 unaffected age-matched donors (n=8) (Lasry et al. 2022) including 10 bone marrow- and 2 peripheral blood-derived samples. We analyzed 261,345 cells (AML n=191,471, CH n=16,766, healthy n=53,108) using the Seurat toolkit utilizing reciprocal principal component analysis based integration method to account for technical variance between individual samples (**Figure 1E**)(Butler et al. 2018). Cells were clustered based on gene expression and each cluster was labeled using established markers that differentiate stem and progenitor cell populations such as *AVP* (Zheng et al. 2018), *HLF*(Komorowska et al. 2017; Karamitros

et al. 2018), and *CRHBP* (Pellin et al. 2019) for hematopoietic stem cells (HSCs), *SELL* and *SPINK2* for multipotent progenitors (MPPs) (Velten et al. 2017; Popescu et al. 2019), *AZU1*, *ELANE*, and *MPO*(Karamitros et al. 2018; Popescu et al. 2019) for granulocyte-macrophage progenitors (GMPs), *PLEK* and *GATA2* for megakaryocyte-erythrocyte progenitor (MEPs)(Y.-C. Lu et al. 2018), and *GATA1* and *HBB* for Erythroblasts (Psaila et al. 2016; Pellin et al. 2019) (**Figures 3-1F, 3-2A**). Combined Indexing of Transcriptomes and Epitopes sequencing (CITE-seq) was performed on a subset of samples (n=9) to facilitate cell type annotation (Stoeckius et al. 2017). The annotated cell types correspond to established cell surface markers such as CD34, CD117, and CD133 in HSC and MPP, CD33 in GMP, and CD71 in MEP (Triana et al. 2021)(**Figure 3-2B)**. Through the integration of diagnostic and treated AML, CH, and normal reference samples, we generate an atlas which enables downstream analysis of transcriptional programs relevant to disease presentation and treatment response. This is the first direct comparison of healthy hematopoiesis, *IDH2* CH, and *IDH1/2* AML at the single cell level. Together with transcriptomic information of leukemic cells, we also profile cells of the immune microenvironment, including myeloid and lymphoid populations (**Figure 3-1F**).

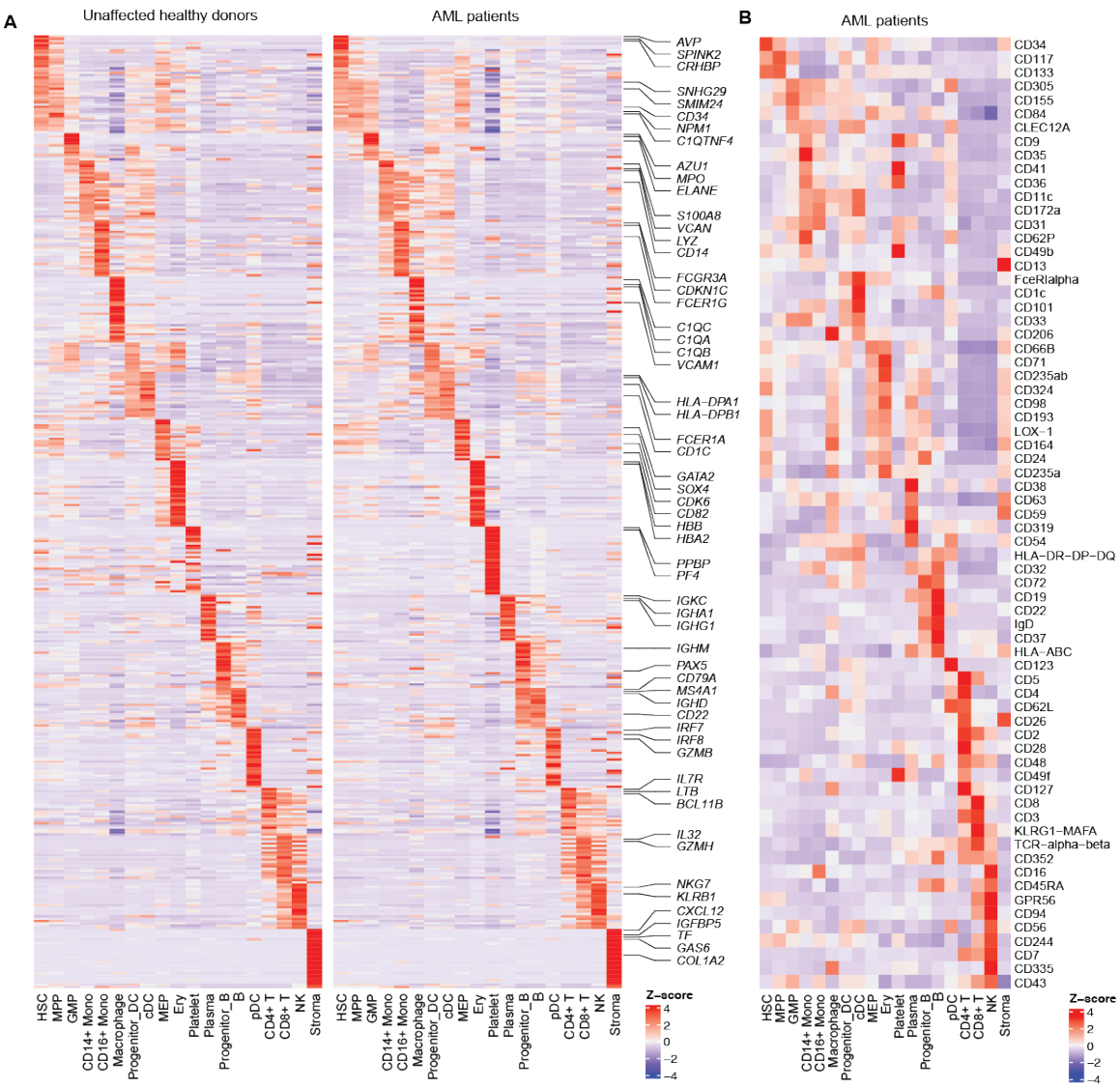


Figure 3-2 Marker expression informs cell type annotation

A) Heatmap of average expression of top gene markers from gene expression analysis across different cell types in unaffected healthy donors (n=12, left) and AML patients (n= 15, right).

B) Heatmap of average expression of top surface protein markers as measured by CITE-seq for different cell types in AML patients.

IDH1/2-Mutant AML at diagnosis is characterized by a stem-like and inflammatory phenotype

Previous studies have demonstrated the role of somatic mutations in shaping the cellular hierarchy in AML (A. G. X. Zeng et al. 2022; van Galen et al. 2019). However, the impact of *IDH1/2* mutations as well as serially acquired mutations on this hierarchy has not been well characterized. To define the cellular hierarchy and transcriptional landscape of *IDH1/2* mutant AML at diagnosis, we compared the cell type representation to healthy donors (**Figures 3-3A** , **3-4A**). We observed that AML patients had an expansion of hematopoietic stem- and progenitor-like populations including HSC/MPP-like, and GMP-like cells, as well as a marked reduction in mature myeloid populations including monocytic and dendritic cells (**Figures 3-3B,C**, **3-4A**). Preclinical models of *IDH1/2*-mutant AML have described *IDH1/2*-mutant blasts as stem-like (Lev M. Kats et al. 2014; Zhang et al. 2020; L. M. Kats et al. 2017), which led us to assess whether *IDH1/2*-mutant AML is particularly stem-like compared to other AML genotypes. We compared *IDH1/2*-mutant to non-IDH mutant AML from our recent study(Lasry et al. 2022). We observed significantly more HSC/MPP-like cells in *IDH*-mutant AML compared to non-IDH mutant AML (n=11) (**Figure 3-3D**). In line with this enrichment, *IDH1/2*-mutant AML had upregulation of the HSC and progenitor-like gene set compared to non-IDH mutant AML within the MPP, GMP and $CD14^{+}$ monocyte compartments (**Figure 3-3E**). To validate this finding, we applied cell type deconvolution(Tsoucas et al. 2019) to the Alliance Cohort (Lasry et al. 2022), an independent, large cohort (n=848) of clinically and molecularly profiled (RNA-(Lasry et al. 2022) and targeted DNA-sequencing) adult AML patients. We found that compared to non-IDH mutant AML patients (n=719), samples harboring *IDH1/2* mutations (n=129) indeed had a higher ratio of HSC/MPP-like cells and a lower ratio of mature myeloid populations (**Figure 3-4B**).

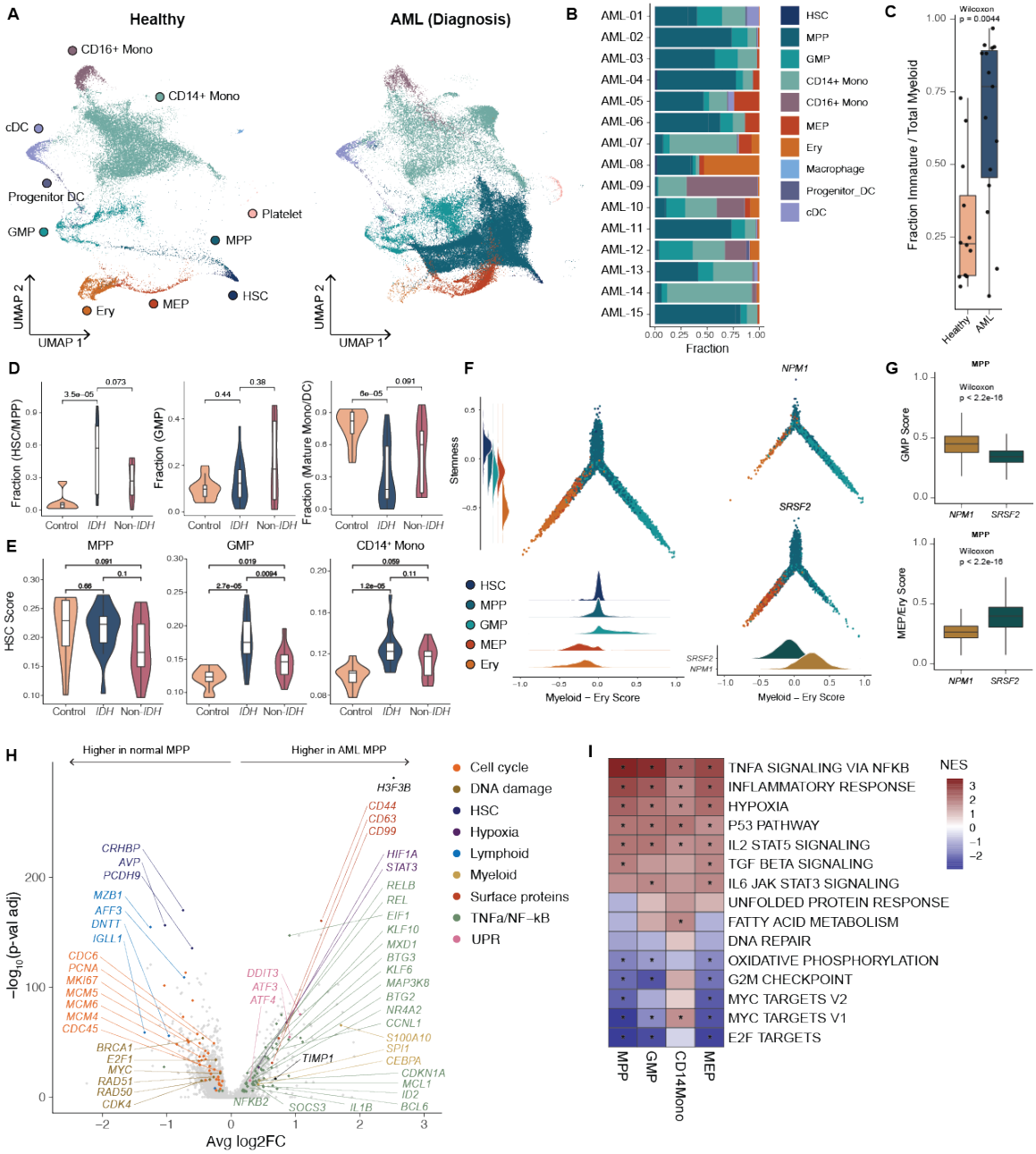


Figure 3-3 Stem-like and inflammatory phenotype of IDH1/2-mutant AML at diagnosis shaped by co-mutations in NPM1 and SRSF2

A) UMAP visualization of single cells from unaffected donors (left, n=12 donors, n= 27,195 cells) and AML patients at diagnosis (right, n=15 donors, 65,037 cells), colored by cell type.

B) Stacked barplot showing cell type representation of diagnosis samples.

C) Boxplot showing the fraction of cells from immature (HSC, MPP, GMP, MEP, Ery, and progenitor DC) over total myeloid (additionally includes $CD14^+$ monocytes, $CD16^+$ monocytes, cDC, and macrophage). Each point represents a single patient. Wilcoxon rank-sum test performed to measure differences in representation between unaffected donors and diagnosis groups, with *p*-value indicated on plots. Horizontal lines in the boxplots represent the median, the lower and upper hinges correspond to the first and third quartiles, and the whiskers extend from the hinge up to 1.5 times the interquartile range from the hinge.
D) Violin plots showing the fraction of cells in the myeloid lineage; HSC/MPP (left), GMP (middle), mature monocytes/DCs (right). Each data point represents each donor from the control group (n=12), the *IDH1/2* mutant AML group (n=17), and the non-*IDH* mutant group (n=11). Wilcoxon rank-sum test was performed for statistical significance.
E) Violin plots showing the average HSC score of MPP (left), GMP (middle), and $CD14^+$ monocytes of each individual. The control group (n=12), the *IDH1/2* mutant AML group (n=17), and non-*IDH* mutant group (n=11). Wilcoxon rank-sum test was performed for statistical significance.
F) Cellular hierarchy of AML diagnosis blasts (n cells =42,994). Lineage and stemness scores from 15 patient samples. Distributions in the margin reflect cell-type specific expression of lineage and stemness modules. Right, Cellular hierarchy similar as in left, split by co-mutation with NPM1 (n=4 patients) or SRSF2 (n=4 patients) with lineage score distribution in the bottom margin.
G) Boxplots of GMP score (top) and MEP/ery score (bottom) in the MPP compartment from *NPM1*- or *SRSF2*-mutant patients. Wilcoxon rank-sum test performed to measure differences in representation between *NPM1*- and *SRSF2*-mutant patients, with *p*-value indicated on plots.
H) Volcano plot depicting genes enriched (right) or depleted (left) in MPPs derived from *IDH1/2*-mutant AML patients (n=15) diagnosis samples compared to unaffected donor MPPs (n=12), cells downsampled to 250 per donor.
I) Gene set enrichment analysis performed per celltype comparing diagnosis AML-derived cells to counterparts from unaffected donors, cells downsampled to 250 per donor per cell type. Heatmap depicts normalized enrichment scores of Hallmark pathways.

To further define hematopoietic lineage biases, individual cells were scored for the expression of stemness and lineage transcriptional modules (van Galen et al. 2019; Velten et al. 2017). We observed that *IDH1/2* mutant AML samples comprised a mixture of HSC/MPP-like, GMP-like and MEP/Ery-like cells (**Figure 3-3F**). Previous studies have demonstrated that driver mutations shape cellular hierarchy in AML (van Galen et al. 2019;

A. G. X. Zeng et al. 2022; Beneyto-Calabuig et al. 2022; Miles et al. 2020). Thus, we next asked how the most frequent co-mutations, *NPM1* (n=4) and *SRSF2* (n=4) in *IDH1/2* mutant AML affect lineage biases. We observed that patients with co-occurring *NPM1* and *SRSF2* mutations exhibited granulo-monocytic (GMP-like) and megakaryocytic/erythroid (MEP/Ery-like) biases, respectively, exhibited by the cell type representation (**Figures 3-3F**). We examined the MPP-like population within *NPM1*- and *SRSF2*-co-mutated patients and found that *NPM1*-mutant AML-derived MPPs upregulated the GMP lineage module compared to *SRSF2*-mutant AML-derived MPPs, which upregulated the Megakaryocyte/Ery lineage module (**Figures 3-3G**). Overall, *IDH1/2*-mutant AML has a stem-like phenotype which is further primed to the GMP or erythroid lineage by the frequent co-mutations *NPM1* and *SRSF2*, respectively.

Next, to identify genes and pathways dysregulated in *IDH1/2*-mutant AML blasts compared to normal hematopoiesis, we performed differential gene expression analysis between blast and normal cell types (e.g. MPP-like blasts from AML vs MPPs from healthy donors) (**Figure 3-3H**). Since the most overrepresented cell population shared amongst mIDH AML patients is MPP-like cells (**Figure 3-3B**), therefore, we initially focused on this subpopulation. AML-derived MPP-like cells exhibited myeloid bias by downregulation of normal HSC markers (*CRHBP, PCDH9*, and *AVP*)(Zheng et al. 2018; Karamitros et al. 2018; Pellin et al. 2019), downregulation of lymphoid factors including *IGLL1*, *DNTT*, *AFF3*, and *MZB1,* and the simultaneous upregulation of myeloid factors (*SPI1*, *CEBPA/B*, *S100A9/10*). AML-derived MPP-like cells also upregulated hypoxia-related factors (e.g *HIF1A*), unfolded protein response genes (*DDIT3*, *ATF3/4*), NF-KB signaling (*NFKB2*, *REL*, *RELB*), as well as cell surface proteins previously identified as upregulated in AML (*CD44*, and *CD99*). The surface receptor *CD63* and its ligand *TIMP1*, associated with pro-survival and *HIF1A* activation (Forte et al. 2017), were also upregulated in AML MPP.

Gene set enrichment analysis using the Molecular Signatures Database Hallmark gene sets showed that AML-derived cells significantly upregulate TNFa/NF-KB signaling, inflammatory response, IL6/JAK/STAT3 signaling, and TGFbeta signaling, demonstrating the inflammatory phenotype of *IDH1/2* mutant AML (**Figure 3-3I**).

AML-derived MPPs downregulated proliferation-associated pathways including G2M Checkpoint, MYC Targets, E2F targets, and Oxidative Phosphorylation, suggesting that *IDH1/2* mutant AML exhibits a less proliferative phenotype (**Figure 3-3I**). Furthermore, AML-derived MPPs downregulated DNA damage response pathway genes (*BRCA1, MYC,* and *RAD50/51)*(**Figure 3-3H**). These trends are shared among progenitor cell types including GMPs and MEPs (**Figure 3-3I**). Finally, we asked whether these pathways broadly reflect AML disease pathogenesis or specifically *IDH1/2*-mutant AML biology. Scoring *IDH1/2*-mutant and other subtypes of AML for the expression of these gene sets across MPPs (**Figure 3-4C**), GMPs (**Figure 3-4D**), and $CD14^+$ Monocytes (**Figure 3-4E**) confirmed that *IDH1/2* AML specifically upregulated inflammation-associated gene sets in MPPs and GMPs, but not $CD14^+$ Monocytes, including TNFα, IL6-JAK-STAT, and TGFβ signaling compared to other AML and downregulated proliferation-associated pathways including G2M Checkpoint and E2F- and MYC-Targets. Together, these results highlight the inflammatory and less proliferative phenotype of *IDH1/2* AML.

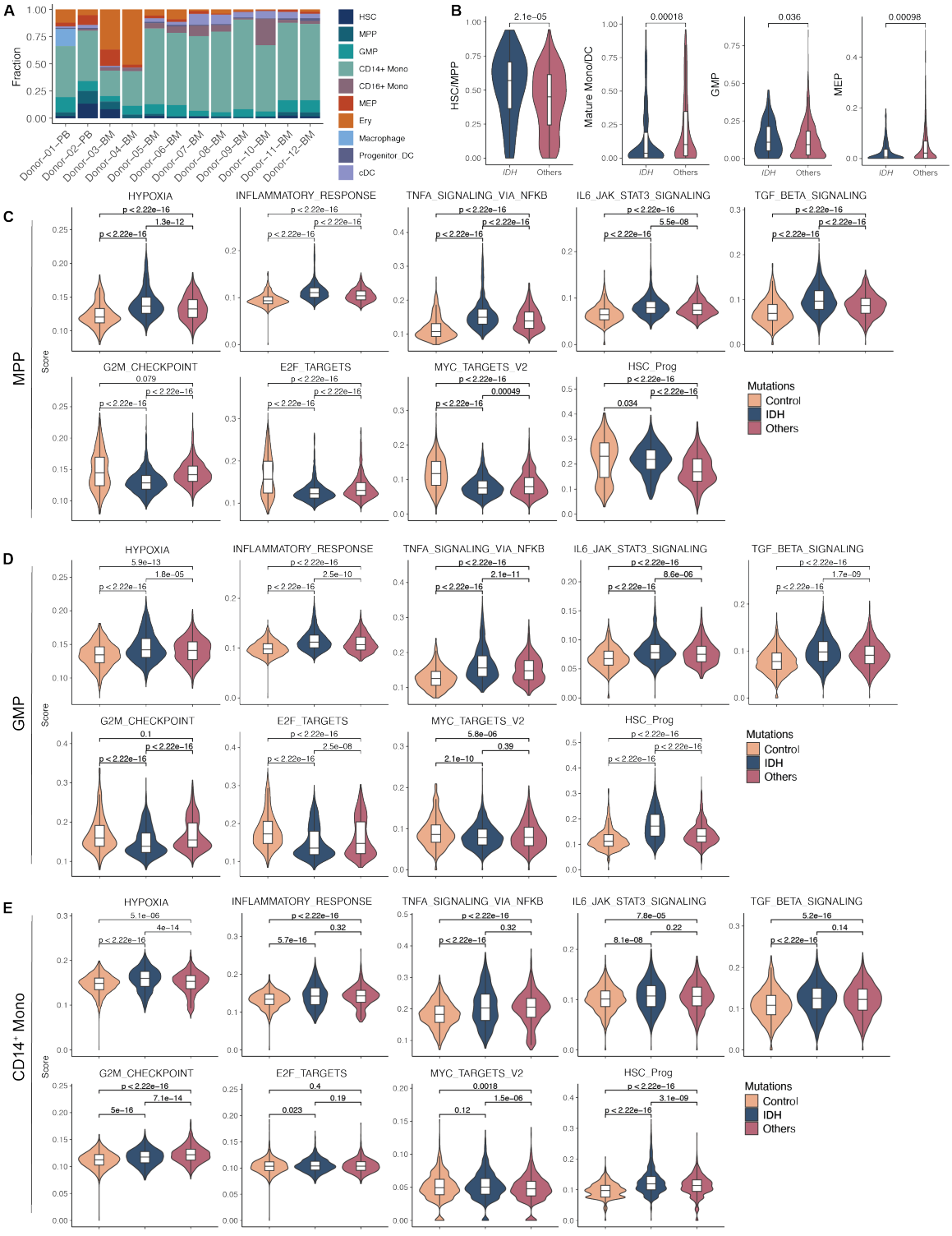


Figure 3-4 Gene expression programs specific to IDH1/2-mutant AML

A) Stacked barplot of cell type representation per sample, including 12 unaffected donors.

B) Violin plots showing the estimated fractions of HSC/MPP, mature mono/DC, GMP, and MEP in the bulk RNA sequencing data from the Alliance cohort (n=872). Each cell compartment of the *IDH1/2*-mutant AML group (n=129) and non-*IDH* mutant AML group (n=719) was compared using unpaired Wilcoxon rank sum test.
C) D) E) Violin plots showing the pathway scores of MPP(C), GMP(D), and $CD14^{+}$ monocytes(E) of 100 downsampled cells per sample: control group (n samples =12), the *IDH1/2* mutant AML group (n=17), and non-*IDH* mutant group (n=11). Unpaired Wilcoxon rank sum test was performed for statistical significance.

Integrative single cell profiling delineates clonal responses to combination of IDH1/2 inhibitors and intensive chemotherapy

Targeted IDH1/2 inhibitors Ivosidenib and Enasidenib have been clinically investigated in combination with intensive chemotherapy for patients harboring *IDH1* or *IDH2* mutations, respectively (Stein et al. 2021). We profiled patients treated with Ivosidenib (n=3) or Enasidenib (n=3) in combination with intensive chemotherapy at diagnosis and a median of 4 weeks after therapy (range: 3.3-4.5 weeks). At post-treatment sampling time, 3 patients had count recovery (**Figure 3-5A**). Five patients achieved complete remission, of which 4 patients ultimately underwent hematopoietic stem cell transplantation. These 5 patients were still alive at the time of last follow up. The 6th patient, AML-04, had partial response at post-treatment sampling time and was deceased after induction and consolidation therapy.

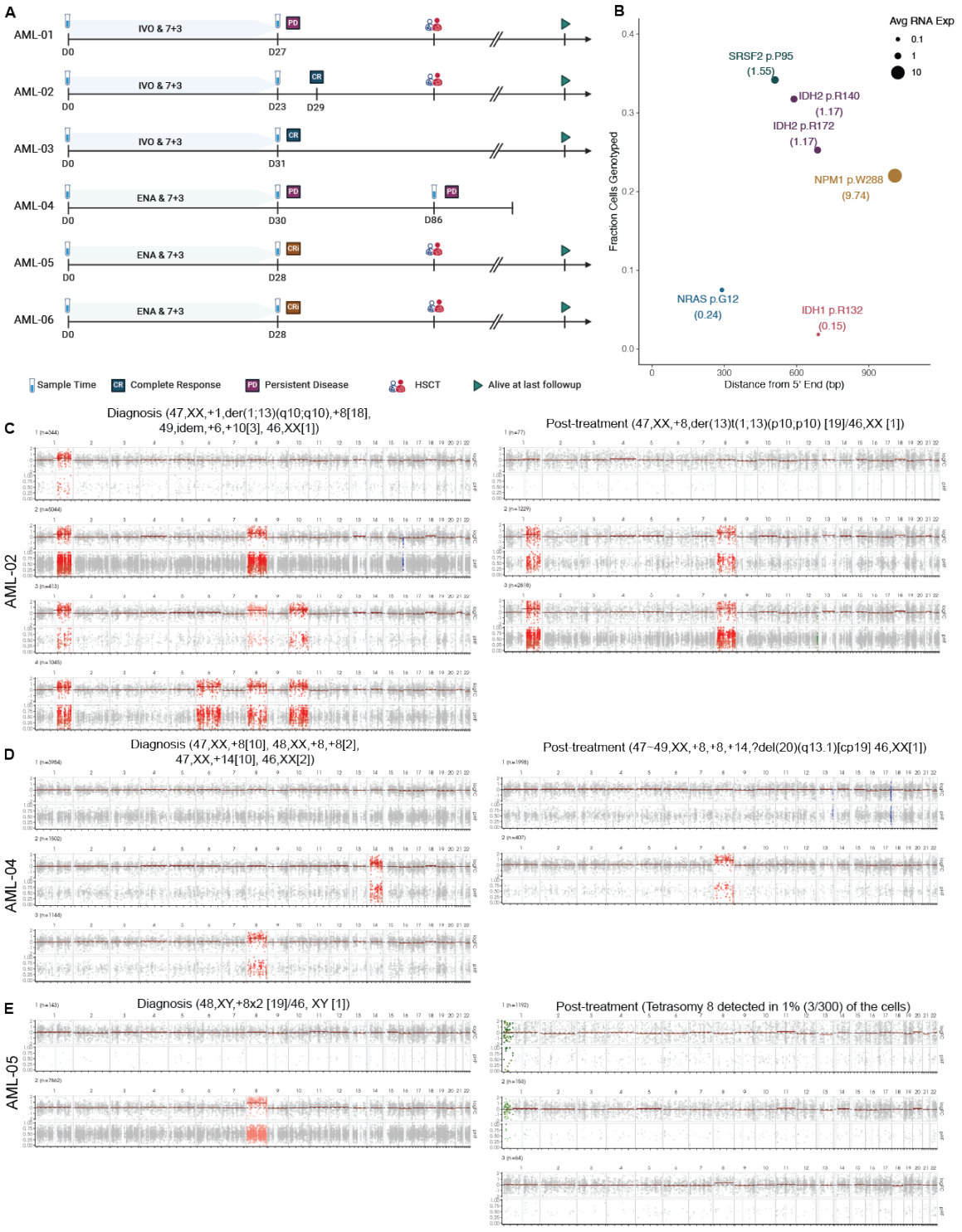


Figure 3-5 Patient treatment overview and identification of somatic mutations and copy number alterations from single cell transcriptomic data

A) Timelines of clinical history, including treatment and response ascertainment, for 6 patients treated with combination of IDH-inhibitors and intensive chemotherapy.

B) Genotyping efficiency of mutations targeted by genotyping of transcriptomes. Distance from 5' end was calculated using the most frequently expressed transcript in hematopoietic cells. Average gene expression for each gene (in parentheses) was calculated within the patients with mutations in that gene.
C) D) E) Numbat clone-specific pseudobulk profiles for n=3 patients at diagnosis (left) and post-treatment (right) with copy number events detected in clinical cytogenetics. Clinical karyotype is annotated for each sample.

Single cell gene expression profiling of these patients at diagnosis (**Figure 3-6A)** and after treatment (**Figure 3-6B)** with IDH1/2 inhibitors and chemotherapy revealed a cell type shift from predominantly immature cells (median fraction of immature cells 0.85) to a recovery of the monocytic population after treatment (median fraction immature cells 0.228)(**Figure 3C, 3D**). While AML-04 showed a reduction of immature cells after therapy, their immature cell fraction (0.662) falls within the range for diagnosis samples (0.657-0.912), consistent with the existence of persistent disease.

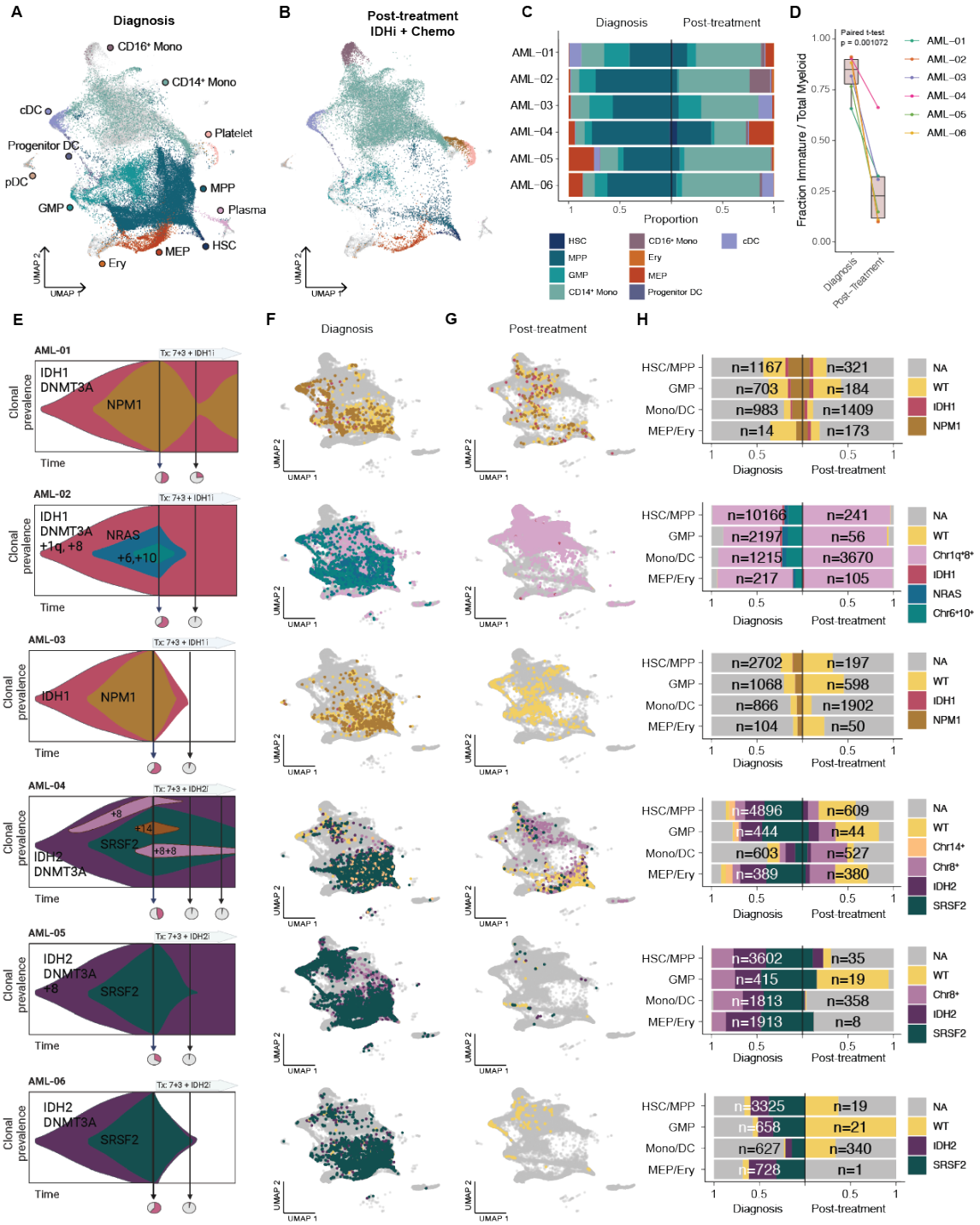


Figure 3-6 Tracking clonal and cellular evolution using integrated single cell gene expression and genotyping under combination of IDHi and intensive chemotherapy

A) UMAP visualization of 41,156 single cells from 6 patients' diagnosis samples.
B) UMAP visualization of 16,913 single cells from 6 patients' post-treatment samples.

C) Stacked barplot showing cell type representation per patient at diagnosis and post-treatment. Total number of cells per sample is indicated on each bar.
D) Boxplot showing the fraction of cells from immature (HSC, MPP, GMP, MEP, Ery, and progenitor DC) over total myeloid (additionally includes CD14+ mono, CD16+ mono, cDC, and macrophage). Each point represents a single patient.
E) Fish plots depicting clonal evolution from diagnosis and after therapy. Sampling timepoints (diagnosis and after therapy) are indicated by vertical arrows and pie chart with corresponding blast percentage by bone marrow aspirate. Fish plots integrate VAFs from targeted DNA sequencing and clinical cytogenetics.
F) Diagnosis sample cells colored by genotype status. Gray cells in the background represent unaffected donor UMAP for reference.
G) UMAP of post-treatment cells colored by genotype as in f).
H) Stacked barplot of cells harboring mutations identified by genotyping of transcriptomes or copy number events detected by Numbat, split by diagnosis (left) and after treatment (right). If cells had multiple events, assignment was prioritized by the most subclonal event or in cells with both mutations and copy number events, cells were colored by mutation. “WT” represents cells with only wildtype alleles detected for mutations assayed and NA represents cells with no identified events by genotyping or copy number state inference. Number of total cells per time point per cell type is indicated on the plot.

To infer evolutionary histories in individual AML patients, we integrated targeted bulk DNA-sequencing and clinical cytogenetics from diagnosis and post-treatment samples (**Figure 3-6E**) with single cell genotyping and gene expression profiling. Briefly, for each cell, multiplexed Genotyping of Transcriptomes (GoT) (Nam et al. 2019) enabled the detection of acquired mutations that were previously identified by clinical sequencing of the primary bulk sample. Inference of chromosomal copy number state from the scRNA-seq data further enabled the identification of distinct copy number alterations (CNA) to include allelic imbalances that were previously detected by clinical cytogenetics or FISH(Gao et al. 2022). Last, integration of mutation and copy number status with gene expression data allowed us to assign each cell across the hematopoietic hierarchy. This analytical workflow enabled us to map the presence of acquired genetic alterations for each cell type.

For GoT, target clone-defining mutations were selected by analysis of VAF from targeted DNA sequencing and sufficient gene expression. Cells were profiled using the 5' 10X Genomics kit to maximize GoT efficiency, as most mutations were within 1000bp of the 5' end (*IDH1/2*, *NRAS*, *NPM1*, and *SRSF2*) (**Figure 3-5B**). By GoT, 23% of cells at diagnosis were genotyped for at least one clone-defining mutation including *IDH1* (n=2%), *IDH2 p.R140 (32%)*, IDH2 p.R172 (25%), *NPM1 (22%)*, *NRAS* (8%) and *SRSF2 (34%)*. We did not genotype *DNMT3A* due to technical limitations of low gene expression (Average expression in single cell data: 0.19; comparable to IDH1 0.15) and >2000 base pair mutation distance from 5' end. Additionally, three patients (AML-02, AML-04, AML-05) had CNA reported by clinical karyotyping and FISH analysis, all of which were detected by inference of copy number profiles from single cell transcriptomes (**Figure 3-5C**-**E**). No cytogenetic events beyond those clinically reported were discovered.

To characterize where in the hematopoietic hierarchy mutant AML cells reside at diagnosis, mutation and CNA status was overlaid on the single cell gene expression-derived clusters (**Figure 3-6F,H**). Indeed, at diagnosis, we observe mutant cells in both progenitor and mature compartments, suggesting that the differentiation block is not complete, and even cells harboring multiple driver mutations including *IDH1/2*, *NPM1*, *NRAS*, and/or *SRSF2* are able to differentiate into the $CD14^{+}$ monocyte compartment (**Figure 3-6F,H**).

Next, we extended these approaches to delineate clonal responses to therapy (**Figure 3-6E, G, H**). We observed that in 4 cases (AML-01, AML-02, AML-04, AML-05), the ancestral clone, defined by *IDH1/2* +/- *DNMT3A* mutations retained clonal dominance after therapy. In all of these cases, the subclones defined by mutations in *NPM1*, *NRAS*, or *SRSF2*, were either markedly reduced compared to the diagnosis timepoint or not detectable by

GoT or targeted DNA sequencing (limit of detection 2% due to NGS error rate). This is consistent with the observations from clinical studies that report persistence of *IDH1/2* mutations after IDH-inhibitor or chemotherapy does not preclude complete remission (Stein et al. 2021; DiNardo et al. 2018; Wiseman et al. 2016). In the other two cases (AML-03, AML-06), we observe that all mutations, including *IDH1/2*, *DNMT3A*, *NPM1* and *SRSF2* were not detected by targeted sequencing or GoT after therapy. We observe that the two *NPM1* co-mutated cases (AML-01 and AML-03) have different trajectories after treatment. While AML-03 does not have detectable *NPM1* mutation after therapy and is relapse free within 6.5 years follow up, AML-01 has persistent *NPM1* at 4% VAF and relapses after drug discontinuation with a concomitant increase in blasts (60%) and *NPM1* VAF (26%) (**Figure 3-5A**, Supp Table 1). These clinical trajectories are aligned with reports that *NPM1*-mutant measurable residual disease after chemotherapy is associated with higher risk of relapse (Ivey et al. 2016; R. Dillon et al. 2020; L. W. Dillon et al. 2022).

To further investigate how these genetic clones reconstituted the hematopoietic hierarchy after IDH inhibitor and chemotherapy, we examined whether mutant cells were stem- and progenitor-like cells or differentiated towards the myeloid lineage (**Figure 3-6H**). In the four samples without mutation clearance post therapy (AML-01, AML-02, AML-04, AML-05), we detected mutant cells in all HSPC-like and mature myeloid compartments. For example, the remission sample for AML-02 reflects the clearance of the *NRAS* and +6/+10 subclones, and the reconstitution of hematopoiesis by the ancestral clone (*IDH1*, *DNMT3A*, +1q/+8) evidenced by clonal *DNMT3A* and *IDH1* VAF (45% and 37%, respectively) and detection of chromosomal gains of Chr1q;8 (in 98% of cells) from the single cell expression data. Patients AML-05 and AML-06 had complete remission with incomplete count recovery, precluding a robust analysis of cell type representation and mutant status at the time of post-treatment sampling. Taken together, we find that parent

clones defined by *IDH1/2* +/- *DNMT3A* or copy number events can respond to combination therapy by differentiation and reconstitution of hematopoiesis.

Defining clone-specific phenotypes in diagnosis AML

Thus far, we demonstrated that common composite genotypes in *IDH1/2*-mutant AML exhibit differential lineage biases and responses to therapy. Next, we investigate the gene expression programs that may underlie these observations. Two patients, AML-02 and AML-04, have multiple discernable clones at diagnosis, including a parental clone defined by *IDH1/2* + *DNMT3A* mutations and subclones with signaling (*NRAS)* or splicing mutations (*SRSF2)*, respectively **(Figure 3-6E)**. This provides the unique opportunity to investigate the impact of sequentially acquired somatic mutations on hematopoiesis while controlling for all potential patient-specific and environmental variables.

AML-02 is a female patient with 66% blasts reported by flow cytometry at diagnosis. Phylogenetic reconstruction of bulk targeted sequencing data and clinical FISH identify three clones: 1) an ancestral clone harboring *IDH1* (VAF 30%), *DNMT3A* (VAF 32%), Gain of Chr 1q (FISH 66% of cells), gain of Chr 8 (FISH 63%); 2) a subclone defined by *NRAS* (VAF 10%) and; 3) a subclone defined by an independent gain of Chr 6 (FISH 17%), Chr 10 (FISH 16%) (Supp Table 1). Gene expression-based clustering of cells from the diagnosis sample revealed two predominant clusters (Population 1 and 2) with mixed composition of MPP-like and GMP-like cells **(Figure 3-7A)**. We evaluated whether these cell clusters could represent different genetic clones within this sample. Overlaying inferred copy number states **(Figure 3-7B)** and the mutation status **(Figure 3-7C)** revealed that *IDH1*-mutant Chr1^{+}8^{+} gain cells (*IDH1*-mutant parent clone) were uniformly distributed throughout all gene expression clusters, while *NRAS* mutant cells (*NRAS* co-mutant

subclone*)* were predominantly confined to Population 2 and a subset of *NRAS* co-mutant subclone harbored gains in Chr 6 and 10 (*NRAS*/Chr6^{+}10^{+} co-mutant subclone*)* formed a subcluster in Population 2 (**Figure 3-7B, C**). Based on this, we assign each cell to a phylogenetic clone.

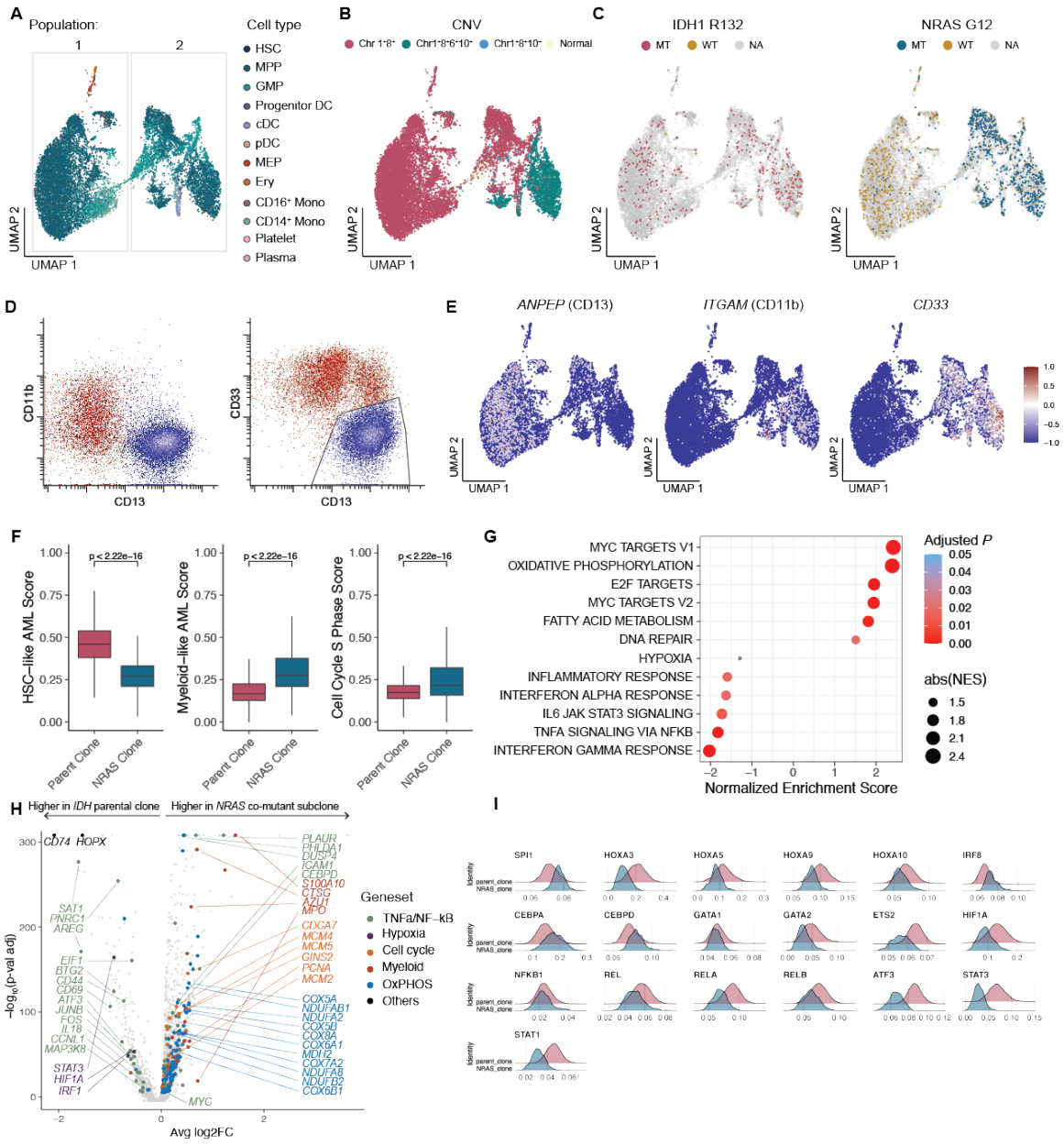


Figure 3-7 Distinct phenotypes of a IDH1 mutant parent clone and NRAS co-mutant subclone in AML patient at diagnosis

A) UMAP of diagnosis sample from patient AML-02 (n=14,155 cells) colored by cell type.
B) UMAP as in A) colored by copy number-defined clones predicted by Numbat.
C) UMAP as in A) colored by genotyping status for IDH1 (n=338 genotyped cells) (left) and NRAS (n = 1184 genotyped cells) (right).
D) Clinical flow cytometry of bone marrow aspirate identifies distinct blast population phenotypes within the $CD34^+$ population.
E) UMAPs colored by scaled gene expression of surface markers from d).
F) Expression of stemness, myeloid, and cell cycle gene modules in the parent clone and NRAS clone. Wilcoxon rank-sum test performed to measure differences in expression between cells from parent vs NRAS-mutant clone, with p value indicated on plots.
G) Gene set enrichment analysis in MPPs belonging to the *NRAS* clone versus parent clone. Size of the point represents the absolute value of the Normalized Enrichment Score (NES) and the shading the adjusted p-value.
H) Volcano plot showing differentially expressed genes between *IDH1* parent clone and *NRAS* clone. Selective genes from the gene set enrichment analysis g) are shown.
I) Activity of selected transcription factors relevant to enriched pathways from g) calculated by Area Under the Curve (AUCell). Expression of transcription factor target genes was ascertained using SCENIC.

Evaluation of the clinical flow profile for this sample also identified three distinct blast populations, which were defined by differential expression of CD11b, CD13, and CD33 in the $CD34^+$ population (Figure 4D). Thus, we next investigated whether these two blast populations were associated with the *IDH1*-mutant parental clone and the *NRAS* co-mutant subclone. We found that CD13 (*ANPEP*) is highly expressed in the parent clone while CD33 and CD11b (*ITGAM*), markers of myeloid differentiation, are expressed in the *NRAS* co-mutant subclone and *NRAS*/chr6^+10^+ co-mutant subclone, respectively (**Figure 3-7E**). This phenotype is aligned with a previous report of upregulated CD11b expression in MAPK/ERK signaling mutation subclones compared to *IDH1* mutant clones (Miles et al. 2020). Next, we investigated whether these markers of differentiation were correlated with transcriptional programs and found that, indeed, the *IDH1*-mutant parental clone had

higher expression of HSC-like modules and the *NRAS*-subclone had higher expression of the more differentiated myeloid-like module (**Figure 3-7F**). Furthermore, the *NRAS* co-mutant subclone had higher cell cycle activity than the *IDH1*-mutant parent clone (**Figure 3-7F**).

To understand how clone-defining mutations shape the transcriptional phenotype of AML blasts, differential gene expression analysis was performed between the *IDH1* mutant parent clone and the *NRAS* co-mutant subclone. While we previously observed that *IDH1/2* AML blasts downregulate *MYC* and MYC targets and E2F targets compared to their wildtype counterparts, we find that the *NRAS* subclone upregulates these pathways (**Figure 3-7G**). *NRAS* co-mutation is associated with non-response in the setting of ivosidenib monotherapy in relapsed/refractory AML (DiNardo et al. 2018), but not in combination chemo- and ivosidenib therapy (Stein et al. 2021). Indeed, the *NRAS* co-mutant clone was not detected at complete remission in AML-02 after combination therapy. Our data suggest that the high chemosensitivity of *NRAS* co-mutant clones (Ball et al. 2021) may be due to their proliferative phenotype.

Conversely, *IDH1/2* and/or *DNMT3A* mutations can persist after therapy, leading us to further characterize the parent clone phenotype at diagnosis. Compared to the *NRAS* subclone, the parental clone had higher expression of the Hypoxia (*HIF1A*) and TNFA Signaling via NF-KB (*CD44, ATF3, AREG, MAP3K8*), inflammatory response, and IL6-JAK-STAT3 signaling pathways **(Figure 3-7G, H)**. To investigate whether the increased expression of these key transcription factors had global effects on shaping the transcriptome, SCENIC (Aibar et al. 2017) was employed to analyze the activity of targets of these transcription factors. The expression of genes with AP-1, HIF1A, and NF-KB transcription factor binding sites were also induced in the parent clone **(Figure 3-7I)**.

Next, we investigated whether this paradigm of a hypoxic and inflammatory parental and proliferative leukemic subclone was mirrored in the clones identified in a second *IDH/DNMT3A*-mutated patient AML-04. AML-04 is a 62-year old female with 46% blasts at diagnosis by bone marrow aspirate smear with two clones identified by bulk targeted sequencing: an ancestral clone harboring *IDH2* (VAF 39%) and *DNMT3A* (VAF 38%) mutations and a leukemic subclone defined by *SRSF2* (VAF 32%) **(Supp Table 1)**. The diagnosis sample comprises MPPs, MkEPs and a small population of $CD14^{+}$ monocytes (**Figure 3-8A**). Cells were genotyped for *IDH2* and *SRSF2* (n = 2,824 *IDH2* mutant cells; 2739 *SRSF2* mutant cells). *IDH2*-mutant cells (*IDH2* mutant parent clone) were uniformly distributed throughout the gene expression clusters, while *SRSF2* mutant cells (*SRSF2* co-mutant subclone) were notably excluded from a subpopulation of MPPs (**Figure 3-8B**). Based on the fraction of *IDH2*- and/or *SRSF2*-mutant cells in each gene expression cluster, we assigned all cells as belonging to either the *IDH2*-mutant parent clone or the *SRSF2*-mutant subclone **(Figure 3-8C, D)**.

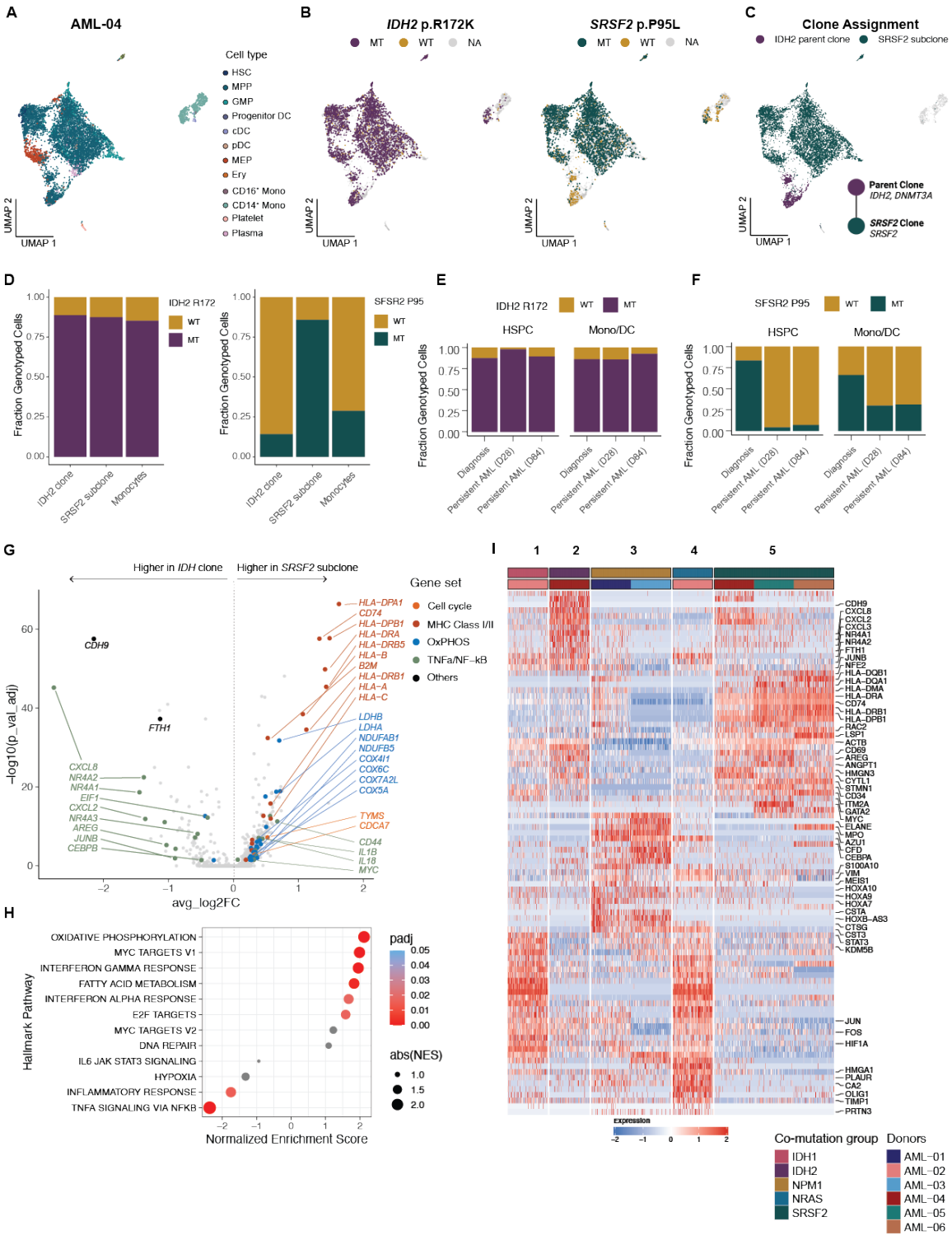


Figure 3-8 Distinct phenotypes of a IDH2 mutant clone and SRSF2 co-mutant subclone

A) UMAP visualization of diagnosis sample from AML-04 (n=6,630 cells) colored by annotated cell type.

B) Cells colored by genotype status for IDH2 p.R172K ascertained by genotyping of transcriptomes (n= 2,978) (left) and cells colored by genotype status for SRSF2 p.P95L ascertained by genotyping of transcriptomes (n= 3,149) (right).
C) Clonal assignment to *IDH2* clone and *SRSF2* subclone.
D) Barplot of fraction genotyped cells for *IDH2* R172 (left) and *SRSF2* P95 (right) per clone.
E) Fraction of genotyped cells (mutant or wildtype) for *IDH2* R172 in progenitor (HSPC) (left) or differentiated myeloid (mono/DC) compartment (right) from diagnosis to post-therapy timepoints.
F) Fraction of genotyped cells (mutant or wildtype) for *SRSF2* P95 in progenitor (HSPC) (left) or differentiated myeloid (mono/DC) compartment (right) from diagnosis to post-therapy timepoints.
G) Volcano plot depicting genes enriched (right) or depleted (left) in the *IDH2*-mutant clone versus the *SRSF2*-mutant subclone.
H) Enriched Molecular Signatures Database Hallmark gene sets in MPPs belonging to the *SRSF2*-mutant subclone versus *IDH2*-mutant clone. Size of the point represents the absolute value of the Normalized Enrichment Score and the shading the adjusted p-value.
I) Mutation-associated gene expression modules derived by assigning all diagnosis blast cells to a phylogenetic subclone. Heatmap of normalized expression of differentially expressed genes between mutation groups.

Next, we looked at how the distribution of *IDH2*- and *SRSF2*- mutant cells across progenitor populations and mature myeloid populations including monocytes and DCs changed across timepoints from diagnosis and post-treatment. Post-treatment, the VAF of *IDH2* remained clonal at 45%, while the *SRSF2* decreased to 22% so we asked whether the mutant cells post-treatment had any changes in cell type distribution. We observed that the fraction of *IDH2*-mutant cells remains constant across both immature and mature cell populations after treatment **(Figure 3-8E)**, reflecting that *IDH2*-mutant cells reconstitute both immature and mature myeloid cell populations after treatment. Conversely, the *SRSF2*-mutant fraction decreases sharply in the progenitors (from 83% of genotyped cells to 4% and 7% at D28 and D84), and to a lesser degree in the monocyte and DCs (66% to 30%), after treatment **(Figure 3-8F)**. This suggests that *SRSF2*-mutant

cells are largely eliminated from the progenitor compartment yet remain within the monocyte population. There is no significant shift in mutation fraction between Day 28 and Day 84, which represents persistent disease after consolidation chemotherapy. Together, this highlights how the parent clone and subclone differentially respond to enasidenib treatment in combination with chemotherapy.

Next, we characterized the clone-specific phenotypes at diagnosis. Differential gene expression analysis revealed that the *SRSF2*-mutant clone upregulated proliferation-associated pathways including MYC Targets and E2F Targets compared to the parental clone, similar to the clone-specific expression pattern in AML-02 **(Figure 3-8G, H)**. In addition, oxidative phosphorylation and antigen presentation machinery are upregulated in the *SRSF2*-mutant clone. Furthermore, the *SRSF2* clone had higher expression of *GATA2* and *NFE2*, erythroid-lineage transcription factors, likely driving the MEP-like phenotype. This *SRSF2* mutant erythroid lineage priming is reflected in the depletion of *SRSF2*-mutant cells in the monocyte population suggesting that monocytes are derived from the *IDH2* parent clone **(Figure 3-8F)**. In the parental clone, we observe upregulation of *HIF1A*, pro-inflammatory chemokines *CXCL2/3/8* **(Figure 3-8G)**. GSEA showed an enrichment of TNFA signaling via NFKB in the parent clone versus the *SRSF2* subclone (**Figure 3-8H**). Interestingly, we also find anti-inflammatory transcription factors *NR4A1/2/3* are overexpressed (**Figure 3-8G**), known as immunomodulators which confer HSPCs' fitness by enhancing resistance to the inflammatory environment (Avagyan et al. 2021).

This suggests that whereas mutations in epigenetic regulators such as *IDH1/2* or *DNMT3A*, which are acquired early, drive a stem-like and non-proliferative phenotype, the acquisition of additional mutations in signaling or splicing genes such as *NRAS* or *SRSF2*

drives cell proliferation, lineage skewing and impaired differentiation. Importantly, deep subclonal phenotyping at the single-cell resolution enables derivation of subclone-specific transcriptomic programs that highlight the contribution of individual somatic mutations in shaping the AML phenotype.

Deriving mutation-associated transcriptomic signatures

Thus far, we have observed the role of mutations in shaping cell type representation and clone-specific transcriptional programs. In the pre-treatment samples besides AML-02 and AML-04, all blast cells belonged to the same genetic clone as defined by mutation VAF, thus, all blast-like cells were assigned to five co-mutation groups: 1) *IDH1+DNMT3A*, 2) *IDH2+DNMT3A*, 3) *IDH1/DNMT3A+NPM1*, 4) *IDH1/2/DNMT3A+NRAS*, and 5) *IDH2/DNMT3A+SRSF2*. Differential gene expression analysis between these subgroups identifies both patient- and genotype-specific gene expression signatures (**Figure 3-8I**). Group 1 (*IDH1/DNMT3A*) had upregulated expression of *STAT3*, hypoxia (*HIF1A*) and AP-1 complex (*FOS*, *JUN*), and Group 2 (*IDH2/DNMT3A*) had overexpression of pro-inflammatory chemokines including *CXCL2/3/8*, and anti-inflammatory transcription factors *NR4A1*. Group 1 and 2 shared expression of *ATF3* and *NR4A2*. Group 3 (*IDH1/DNMT3A+NPM1*), had overexpression of GMP-like genes including *AZU1*, *ELANE*, *CTSG*, and *MPO* as well as *MLL*-target genes including the *HOX* cluster (*HOXA7/9/10, HOXB3/4/5*) and *MEIS1 (Armstrong et al. 2002)*. These cells had the lowest expression of CD34. Group 5 (*SRSF2*-co-mutant) cells showed overexpression of erythroid transcription factors including *GATA2*, *NFE2*, and *MAX*. *IDH2+SRSF2*-mutant patients also shared high expression of MHC-II antigen presentation (*CD74*, *HLA-DRA*, *HLA-DPB1*, *HLA-DRB1*) compared to other genotype groups, possibly reflecting the downregulation of MHC-II via methylation of the HLA Class II transactivator *CIITA* in *IDH1-*

and *NPM1*-mutated AML (Dufva et al. 2020). Taken together, these clonal analyses provide early insights on the role of IDH specific co-mutators in further shaping *IDH1/2*-mutant phenotype in AML.

Relapse after IDHi Monotherapy Occurs in the Context of Stemness and Inflammation and downregulation of MHC Class II

To characterize patterns of response and relapse of m*IDH1/2* AML patients upon mIDH inhibitors treatment, we profiled longitudinal samples of AML patients who were treated with mIDH inhibitors monotherapy or in combination with azacytidine. Among these 7 patients, three patients achieved complete remission while four patients eventually relapsed from treatment (**Figure 3-9A**). For AML-13, there are no relapse samples available for sequencing. Three patients (AML-07, AML-08, AML-10) relapsed an average of 352 days after treatment (range: day 112 - day 756) initiation with mIDH inhibitor monotherapy. In two cases, relapse was accompanied by the acquisition of new mutations, *KRAS* (AML-07) and *GNAS* (AML-08).

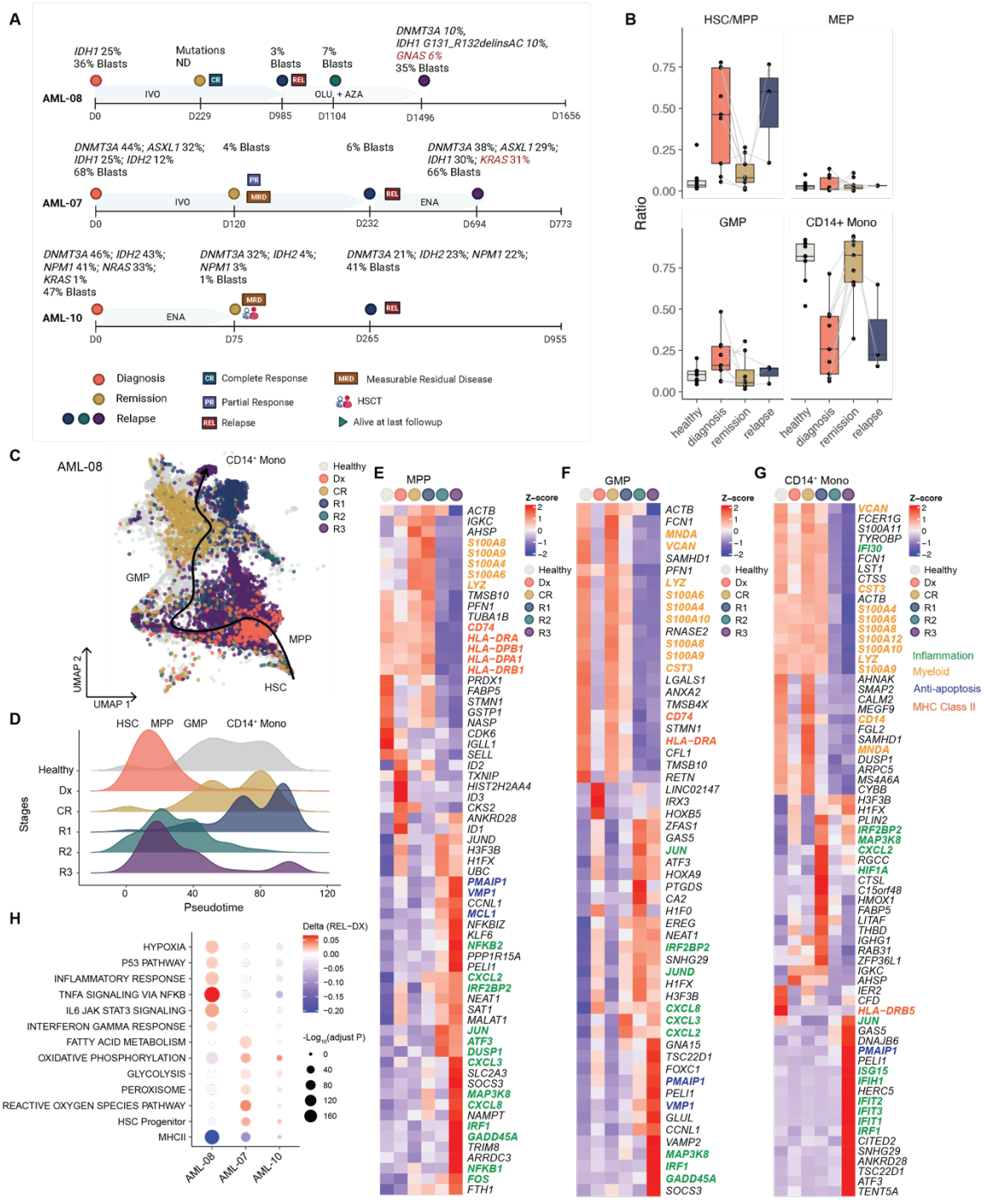


Figure 3-9 Genetic and transcriptomic patterns of relapse after IDHi monotherapy

A) Clinical time course of three AML patients who relapsed. Genes labeled as red represent RTK-RAS pathway mutations newly acquired in the relapse stage.

B) Boxplot showing ratio of myeloid subsets in total myeloid cells across different treatment stages. Each dot represents one patient and different timepoints within the same patient are connected by lines.

C) UMAP of pseudotime projecting differentiation hierarchy during longitudinal treatment stages.
D) Pseudotime projecting differentiation hierarchy during longitudinal treatment stages.
E) Heatmap showing DEGs in the MPP subset of AML-08 across different stages of treatment.
F) Heatmap showing DEGs in the GMP subset of AML-08 across different stages of treatment.
G) Heatmap showing DEGs in the $CD14^{+}$ Monocytes subset of AML-08 across different stages of treatment.
H) Pathways that are commonly and uniquely upregulated among MPP cells in relapse stages versus diagnosis stages in AML-07, AML-08 and AML-10.

Our previous analysis of the cell type composition under combination of IDHi and chemotherapy revealed a shift from predominantly immature cells to mature myeloid cells after treatment, particularly in patients who achieved complete remission. Next, we investigated whether patients treated with IDHi monotherapy +/- hypomethylating agents follow a similar trajectory, and whether this is reversed at relapse. Single cell gene expression profiling of this cohort of patients at diagnosis (n=9), remission (n=9), and relapse (n=3) revealed that although the cell type composition at remission resembles healthy bone marrow, relapse is marked by the significant outgrowth of MPP-like blasts among all three relapse patients (**Figure 3-9B**). We also observed a modest increase in GMPs but not in the other cell compartments including erythroid progenitor, DC progenitor, or mature monocytes, suggesting that MPP-like blasts are the compartment emerging at resistance.

While genetic mechanisms of relapse have been well described in *IDH1/2*-mutant AML, the transcriptional patterns of relapse after mIDH inhibition have not been well characterized. To address this, we compared myeloid cell types present at diagnosis, remission, and multiple relapse stages from three patients. Because these patients had

highly variable genomic evolution and clinical responses, we first performed the analysis on serial samples on a per-patient basis.

For AML-08, we analyzed a total five timepoints from diagnosis, complete remission, and three different relapse stages in which 3% (R1), 7% (R2), and 35% (R3) blasts were detected by flow cytometry, with R1 and R2 after Ivosidenib and R3 after Olutasidenib (second generation mIDH1i) and Azacytidine treatment (**Figure 3-9A**). Pseudotime ordering showed a shift from predominantly MPPs at diagnosis, to a distribution resembling healthy bone marrow at CR, followed by the reversion toward MPP-like cells from the early time point of relapse (R1) to the later time points (R2, R3). (**Figure 3-9C, D**). Within the MPP compartment, relapse cells significantly upregulated levels of pro-inflammatory genes including pro-inflammatory cytokines such as *CXCL8, CXCL2,* and several transcription factors including AP-1 complex (*FOS, JUN)*, *IRF1,* NF-kB complex (*NFKB1, NFKB2)*, whereas diagnosis MPPs showed the high expression level of ID protein family (*ID1/2/3*), which have been implicated in cancer initiation, maintenance and progression of various cancer types (Lasorella, Benezra, and Iavarone 2014). (**Figure 3-9E)**. We observed that a few of those genes including *CXCL2, CXCL8, IRF2BP2* were increased at R1 and further elevated at R2 and R3, but some inflammatory genes (*NFKB2, IRF1, IFIT1/2/3*) were specifically upregulated at R3, suggesting a correlation between the inflammation level and disease progression. In addition, we found *MCL1* and *PMAIP1*, pro-survival members of the BCL-2 family, are increasingly upregulated through the three relapse stages (**Figure 3-9E**). Next, we investigated the gene expression patterns of GMP- and monocyte-like cells over time. GMP-like cells at relapse generally share the pattern of upregulated pro-inflammatory genes with MPP-like cells in respective relapse time points (**Figure 3-9F**). Interestingly, $CD14^{+}$ monocytes showed two distinct inflammation-associated gene modules at early and late relapse time points (**Figure 3-**

9G). At R1, which pre-dates the outgrowth of MPPs in later stages of relapse, a set of hypoxia and inflammation related genes including *CXCL2, HIF1A, CTSL, HMOX1, THBD* were overexpressed in $CD14^+$ monocytes. The expression of these genes was diminished in later relapse stages R2 and R3, and R3 instead uniquely exhibited elevated expression of interferon-related genes (*IRF1, IFITH1/2/3).* In summary, we find distinct inflammatory pathways active at different disease stages in AML-08.

Longitudinal profiling of AML-07 provided a unique opportunity to characterize phenotype changes of sequential IDH1m and IDH2m-directed treatment in a patient with co-existing *IDH1* (VAF 25%) and *IDH2* (VAF 12%) mutations at diagnosis. AML–07 relapsed under Ivosidenib treatment (R1: 6% blasts), and exhibited further disease progression on Enasidenib (R2: 66% blasts). At this R2 time point, *IDH2* was not detected, while *IDH1* was retained at high VAF (30%), together with the newly acquired *KRAS* mutation (VAF 31%). Consistent with progression, we found that the MPP compartment was profoundly expanded at R2 (**Figure 3-10A, B**). This was accompanied by elevated expression of stemness-associated genes including *HOXA9*, *CD99* throughout myeloid lineages (**Figure 3-10C**). Additionally, we found that genes such as *LGALS1*, *CPT1A,* involved in fatty acid metabolism and mitochondrial complex I (*NDUFA4, NDUFC2, NDUFB4/7*) and IV (*COX5A, COX6B1, COX5B, COX7C*) of oxidative phosphorylation (OxPHOS) were upregulated in R1 and further elevated in R2 compared to diagnosis, suggesting the potential role of mitochondria metabolism in resistance to mIDH inhibitors (**Figure 3-10C**). In addition, GMP-like cells and $CD14^+$ monocytes at relapse also largely shared the dysregulated genes involved in stemness and mitochondria metabolism with the MPP compartment (**Figure 3-10C**). Interestingly, while we found AML-10 relapsed with a majority of monocytic-like blasts (**Figure 3-10D, E**), they shared the upregulated pathways

with relapse MPPs in AML-07, including stemness, fatty acid metabolism and OxPHOS (**Figure 3-9H**, **Figure 3-10F**).

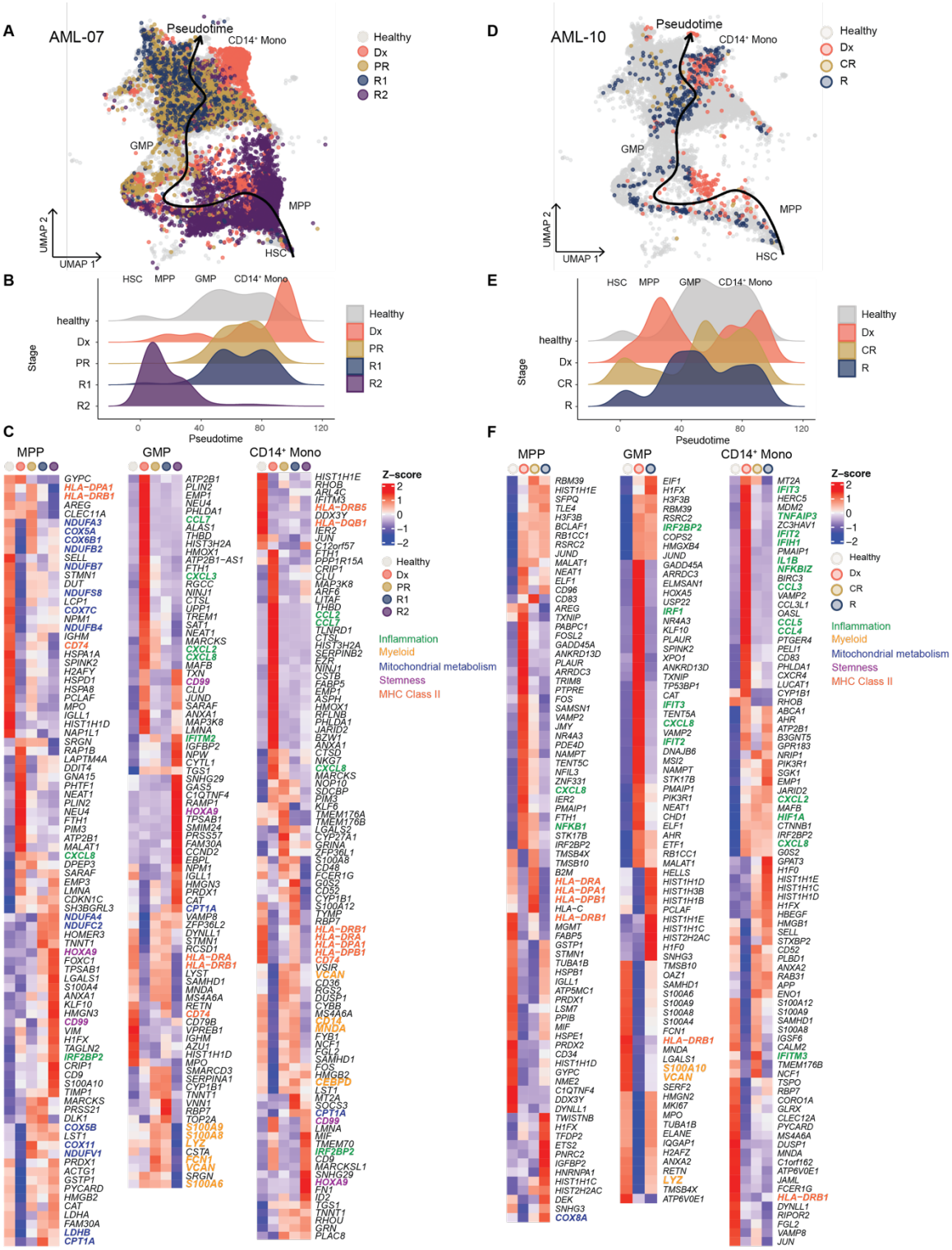


Figure 3-10 Transcriptomic and cellular composition shifts in serial samples under IDHi monotherapy

A) UMAP of pseudotime projecting differentiation hierarchy during longitudinal treatment stages of AML-07.
B) Pseudotime projecting differentiation hierarchy during longitudinal treatment stages of AML-07.
C) Heatmap showing DEGs in each myeloid subset of AML-07 across different stages of treatment.
D) UMAP of pseudotime projecting differentiation hierarchy during longitudinal treatment stages of AML-10.
E) Pseudotime projecting differentiation hierarchy during longitudinal treatment stages of AML-10.
F) Heatmap showing DEGs in each myeloid subset of AML-10 across different stages of treatment.

Thus far, we identified heterogeneous genetic and transcriptomic patterns of relapse to IDH inhibitors, which highlights the need to dissect treatment response and relapse in a patient-specific manner and in the context of their clinical progress. However, we next interrogated whether any transcriptional changes associated with relapse were shared by all patients. MPPs derived from all patients exhibited reduced expression of major histocompatibility complex (MHC) class II (*HLA-DRA, HLA-DRB1)* and *CD74*, MHC II-associated protein invariant chain, in relapse (**Figure 3-9H, E**, **Figure 3-10C**, **F**). The distinctive transcriptomic shifts occurring during therapy response and relapse, including upregulation of inflammation pathways, gain of stemness, and downregulation of antigen presentation highlights the need for further investigation into the role of each of these pathways in disease relapse.

IDH2-mutant Clonal Hematopoiesis Exhibits Dysregulation of *IDH2*-associated Pathways

IDH2 mutations are rare in CH, compared to more frequent CH mutations such as *DNMT3A* or *TET2*, possibly due to high risk of and rapid transformation to overt leukemia (Desai et al. 2018; Abelson et al. 2018; Young et al. 2019). However, we were able to identify two individuals with *IDH2*-mutant CH. Because these individuals did not exhibit hematologic abnormality that would meet criteria for bone marrow assessment, only peripheral blood samples were readily available for bulk targeted DNA sequencing and single cell gene expression profiling.

Clinical sequencing of these samples reported VAFs for *IDH2* and *SRSF2* at 16-17% (Supp Table 1). To determine the VAF of *IDH2* and *SRSF2* within the myeloid compartment, we sequenced lymphoid-depleted cells on a targeted panel of myeloid genes. In CH-01, *IDH2* VAF was 45%, reflecting that IDH2 mutant cells are clonal in this population (Supp Table 1). In CH-02, the VAF for *IDH2* and *SRSF2* were 34% and 29%, respectively. This demonstrates that the VAFs reported by clinical sequencing of unsorted samples underestimates the prevalence of mutant cells in the myeloid lineage, with potential implications for assessing risk of AML transformation or non-hematologic sequelae.

To investigate the impact of *IDH2* mutations at CH, we profiled lymphoid-depleted peripheral blood derived mononuclear cells from two *IDH2* CH donors (**Figure 3-11A)**. The samples are predominantly comprised of $CD14^+$ monocytes, $CD16^+$ monocytes, and dendritic cells (**Figure 3-11B, C**).

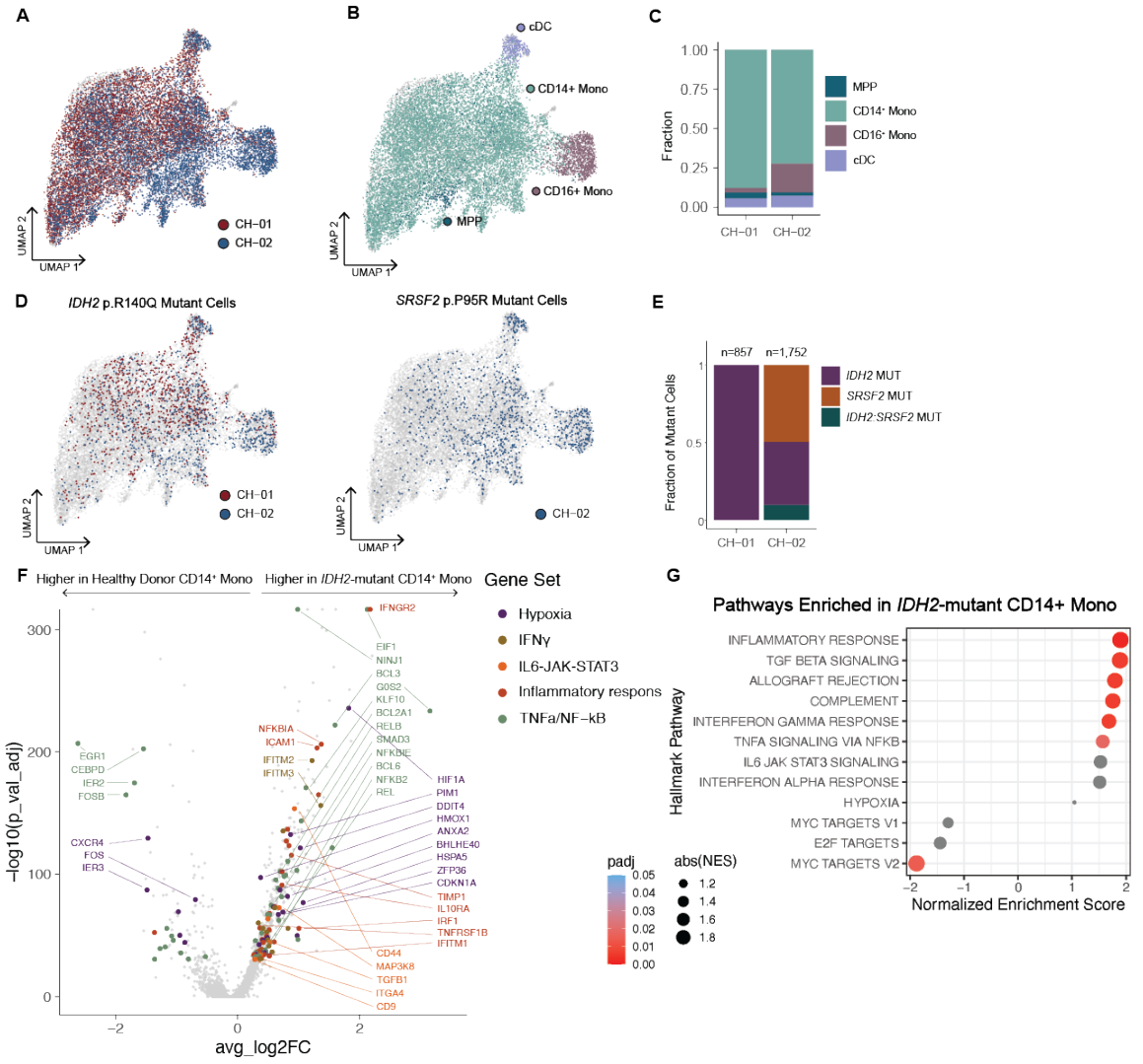


Figure 3-11 Joint single cell gene expression profiling and genotyping in IDH2 mutant CH reveals upregulated inflammation and hypoxia pathways

A) UMAP of 16,288 cells from two *IDH2*-mutant CH donors colored by donor identifier.
B) UMAP as in a) colored by annotated cell type.
C) Stacked barplot of cell type composition of two *IDH2*-mutant CH samples.
D) Cells with mutation in *IDH2* (left) and *SRSF2* (right) are colored by donor identifier. Cells in gray did not have mutant transcripts detected.
E) Stacked barplot of cells mutant for *IDH2* and *SRSF2* per donor. Number of total mutant cells annotated in the plot.
F) Volcano plot of genes enriched (right) or depleted (left) in *IDH2*-mutant $CD14^+$ monocytes from CH donors vs $CD14^+$ monocytes from unaffected donors.
G) Enriched Molecular Signatures Database Hallmark gene sets in *IDH2*-mutant CH versus unaffected donor cells within each of the hematopoietic cell types.

Single cell genotyping (GoT) was deployed on these samples to identify mutant cells and reconstruct the phylogeny of co-occurring mutations. *IDH2* p.R140Q was genotyped in both CH-01 and CH-02 and 2,186 mutant cells were identified, distributed throughout the gene expression-derived clusters (**Figure 3-11D**). CH-02 had a co-occuring *SRSF2* p.P95H mutation and by single cell genotyping it was confirmed that the *IDH2* and *SRSF2* mutations occur in the same cells (**Figure 3-11E**). This demonstrates that stem or progenitor cells harboring *IDH2* mutation in isolation or in combination with *SRSF2* mutation are able to differentiate into mature myeloid cells at CH. This is consistent with murine models of mutant *IDH2* which exhibit impaired but not abrogated myeloid differentiation (Figueroa et al. 2010; Chen et al. 2013; Mylonas et al. 2014).

To identify the transcriptomic consequences of the acquired *IDH2* mutations on monocytes in CH, differential gene expression analysis was performed between the mutated monocytes and monocytes from the normal reference samples (**Figure 3-11F**). Compared to healthy donor-derived monocytes, *IDH2*-mutant monocytes exhibited upregulation of upregulation of the hypoxia (*HIF1A, HMOX1*), Inflammatory Response (*TIMP1, IRF1, ICAM1, IFNGR2*), and Interferon Gamma Response (*IFITM2/3*), IL6-JAK-STAT3 (*CD44, MAP3K8, TGFB1)* and TNFA signaling via NFKB (*NFKB2, NFKBIE, REL, NFKBIA*, *RELB*) (**Figure 3-11F, G**). This data supports the paradigm whereby pro-inflammatory mutant mature myeloid cells facilitate clonal expansion in CH (Avagyan et al. 2021; Yeaton et al. 2022) and may ultimately contribute to cardiovascular and other inflammation-associated disease (Sano et al. 2018). Importantly, many of these pathways are shared with *IDH1/2* mutant AML-derived monocytes at diagnosis (**Figure 3-11I**), suggesting that aspects of the mutant IDH pathogenic phenotype are already present at CH.

Discussion

While characterization of the molecular landscape of AML has enabled the development of novel targeted agents for recurrent genomic alterations, patient responses to these agents are highly variable and many patients develop resistance. Overcoming this challenge is complicated in part by complex intra-tumoral heterogeneity as well as clonal evolution under the selective pressures of treatment. To address the unique challenges posed by the clonal architecture in AML, we established a framework to delineate clone-specific transcriptomic phenotypes and delivered a detailed characterization of *IDH1/2* mutant AML phylogeny at diagnosis, remission, and relapse.

Our analysis of diagnosis samples demonstrate a stem-like, quiescent, and inflammatory phenotype conferred by *IDH1/2* mutations in AML, corroborating similar observations in AML preclinical models (Gruber et al. 2022; Sasaki, Knobbe, Munger, et al. 2012; Zhang et al. 2020), as well as in models of IDH-mutant glioma (Bralten et al. 2011; Sasaki, Knobbe, Itsumi, et al. 2012; Fu et al. 2015). Specifically, this provides validation of previous *in vitro* studies reporting mIDH1/2- or R-2-HG-mediated downregulation of proliferation (Su et al. 2018; Bralten et al. 2011; Gruber et al. 2022) and MYC (Su et al. 2018; P. Zeng et al. 2022) in primary patient samples at a single cell resolution. We corroborated previous reports of DNA repair deficiency and consequently, efficacy of PARP inhibitors in *IDH1/2* AML pre-clinical models (Sulkowski et al. 2017; Gbyli et al. 2022). This analysis supports the relevance of these pathways in primary patient samples.

To further understand how the most frequent co-mutations in *IDH1/2* AML modify this phenotype, we established an integrative approach to assign AML cells to phylogenetic subclones and interrogate clone-specific transcriptional programs. We observed that mutations in *SRSF2* and *NPM1* modulate the AML blast phenotype, towards erythroid (*SRSF2*) or GMP-like (*NPM1*). Subclones defined by mutations in *NRAS* or *SRSF2* exhibited elevated proliferation and metabolic pathways, in contrast to *IDH1/2*-mutant parent clones which upregulated stemness and inflammation. The highly proliferative, and thus likely chemosensitive, phenotype of the *NRAS*-mutant subclone provides a potential explanation for the observation that NRAS-mutations were no longer associated with non-response when IDH1/2 inhibitors were combined with chemotherapy(Amatangelo et al. 2017; DiNardo et al. 2018; Ball et al. 2021; Stein et al. 2021). Furthermore, we find that in the context of *NPM1* and *IDH1* co-mutation, AML blasts upregulate MLL-target genes, supporting the relevance of Menin inhibition in *IDH/NPM1* co-mutated AML(Fiskus et al. 2022). Indeed, there is promising clinical evidence of the synergistic combination of IDHi and Menin inhibitors (Yoon et al. 2022) that warrant further investigation in larger well-powered studies.

We next profiled serial samples from patients treated with IDHi monotherapy from diagnosis to remission to relapse. In all cases, relapse was accompanied by the acquisition of new mutations, including in the RTK-*RAS* pathway, which has been previously implicated as driving resistance (Feng Wang et al. 2021; Amatangelo et al. 2017; Choe et al. 2020). Our analysis of transcriptomic patterns of response nominates several pathways for further investigation including stemness, mitochondrial metabolism, inflammation, and immune evasion by downregulation of MHC Class II. Our finding of increasing stemness at relapse builds upon growing evidence for the role of stemness in mediating resistance to IDHi (Feng Wang et al. 2021) as well as other therapeutic agents.

Furthermore, our analysis identified higher levels of fatty acid metabolism and OxPHOS upon disease relapse, adding evidence to earlier reports (Stuani, et al 2021) to provide more rationale for investigating mitochondria metabolism as a therapeutic vulnerability. We also identified inflammation-associated pathways upregulated at relapse, pointing to a potential role for inflammation in mediating relapse. This adds to a growing body of evidence from our group and others that inflammation may contribute to AML initiation and progression (Avagyan et al. 2021; Yeaton et al. 2022; Lasry et al. 2022). A striking feature shared by all three relapse patients was the downregulation of MHC Class II at relapse, pointing to immune evasion of AML as a potential relapse mechanism. MHC Class II downregulation has been previously reported in other contexts, such as AML relapse after allogeneic transplantation (Christopher et al. 2018; Toffalori et al. 2019; Shahid et al. 2022). Together, these findings motivate and guide future efforts to define the functions of immunomodulatory AML populations, and ultimately, therapeutic interventions to modify their activity.

Finally, we investigated the role of *IDH2* mutations at the preleukemic stage of CH using joint transcriptomic profiling and genotyping. Our analysis revealed elevated inflammation at CH, consistent with reports in other more well-studied CH genotypes (*DNMT3A*, *TET2*)(Jaiswal and Ebert 2019). Because *IDH1/2* mutations confer high risk of AML transformation and there exist therapeutic interventions with reasonable side effect profiles, IDH1/2 inhibitors are currently being investigated in patients with *IDH1/2*-mutated clonal cytopenia of unknown significance (NCT05030441, NCT05102370 (“Ivosidenib for Patients With Clonal Cytopenia of Undetermined Significance and Mutations in IDH1” n.d., “A Study of Enasidenib in People With Clonal Cytopenia of Undetermined Significance” n.d.)). Importantly, we show that targeted bulk DNA sequencing in unsorted mononuclear samples underestimates the expansion of *IDH1/2*-mutant cells within the myeloid

compartment. Our methods and findings provide the foundation to identify biomarkers and assess pathogenic consequences of *IDH1/2* mutations at CH in a larger patient cohort. This will support detailed correlative studies under IDHi therapy to track responses and identify additional therapeutic vulnerabilities in patients who may ultimately transform to AML.

In summary, we leveraged innovative single cell transcriptomic and genotyping methodologies to characterize the phenotypic consequences of somatic mutations in *IDH1/2*-mutant AML, a high-risk subtype. We delineate the transcriptional heterogeneity of diagnosis *IDH1/2*-mutant AML and the striking impact of cooperating mutations in shaping blast phenotypes. Our study highlights the role of inflammation at CH and AML at multiple stages from diagnosis to relapse. Our findings can guide development of therapeutic strategies to achieve better responses in *IDH1/2*-mutant AML and address mechanisms of resistance that make this disease so challenging to treat.

Methods

Meta-analysis of large genomic cohorts of AML patients

Genomic data from two large cohort studies of AML patients (n=3,653) with comparable molecular profiling was integrated (Tazi et al. 2022; Papaemmanuil et al. 2016). We performed a systematic evaluation of all potential pairwise interactions between gene mutations and/or cytogenetic abnormalities (**Figure 3-1A**) using methods previously described in Papaemmanuil et al. 2016. The analysis was restricted to mutations/events occurring at least 8 times in the dataset. Gene-pairs were said to co-occur if both genes were mutated in more patients than expected by random chance (Odds Ratio > 1).

Conversely, gene-pairs were classified as mutually exclusive if they were mutated in fewer patients than expected by chance (Odds Ratio < 1). Fisher's exact tests were performed for each interaction pair and adjustment for multiple hypothesis testing was performed using the Benjamini Hochberg approach to identify statistically significant interactions.

Mutational ordering (**Figure 3-1B, C**) was performed as follows. The analysis was performed within patients with *IDH1* (n= 247) or *IDH2* (n=453) mutations and restricted to genes with more than 15 occurrences within the dataset. Gene-pairs found to be statistically significantly mutually exclusive with *IDH1* or *IDH2*, such as *TET2,* (Odds Ratio < 1) were excluded from the analysis. Variant allele fraction estimates were adjusted for local ploidy status and 95% confidence intervals were drawn for the variant allele fraction using sequencing depth, under a bimodal distribution. Variant allele fraction of gene mutations mapping on the X chromosome of male patients were adjusted to account for a single copy. Within each sample, a Fisher test with cut off $p< 0.01$ was performed to test whether mutations occurred in the same clone (null hypothesis) or whether they were in separate clones. From the set of genes in which 10 or more precedences were observed, Bradley-Terry models were applied using maximum likelihood to the observed precedences using the R package BradleyTerry2 version 1.1-2.

Cohort selection and sample acquisition

To study the role of co-mutations in shaping disease presentation and treatment response in *IDH1/2*-mutant AML to either conventional cytotoxic chemotherapy and/or IDH-directed inhibitors +/- hypomethylating agents, we screened the MSKCC and OSU biobanks for samples from patients receiving targeted IDH1/2 inhibitors and prioritized samples on the basis of availability of sequential samples, including matched diagnostic and post-

treatment samples. In total, we obtained 43 cryopreserved bone marrow aspirate or peripheral blood samples from 15 patients at longitudinal timepoints beginning at diagnosis.

We profiled 6 patients (median age = 58 years) on clinical trial NCT02632708 for the treatment of newly diagnosed *IDH1/2*-mutant AML with IDH1/2 inhibitors in combination with standard of care chemotherapy (Stein et al. 2021). Specimens were collected under MSKCC IRB protocol 16-012. Post-treatment samples were selected at ~4 weeks after treatment started. For one patient with persistent disease, AML-04, a third sample was profiled at 12 weeks after consolidation.

We also selected 9 newly diagnosed patients (median age = 70 years) receiving IDH1/2 inhibitor monotherapy (n=5) or in combination with Azacytidine (n=2) or chemotherapy alone (n=2). Serial samples were collected at remission (n= 9 patients, median time after diagnosis (41 days)) and relapse timepoints (n=4 patients; median time 985 days).
Healthy control specimens were collected from unaffected individuals undergoing hip-replacement surgery (n=4) or bought from STEMCELL Technologies (n=6) (catalog# 70001)

Sample Preparation and Single Cell RNA-seq

Samples from AML-01, AML-02, AML-03, AML-04, AML-05, AML-06, CH-01, CH-02 were processed as follows. Frozen bone marrow and peripheral blood samples were thawed under standard procedures and prepared for sorting by flow cytometry. Briefly, cryopreserved samples were thawed and transferred to 15 ml conical tubes containing 10 ml of RPMI media with 10% fetal bovine serum (FBS) and 0.1mL of a 1 mg/mL DNAse I solution. Cells were centrifuged at 300 x g for 4 minutes at 4℃, and supernatant was

removed. Cells were resuspended in FACS buffer and incubated with Fc blocker for 15 minutes. Next, cells were stained with PE-conjugated CD45, and FITC-conjugated CD3, CD19, and CD56 to enable lymphoid-depleted sorting. Cells were stained for 15 minutes and resuspended in FACS-DAPI to enable enrichment for live cells. Cells were then sorted for DAPI-, CD3-/CD19-/CD56- cells using the BD Influx at the Weill Cornell Medicine flow cytometry core. Sorted cells were processed using the 10X Genomics Chromium controller with the Single Cell Gene Expression 5' V1 protocol (except sample CH-02 which was processed on 5' V2 after V1 was discontinued). The 5' kit was selected to maximize genotyping efficiency because most mutations, including IDH1, IDH2, NRAS, and SRSF2, were closer to the 5' end of the mRNA. Single cell gene expression libraries were sequenced on a Nova-seq 6000 with 150:8:16:150 paired end format.

Sample Preparation and CITE-seq

Samples from AML-07, AML-08, AML-09, AML-10, AML-11, AML-12, AML-13, AML-14, AML-15 were processed as follows. Frozen bone marrow samples were thawed and transferred into 50ml conical tubes containing 1x PBS +2% fetal bovine serum(FBS). Cell suspensions were centrifuged at 300 x g for 5 minutes at 4℃, and supernatant was removed. Samples were then enriched for >90% alive cells by using a dead cell removal kit, or stained with DAPI before sorted for alive cells using SY3200 Cell Sorter (SONY). For cell sorting, DAPI low cells are collected into 15ml conical tubes containing 3ml FBS on ice. Samples were then centrifuged at 300 x g for 5 mins at 4℃.

For CITE-seq, enriched alive cells were first stained with cell-hashing oligo-tagged antibodies(Biolegend) according to manufacturer's instructions. Then samples were counted and pooled together and stained with a CITE-seq antibody cocktail (Biolegend)

according to the manufacturer's instructions. After staining, samples were processed using 10X Genomics Chromium controller with the Single Cell Gene Expression 3' Reagent Kits (v3 and v3.1). Libraries were sequenced on an Illumina NovaSeq 6000.

Targeted DNA sequencing

For a subset of samples from 8 patients/donors (AML-01, AML-02, AML-03, AML-04, AML-05, AML-06, CH-01, CH-02), sorted cells remaining after the 10X Genomics Single Cell protocol were sequenced on a bulk targeted DNA panel to confirm clinically reported mutations and ascertain the VAF within the lymphoid depleted population. Remaining cells were pelleted, flash frozen, and DNA-extracted. Targeted-DNA libraries were made using a panel of genes recurrently mutated in myeloid malignancies (n=152). The approaches for targeting sequencing, sample quality control, bioinformatics processing, and variant pathogenicity annotation have been previously described in detail (Bernard et al. 2021; Medina-Martínez et al. 2020). In summary, libraries were sequenced at 600x coverage. Putative oncogenic mutations (Variant Allele Fraction > 2%) were annotated, including single nucleotide variants, insertions and deletions. Chromosomal alterations, including arm-level and whole chromosome events, were assessed. These data were used to validate and complement the clinical mutation calls and karyotype and FISH assays. A putative clonal structure was derived for each patient using a proportion test to test whether the proportion of mutant reads over total reads for each mutation were the same, and therefore in the same clone. For each sample this led to a complete catalog of mutations and the anticipated clonal structure.

Single cell genotyping

Single cell genotyping was applied to 8 patients/donors (AML-01, AML-02, AML-03, AML-04, AML-05, AML-06, CH-01, CH-02). Mutations were prioritized for Genotyping of Transcriptomes (GoT) on the basis of being clone-defining (see Targeted DNA sequencing), proximity to the 5' end, and expression levels. GoT libraries were prepared for *IDH1*, *IDH2*, *NRAS*, *NPM1*, and *SRSF2* as described in Nam et al 2019. Primers and number of PCR amplification cycles are provided in Supp Table 2. GoT Libraries were pooled and sequenced on Illumina MiniSeq 300 cycle mid-output kit in 147:8:16:147 paired end configuration. FASTQs were aligned to reference genome GRCh38 using CellRanger 5.0.1. Mutation status for single cells was ascertained by performing a pileup on each mutation while tabulating cell barcodes and UMIs (automatically stored in BAM file by CellRanger) in both the GoT and gene expression libraries (Code provided on github). Cells were assigned Mutant if one or more mutant UMIs was detected. Due to allelic dropout, increased stringency was applied for wildtype cells. Cells were assigned as wildtype if 2 or more UMIs had only wildtype alleles.

Single cell gene expression

FASTQs were aligned to GRCh38 and cell by gene count matrices were generated using CellRanger 5.0.1 with default parameters. Gene expression analysis was conducted using Seurat v4 in R4.0. Analysis scripts will be provided on GitHub. Details for specific analyses are described below.

Inference of copy number state from gene expression data

Inference of chromosomal alterations from single cell gene expression data was performed using a haplotype-aware caller, Numbat (Gao et al. 2022). AML samples with normal karyotype were used as the reference. The clusters output from Numbat were examined and compared against the clinically reported karyotype and FISH reports. Clusters were annotated as bellowing to clones defined by the copy number alterations present in both the Numbat results and clinical reports. No new events were discovered. This enabled mapping of single cells to phylogenetic clones defined by chromosomal alterations.

Clustering and cell type annotation

We first filtered out the cells with more than 20 percent of mitochondrial genes, less than 200 number of genes detected, and greater than 4000 number of genes detected, assuming the dead cells, empty droplets, and multiplets, respectively. For the hashtagged libraries, we additionally identified singlets using the `HTODemux` function implemented in the Seurat package with a default setting. To account the technical batch effects of 3' sequencing, 5' sequencing, and different sequencing depths, we integrated the datasets using Seurat integration. Briefly, after log-normalization of each dataset using `NormalizeData`, we selected the top 2,000 highly variable genes across the datasets using `SelectIntegrationFeatures` and scale the normalized data of the 2,000 genes and regressed out the percentage of mitochondrial genes and ribosomal genes. Next, we used Reciprocal Principal Component Analysis (RPCA) for the integration using the `FindIntegrationAnchors` function with k=10 and `IntegrateData`, followed by PCA analysis and Uniform Manifold Approximation and Projection for Dimension Reduction (UMAP) for dimensionality reduction. Next, we clustered the cells using the Louvain

algorithm based on the k-nearest neighbor graph derived from the integrated data. We used `FindClusters` with resolution = 3, which generates a total 78 clusters. We assessed the marker genes of 78 clusters and identified 7 clusters as the mixture of different biological cell types, potentially multiplets. After filtering out these cells, we performed another iteration of integration described below, confirming that there are no mixed biological cell types as a single unbiased cluster. With various resolutions of the Louvain algorithm, we manually annotated the cell types based on the pre-defined gene markers. For the comparison with non-*IDH1/2* mutated AML samples (Lasry et al, 2022), we used Seurat reference mapping, and transferred the cell type annotations to non-*IDH1/2* mutated AML scRNA-seq data.

Differential gene expression and gene set enrichment analysis

Comparison of AML-derived cell types to healthy control counterparts was performed by first downsampling the number of cells per cell type per patient (250 cells) and then using Seurat's FindMarkers function with log fold change set to 0 and all other default parameters. For the comparison of clonal hematopoiesis samples, we compared *IDH2* mutant CD14+ monocytes from two different donors (total 1,391 cells, 753 cells, 638 cells, respectively) and healthy CD14+ monocytes (total 2,000 cells, each 1,000 cells from two control peripheral blood samples). For the longitudinal analysis comparing healthy, diagnosis, remission, and relapse, we compared two different stages in pair-wise manner Top 20 differentially expressed genes were selected for visualization based on their log fold changes and adjusted p-values from every pairwise comparison. To visualize the gene expression level, we used Z-score standardization across the stages for the selected genes. We used the Wilcoxon rank sum test implemented by `FindMarkers` in the Seurat package throughout the study. Gene set enrichment analysis was performed using the fgsea R package on msigdb's Hallmark Gene Sets (minSize 15, maxSize 500,

nPermSimple 100000). Gene ontology (GO) enrichment analysis was performed using the clusterProfiler R package with minGSSize =3, maxGSSize=800.

Pseudotime analysis

Pseudotime analysis was performed as described in the previous study (Granja et al, 2020). Briefly, since the differentiation order of stem cells to mature cells are preserved in the UMAP coordinates, the mean coordinates of HSC, MPP, GMP, and CD14+ monocytes were used to align single cells across the trajectory. First, the euclidean distance between each cell to the mean coordinate of its corresponding cell type and to the coordinate of further differentiated cell type was computed. Next, the pseudotime vector was computed based on the quantiles of the distance of each cell to the next cell type cluster. Lastly, we used `smooth.spline` to fit a continuous trajectory and aligned all cells to the trajectory using their nearest point to the fitted trajectory.

Gene set scoring analysis

Individual cells were scored with Hallmark Gene Sets from Molecular Signatures Database using the UCell R package (Andreatta and Carmona 2021; Liberzon et al. 2015). For the statistical test between the patient groups, the average score of each gene set was calculated per cell type and Wilcoxon Rank Sum test was performed on the average score. Lineage score was calculated based on the lineage gene modules for HSC-like and GMP-like (van Galen et al. 2019)and MEP(Velten et al. 2017). For the lineage score plot in Figure 2F-G, GMP and MEP scores were scaled 0-1. The x-axis is the difference between the GMP and Ery score (yields negative values for MEP/Ery-like blasts). HSC score was scaled from -0.8 to 0.5. HSC-like cells, based on cell type annotation, scored >0 so the top fraction of stem-like cells was assigned to the stemness/vertical axis and for

visualization purposes, scores (stemness and GMP-MEP) were plotted on a unique axis. Cells were also scored for expression of cell cycle gene modules (Tirosh et al. 2016).

Assignment of single cells to phylogenetic subclones

Assignment of single cells to phylogenetic clones was performed by integrating copy number alteration and mutation genotype status. Copy number alteration-defined clones were defined based on the cell clustering and output of Numbat. Next, the fraction of mutant cells by Genotyping of Transcriptomes per gene expression-derived cluster was tabulated. Gene expression derived clusters were assigned as belonging to the parent clone or subclone based on having predominantly mutant or wildtype/NA composition. Given the clonal variant allele fractions of *IDH1/2* and *DNMT3A* mutations at diagnosis, all cells were assumed to harbor these clonal events. This could be confirmed in cases with clonal copy number alterations where the events were detected in nearly all cells. Differential gene expression analysis between parent clones and subclones were performed by subsetting for MPPs, the majority cell type, to avoid confounding by unbalanced cell type representation per clone. FindMarkers was used with logFC set to 0 and all other default parameters. GSEA was performed as described above.

Gene set scoring analysis and cell type deconvolution from bulk RNA-sequencing data

Bulk RNA-seq from Alliance cohort (Lasry et al. 2022) was used to study IDH1/2 mutant AML in a large cohort. The RNA-seq was log2-transformed and each sample was scored with lineage markers and hallmark gene sets using the UCell R package. Wilcoxon Rank Sum test was used to compare the IDH mutant group and the non-IDH mutant group. To identify cell type fractions based on our single-cell RNA-seq annotations, we used

dampened weighted least squares (DWLS). DWLS first identifies cell type-specific gene expression profiles in the annotated scRNA-seq data set, and then uses the derived signatures to deconvolute cell type fractions from the bulk RNA-seq data using a weighted least squares approach, which improves upon commonly used ordinary least square methods to control for large estimation errors in rare cell types and improves the contribution of lowly expressed genes with highly differential expression between cell types.

Table 3-1 Supplementary Table 1: Cohort Characteristics

Patient ID	Sequencing	Gender	Age	Stage	Blast	Treatment
AML-01	scRNA-seq+GoT	M	61	Dx	53%	IDHi&7+3
				CR	24%	
AML-02	scRNA-seq+GoT	F	55	Dx	63%	IDHi&7+3
				CR	0.05%	
AML-03	scRNA-seq+GoT	F	25	Dx	60%	IDHi&7+3
				CR	4%	
AML-04	scRNA-seq+GoT	F	62	Dx	46%	IDHi&7+3
				CR	0.27%	
				R1	2%	
AML-05	scRNA-seq+GoT	F	65	Dx	33%	IDHi&7+3
				CR	2%	
AML-06	scRNA-seq+GoT	M	45	Dx	60%	IDHi&7+3
				CR	3%	
AML-07	CITE-seq	F	70y	Dx	68%	IDH1i
				MRD	4%	
				CR		
				Relapse1	6%	IDH2i
				Relapse2	66%	
AML-08	CITE-seq	F	71y	Dx	36%	IDH1i
				CR, cycle 8	0%	
				CR, cycle 23	0%	
				CR, cycle 33	0%	
				elapse, cycle 3	3%	Secondary IDHi+Aza
				Relapse	7%	FT-2102+Aza
				P1	35%	LP-108
				P2	56%	APV0436
				P3	78%	
AML-09	CITE-seq	M	70y	Dx	60%	IDH1i
				CR	CR per cytology	
				CR	CR post allo PBSCT	
AML-10	CITE-seq	F	70yr	Dx	47%	IDH2i
				MRD	1%	BMT
				CR	0%	
				Relapse	2%	
AML-11	CITE-seq	M	73yr	Dx	N/A	IDH2i
				PR	30%	
				CR	CR per cytology	
AML-12	CITE-seq	M	78yr	Dx	85%	IDHi+Vidaza
				CR	0%	
				CR	0%	
AML-13	CITE-seq	M	80yr	Dx	35%	IDHi+Aza
				PR	9%	
				CR	0%	
				Relapse	N/A	
				Relapse	N/A	
AML-14	CITE-seq	M	64yr	Dx	36%	7+3
				CR	0%	
				Relapse	N/A	
AML-15	CITE-seq	F	49yr	Dx		7+3
				CR		
				Relapse	35%	
CH-01	scRNA-seq+GoT					
CH-02	scRNA-seq+GoT					

(continued) iii

Patient ID	Mutations	Cytogenetic
AML-01	IDH1(0.42),DNMT3A(0.19),NPM1(0.43)	
	IDH1(0.51),DNMT3A(0.41),NPM1(0.04)	
AML-02	IDH1(0.3),DNMT3A(0.32),NRAS(0.1)	47,XX,+1,der(1;13)(q10;q10),+8[18]; 49,idem,+6,+10[3] 46,XX[1]
	IDH1(0.37),DNMT3A(0.45),NRAS(ND)	
AML-03	IDH1(0.461),NPM1(0.478),KRAS(0.024)	
AML-04	IDH2(0.394), DNMT3A(0.383),SRSF2(0.329),KRAS(0.017)	
	IDH2(0.45), DNMT3A(0.41),SRSF2(0.22)	
AML-05	IDH2(0.5),DNMT3A(0.467),SRSF2(0.502),RUNX1(0.434)	47,XX,+8[10]; 48,XX,+8,+8[2]; 47,XX,+14[10]; 46,XX[2]
	IDH2(0.48),DNMT3A(ND),SRSF2(ND),RUNX1(ND)	
AML-06	IDH2(0.48),DNMT3A(0.503), CEBPA(0.453),SRSF2(0.496)	
	IDH2(0.09),DNMT3A(0,11), CEBPA(ND),SRSF2(0.06)	
AML-07	ASXL1 (0.32),DNMT3A (0.44),IDH1(0.25),IDH2 (0.12)	47,XX,del(5)(q22q35),+8[20]
	N/A	47,XX,del(5)(q22q35),+8 [20,two w/nonclonal abnormalities]
	N/A	47,XX,+8[11]/47,sl,del(5)(q22q35)[3]
	ASXL(0.29),DNMT3A(0.38),IDH1(0.30),KRAS(0.31)	47,XX,+8[6]/47,sl,del(5)(q22q35)[7]/46,XX[5]/
AML-08	IDH1(0.25)	46,XX[20]
	NGS done, no mutation detected	
	N/A	
	N/A	46,XX[19,one is 4n]/nonclonal[1]
	DNMT3A(0.12), GNAS(0.06), IDH1(0.10)	
AML-09	IDH1(0.38), NPM1(0.4),FLT-ITD(0.21)	46,XY[18]/nonclonal[2]
	IDH1 (0.21), NPM1(0.2)	
AML-10	DNMT3A(0.46),IDH2(0.43),KRAS(0.01),NPM1(0.34),NRAS(0.23),FLT-ITD(0.03)	N/A
	DNMT3A(0.32),IDH2(0.04),NPM1(0.03),FLT-ITD(0.05)	N/A
	DNMT3A(0.02),FLT-ITD(0.04),IDH2(0.02),NPM1(0.02)	N/A
AML-11	ASXL1(0.5), CEBPA (0.16), IDH2 (0.48), SRSF2 (0.54),KRAS(0.47)	46,XY[20]
AML-12	IDH2(0.48),NPM1(0.42),SRSF2(0.47)	46,XY (17 metas), -Y in 3 metas
	IDH2(0.17),SRSF2(0.14)	46,XY (17 metas), -Y in 3 metas
	IDH2(0.2),SRSF2(0.22)	All cytogenetically normal
AML-13	DNMT3A(0.34),IDH1(0.37),SMC1A(0.24)	46,XY[20]
	N/A	46,XY,del(7)(q22q32)[4]/46,XY[16]
	DNMT3A(0.14),FLT-ITD(0.06),IDH1(0.14),RUNX1(0.09)	46,XY,del(7)(q22q32)[5]/46,XY[15,one is 4n
	IDH1(N/A)	
AML-14	IDH2(0.46),BCOR(0.99),FLT3TKD(0.06),PTPN11(0.17),U2AF1(0.45)	48,XY,+8,+8,i(8)(p10),t(9;15)(q33;q15),del(10)(q22.1q24),del(13)(q14q21)(cp14)/44-48,idem,add(13)(p11.2)(cp2)
		48,XY,+8,+8,i(8)(p10),t(9;15)(q33;q15),del(10)(q22.1q24),del(13)(q14q21)(1)
		/46,XY(19)
AML-15		47,XX,+8[19]/nonclonal w/clonal abnormalities[1]
		46,XX[18]/nonclonal[2]
	DNMT3AR882 (0.15), IDH2 (0.11)	47,XX,+8[1]/46,XX[19]
CH-01	IDH2(0.458)	
CH-02	IDH2(0.341),SRSF2(0.298)	

(continued) iv

Patient ID	IDH1	IDH2	ASXL1	BCOR	CEBPAdm	DNMT3A	FLT3 TKD	FLT3-ITD	NPM1	NRAS	KRAS	PTPN11	RUNX1	SMC1A	SRSF2	U2AF1
AML-01	0.42					0.19			0.43							
	0.51					0.41			0.04							
AML-02	0.3					0.32				0.1						
	0.37					0.45				ND						
AML-03	0.461								0.478		0.024					
AML-04		0.394				0.383					0.017				0.329	
		0.45				0.14									0.22	
AML-05		0.5				0.467							0.434		0.502	
		0.48				ND									ND	
AML-06		0.48			0.453	0.503									0.496	
		0.09			ND	0.11									0.06	
AML-07	0.25	0.12	0.32			0.44										
	0.3		0.29			0.38					0.31					
AML-08	0.25															
AML-09	0.38							0.21	0.4							
	0.21								0.2							
AML-10		0.43				0.46		0.03	0.34	0.23						
		0.04				0.32		0.05	0.03							
		0.02				0.02		0.04	0.02							
AML-11		0.48	0.5		0.05,0.05,0.06						0.47				0.54	
AML-12		0.48							0.42						0.47	
		0.17													0.14	
		0.2													0.22	
AML-13	0.37					0.34								0.24		
	0.14					0.14		0.06					0.09			
AML-14		0.46		0.99			0.06					0.17				0.45
AML-15																
		0.11				0.15										
CH-01		0.458														
CH-02		0.341													0.298	

Chapter 4: Perspectives and Future Directions

Summary

This thesis leveraged integrated genomic technologies to characterize *UBA1* mutations in MDS and recurrent *IDH1/2* clonal phylogenies in AML.

In the *UBA1*-mutated MDS study, we find that *UBA1* p.M41/T/V/L hotspot mutations are present in 7% of male MDS patients who either have no molecular disease-defining alterations or are unclassifiable by WHO 2016. Among *UBA1*-mutated patients, those with high *UBA1* VAF (>5%) were the most likely to have inflammatory clinical presentation associated with VEXAS syndrome. This directly informs clinical implementation of *UBA1*-mutation diagnostic screening in MDS. We identify frequently co-occurring mutations in *TET2* and *DNMT3A* and find distinct patterns of clonal architecture for these co-mutations that point to differential mechanisms of selection. This informs future mechanistic studies of *UBA1*, *TET2*, and *DNMT3A* interactions.

In the IDH1/2 AML study, we 1) define recurrent co-mutation patterns in *IDH1/2* mutant AML, 2) develop a framework for phylogenetic clone-specific transcriptomic phenotyping in primary *IDH1/2* mutant AML patient samples to investigate how these co-mutations shape the blast phenotype at diagnosis, 3) characterize transcriptomic patterns of diagnosis, remission, and relapse samples from patients treated with *IDH1/2* inhibitors, and 4) identify *IDH2*-associated gene expression programs dysregulated at the pre-leukemic stage of CH.

This thesis highlights the importance of incorporating learnings from population genomics studies and individual patient clonal architecture to understand disease phenotypes. The frameworks developed to characterize the impact of *IDH1/2* and other frequently co-occurring mutations in AML in Chapter 3 of this thesis can be applied across cancer subtypes. In particular, this thesis identifies an opportunity to study the role of *UBA1* and co-occurring mutations such as *TET2* and *DNMT3A* in VEXAS and MDS.

Advances in single cell genomics technologies will enable clone-specific phenotyping within more complex clonal architectures

In the *IDH1/2* AML study, the distinct clonal phenotypes exhibited by the *IDH1/2* mutant parent clones and *NRAS*- *NPM1*- and *SRSF2*-mutant subclones highlight the importance of incorporating clonal architecture into the analysis of primary patient samples. Most single cell gene expression profiling studies to date in AML have integrated samples with heterogeneous genomic alterations and complex clonal phylogenies. Thus, clone-specific transcriptomic programs, which can be relevant for identification of biomarkers or therapeutic vulnerabilities, are obscured.

The most significant technical limitation of the *IDH1/2* AML study is that currently, single cell gene expression technologies capture a small fraction of the mRNA within a particular cell (10-20%, depending on cell type and technology platform)(G. X. Y. Zheng et al. 2017). This poses challenges for the analysis of rare cell populations as differential gene

expression requires pseudo-bulk analysis of hundreds of transcriptionally similar cells. Furthermore, the genotyping efficiency of somatic mutations was limited by the level of gene expression. Many somatic mutations occur in genes which are quite lowly expressed relative to genes involved in cellular processes such as cell cycle or differentiation. To overcome these limitations, there are promising ongoing efforts to leverage the additional information from the clonal evolution in mitochondrial variants for lineage tracing which will improve the confidence of single cell clonal assignment (J. Xu et al. 2019; Ludwig et al. 2019; Velten et al. 2021). Another approach is to derive mutation status from DNA, rather than mRNA. Emerging multi-modal technologies which include profiling epigenetic regulation (Gaiti et al. 2019) and chromatin accessibility (Myers et al. 2022) simultaneously with genotyping will further expand our understanding of how somatic mutations dysregulate hematopoiesis and drive leukemogenesis. The genotyping efficiency from the single cell chromatin accessibility assay is not limited by gene expression, expanding the catalog of mutations amenable to genotype-phenotype profiling. Furthermore, technical advances increasing the number of cells profiled per sample, the fraction of transcripts captured per cell and genotyping efficiency will facilitate phenotyping of rare subclonal populations. This will be relevant for tracking the emergence of treatment resistance-associated or measurable residual disease clones.

The genotype-phenotype single cell technologies have been predominantly applied to hematologic malignancies which have a median of 3 driver mutations in myeloid neoplasia-related genes(Gaiti et al. 2019; Nam et al. 2019; Petti et al. 2019; Rodriguez-Meira et al. 2019; Miles et al. 2020; Demaree et al. 2021; Myers et al. 2022). As the technical specifications and performance of these platforms improve, it will be possible to apply these frameworks to other cancers with higher genomic instability and mutational burden, and more complex clonal architecture(Kandoth et al. 2013).

The in-depth characterization of primary patient samples identifies patient-specific patterns of treatment response and resistance in *IDH1/2*-mutant AML

In this study, as well as other published studies, patient responses to *IDH1/2* inhibitors are highly variable. Published primary mechanisms associated with non-response include leukemia stemness (Feng Wang et al. 2021) and mutations in Receptor Tyrosine Kinase (RTK) pathway genes (Amatangelo et al. 2017; DiNardo et al. 2018). Secondary relapse-associated mechanisms include clonal evolution (Quek et al. 2018), especially acquisition of RTK pathway mutations particularly *KRAS* which is mutated less frequently in primary resistance(Choe et al. 2020), acquisition of second-site mutations precluding drug binding(Intlekofer et al. 2018), isoform switching (ie. cytosolic *IDH1* to mitochondrial *IDH2*)(Harding et al. 2018), and changes in mitochondrial metabolism (Stuani et al. 2021).

In the *IDH1/2* AML study, the three patients who relapsed after IDHi therapy had concomitant acquisition of additional mutations, including in *KRAS*. Single cell gene expression profiling of patients' serial samples from diagnosis to remission to relapse identifies transcriptomic patterns associated with these different disease stages. While many of these shifts were patient specific, such as upregulated mitochondrial metabolism, all three patients shared downregulation of Major Histocompatibility Complex (MHC) Class II components at relapse. As immune escape mediated by MHC Class II downregulation has been previously reported in AML outside of the *IDH1/2* inhibitor context (Christopher et al. 2018; Toffalori et al. 2019; Shahid et al. 2022), this is unlikely to be a relapse mechanism specific to *IDH1/2* inhibitors. From comparison of the different co-mutation

associated transcriptomic programs, we find that *SRSF2* co-mutated patients expressed the highest levels of MHC-II components at diagnosis compared to *IDH1/2*, *NPM1* and *NRAS* mutated clones, consistent with previous reports that *IDH1/2*- and *NPM1*-mutant AML downregulate MHC Class II at diagnosis(Dufva et al. 2020). The role of MHC-II in immunosurveillance of mutant HSPCs has recently been described in the context of AML disease initiation (Hernández-Malmierca et al. 2022). It will be important to extend this analysis to determine the immunogenicity of different CH and AML founding mutations and investigate how the activity of these immune surveillance mechanisms shifts at different stages from CH to AML transformation, then under therapy and at relapse. Further studies that incorporate close examination of the T cell populations and functional validation are needed to investigate this potential mechanism.

Single-cell analyses of *IDH*-mutant CH and AML to understand high risk of *IDH* CH transformation

While somatic mutations in genes such as *DNMT3A* and *TET2* are found at high frequency in aging populations, the risk of progression or transformation to an overt hematologic malignancy is relatively low(Jaiswal et al. 2014; Young et al. 2016; Desai et al. 2018; Abelson et al. 2018). Furthermore, it is unclear which factors determine an individual's risk of transformation as hotspot mutations alone are insufficient predictors. In contrast, *IDH* mutations are rarely found at the pre-leukemic stage of CH but when they are found, they confer high risk of transformation to AML. Furthermore, *IDH2* p.R140Q mutations were identified as one of the fittest CH variants (Watson et al. 2020). It is possible that rapid progression of *IDH*-mutant CH to over leukemia explains the low prevalence of *IDH* CH in

the general population. However, it is unclear which mechanisms under this rapid disease progression. The rapid acquisition of additional somatic mutations in *IDH1/2* mutant cells, leading to rapid clonal expansion may explain this phenomenon. This explanation is supported by our observation of a stemlike and non-proliferative phenotype in *IDH1/2* ancestral clones in AML as well as decreased DNA damage repair in *IDH1/2* mutant AML. Our study was limited by availability of only peripheral blood samples from a small number of CH donors and could not definitely address this question. Further study of the mechanisms underlying *IDH* mutant CH progression to MDS or AML is needed. To address this, a longitudinal study profiling *IDH* progression from CH to MDS and/or AML incorporating clone-specific phenotyping frameworks as established here can identify the transcriptomic shifts within clonal populations and the immune microenvironment during disease progression.

In the meantime, IDH inhibitors, which have a reasonable safety profile for long-term use, are being investigated clinically for patients with *IDH1/2*-mutant clonal cytopenia of unknown significance (CCUS) (NCT05102370, NCT05030441). It remains to be seen whether there is a clinical benefit for this approach either in limiting pathological sequelae of *IDH* CH or in preventing transformation. While promising, there is a risk that the same resistance mechanisms observed in AML treated with IDH inhibitors (e.g. isoform switching(Harding et al. 2018), second-site mutations(Intlekofer et al. 2018)) will pose a challenge here.

Inflammation plays a key role in disease initiation, maintenance, and progression across the spectrum of CH, MDS and AML.

Together, this thesis builds evidence for the important role for inflammation in hematologic malignancies, including *IDH1/2*-mutant AML and *UBA1*-mutant MDS. In the *IDH1/2* AML study, *IDH1/2* AML upregulated inflammation-associated pathways and this was often specific to the *IDH1/2*-mutated ancestral clone. Furthermore, relapse under *IDH1/2* inhibitors was accompanied by increased inflammation. Lastly, *IDH2*-mutant CH exhibited upregulation of inflammation pathways. In the *UBA1*-mutant MDS study, we identify frequent co-mutation with *TET2*, a mutation often found in CH as well as overt hematologic malignancies, and associated with increased inflammation(Yeaton et al. 2022).

Previous studies have implicated inflammation in driving disease initiation, progression, and relapse. CH has been associated with elevated inflammation and increased incidence of systemic autoinflammatory diseases (Hecker et al. 2021; Ricard et al. 2020; Marnell, Bick, and Natarajan 2021). Inflammation has been implicated in pre-clinical models of clonal selection in CH due to increased fitness of mutant clones in inflammatory environments(Avagyan et al. 2021). It will be important to understand how (if) the interaction between *UBA1*-mutant and proinflammatory *TET2*-mutant clones may shape MDS progression. Patients with autoimmune diseases have increased risk of MDS(Boddu and Zeidan 2019) and *IDH1/2* and *TET2* mutations specifically have been associated with inflammatory clinical presentation in the context of MDS(Zhao, Boy, et al. 2021). Furthermore, high levels of inflammation have been associated with poor patient outcomes in AML(Lasry et al. 2022). Understanding how inflammation drives disease

initiation, progression, and relapse in hematological malignancies will be the focus of future research.

Co-mutation analysis highlights *DNMT3A* and *TET2* as key interacting partners with *UBA1* mutations, highlighting the need for further investigation in VEXAS/MDS

From the MDS population genomics study, we find different patterns of co-mutations in these genes: *DNMT3A* and *UBA1* are likely acquired within the same cell and lead to rapid expansion of that clone, suggestive of a fitness advantage **(Figure 4-1A)**. In other contexts, such as the frequently observed co-mutation with *TET2*, we observe either low VAF *TET2*/high VAF *UBA1* **(Figure 4-1B)** or the opposite trend **(Figure 4-1C)**. This clonal architecture has also been recently described in a study leveraging single cell DNA-sequencing of *UBA1*- and *TET2*-mutant CH(Fernandez et al. 2022; Patel et al. 2022). Since *TET2* is also a pro-inflammatory driver, it would be important to understand how these two cell populations interact and if there are any cell non-autonomous interactions driving disease progression. Furthermore, investigation of shifts in clonal composition in longitudinal samples correlated with detailed annotation of inflammatory symptoms would elucidate mechanisms of cooperation between these co-occurring mutations.

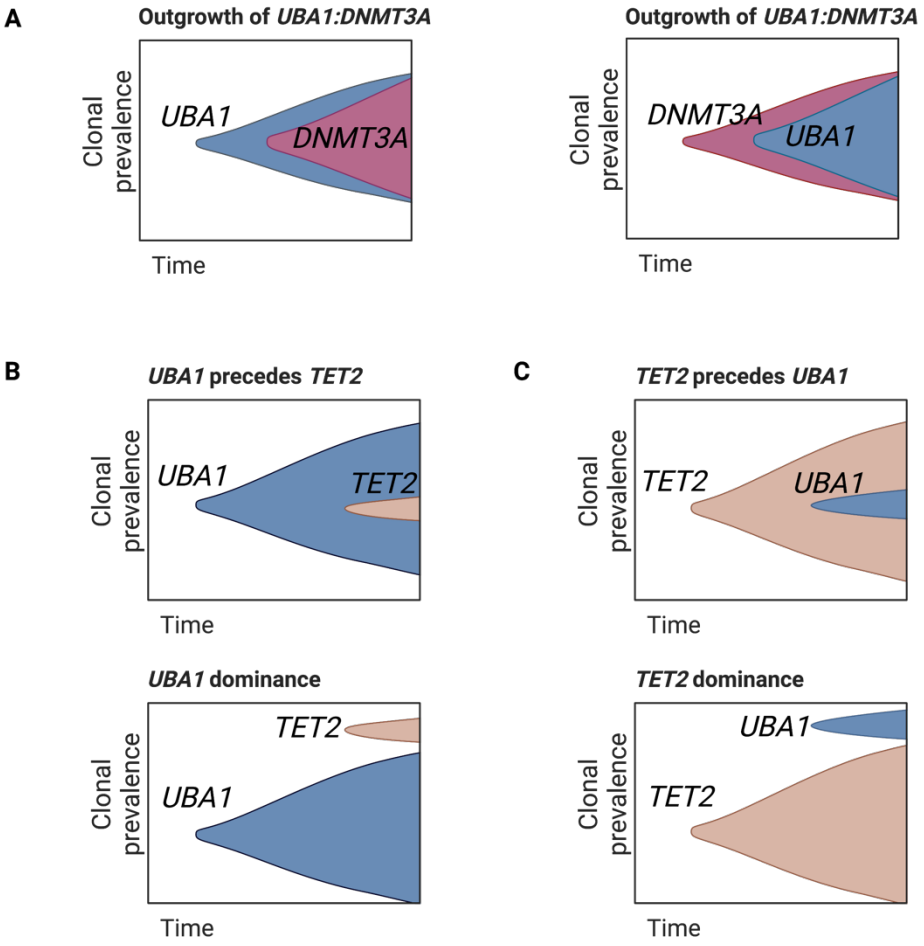


Figure 4-1 Schematics of putative recurrent clonal structures in MDS: UBA1 co-mutation with A) DNMT3A and B,C) TET2.

The methodologies of clonal phenotyping established in the *IDH1/2*-mutant AML study could be applied to study the distinct clonal phenotypes of *UBA1*-mutant vs *TET2*-mutant vs *UBA1-DNMT3A* co-mutated populations. Characterization of the pathobiology of these co-occurring genomic events could enable identification of therapeutic vulnerabilities in *UBA1*-mutant VEXAS and MDS.

On the other hand, from the UBA1 study, we find that in some cases of MDS, *UBA1* is the sole genetic abnormality detected. Given the high incidence of MDS among VEXAS patients (25-55%)(Beck et al. 2020; Bourbon et al. 2021; Georgin-Lavialle et al. 2022), MDS may represent progression of the VEXAS disease. Longitudinal mutation profiling of VEXAS patients will be needed to understand risks of developing MDS.

Integrative clone-specific gene expression analysis may identify therapeutic vulnerabilities in VEXAS/MDS

Due to the inflammatory clinical presentation of VEXAS, patients are often treated with high-dose glucocorticoids, which have an unacceptable safety profile for long-term use and result in symptom relapse if reduced or discontinued. Thus, it is imperative to identify therapeutic strategies that are curative and/or safe for long-term use. Much of the data to date on treatment response in VEXAS is from retrospective analysis of patients with heterogeneous clinical management of associated clinical manifestations of the disease. From retrospective analyses, there is some evidence of the benefit of hypomethylating agents, Janus kinase (JAK) inhibitors, and allogeneic hematopoietic stem cell transplant in VEXAS.

Hypomethylating agents, specifically Azacytadine, are used in the clinical management of high-risk MDS, and have shown benefit in VEXAS patients (Comont et al. 2022; Raaijmakers et al. 2021; Arsene Mekinian et al. 2022). As many of the patients we

identified in the *UBA1*-profiling study actually have low-risk subtypes of MDS, it will be crucial to screen MDS patients with inflammatory presentation for *UBA1* mutations and incorporate this into clinical management of VEXAS/MDS.

The JAK inhibitor Ruxolitinib has been found to show particular benefit over other JAK inhibitors in VEXAS (Heiblig et al. 2022). However, responses to JAK inhibitors are highly variable and not yet well-characterized in a prospective cohort. Furthermore, this approach does not eradicate the *UBA1*-mutant clone, which poses a risk for patients eventually acquiring resistance to JAK inhibition as has been seen in other hematologic malignancies (Bhagwat, Levine, and Koppikar 2013).

Allogeneic stem cell transplant has the potential to be a curative treatment as it can eradicate the disease-causing *UBA1*-mutant clone from the hematopoietic stem and progenitor cell reservoir(Diarra et al. 2022; Loschi et al. 2022; Al-Hakim et al. 2022; Mangaonkar et al. 2023). However, many patients are not eligible for allogeneic hematopoietic stem cell transplant, which is an intensive procedure with high mortality rate. A clinical trial to evaluate the risk-benefit of this therapeutic option is ongoing (NCT05027945).

Given the limited therapeutic options for VEXAS patients, identification of additional interventions is a significant unmet need. To identify therapeutic vulnerabilities in VEXAS, further characterization of primary patient samples and development of preclinical models is needed. As described above, a mechanistic understanding of the association between MDS and VEXAS is also needed to manage the disease course. A case report of an MDS patient with a *UBA1* mutation treated with azacitidine and venetoclax showed concomitant reduction of the *UBA1* mutation and autoinflammatory symptoms, suggesting that

therapeutic approaches targeting *UBA1*-mutated cells can ameliorate the disease. Clone-specific phenotyping methods such as those established in this thesis may be able to identify targetable cell surface proteins or dysregulated biological pathways. Given the role of *UBA1* in initiating protein ubiquitination and degradation, therapeutic strategies targeting the proteasome or cellular stress pathways such as the unfolded protein response may also be beneficial.

Novel genomic drivers of hematologic malignancies and autoinflammatory diseases

While previous genomic studies have made huge strides in characterizing recurrently mutated genes in hematological malignancies, there are still a significant number of patients (5-10%) with no identified genetic alterations(Papaemmanuil et al. 2016; Tazi et al. 2022; Bernard et al. 2022). Newly recognized mutations in *UBA1* only explain a fraction of the mutation-negative MDS cases in our study and a small fraction of the disease etiology of the Undiagnosed Diseases cohort(Beck et al. 2020). Furthermore, there have been reports of patients with VEXAS-like clinical presentation that do not have *UBA1* mutations(Duminuco et al. 2022). There may be somatic mutations yet to be discovered and characterized whose identification will eventually unlock new diagnostic, prognostic and therapeutic strategies in AML and MDS, and overlapping autoinflammatory and autoimmune diseases. Systematic whole exome or whole genomic profiling of these cases may identify novel drivers of hematologic malignancies and/or autoinflammatory diseases

and enable improved risk stratification and clinical management for these patients. Biobanks with detailed and uniform clinical annotation are essential to supporting these large cohort discovery studies.

Additionally, there is building evidence for non-genetic drivers of hematologic malignancies including the microenvironment(Baryawno et al. 2019; Lasry et al. 2022; Barreyro, Chlon, and Starczynowski 2018; Galán-Díez et al. 2022). Rapidly advancing single cell technologies that enable high resolution profiling of the bone marrow niche in primary patient samples, including with spatial resolution(Baccin et al. 2020), will shed light on the complex and dynamic networks driving hematologic malignancies and autoinflammatory diseases.

In conclusion, in this thesis I 1) designed and executed population-based studies for biomarker characterization; 2) identified disease-defining genotypes in hematologic neoplasms; 3) delivered genotype-focused analysis of phylogenetic subclones and hematopoietic cellular hierarchies in primary patient samples; 4) deployed these approaches across distinct disease transitions including preleukemic, diagnosis, and on therapy. Scaling up these approaches to larger patient cohorts will enable identification of prognostic biomarkers or therapeutic vulnerabilities.

Bibliography

Abelson, Sagi, Grace Collord, Stanley W. K. Ng, Omer Weissbrod, Netta Mendelson Cohen, Elisabeth Niemeyer, Noam Barda, et al. 2018. “Prediction of Acute Myeloid Leukaemia Risk in Healthy Individuals.” *Nature* 559 (7714): 400–404.

Aibar, Sara, Carmen Bravo González-Blas, Thomas Moerman, Vân Anh Huynh-Thu, Hana Imrichova, Gert Hulselmans, Florian Rambow, et al. 2017. “SCENIC: Single-Cell Regulatory Network Inference and Clustering.” *Nature Methods* 14 (11): 1083–86.

Al-Hakim, Adam, James A. Poulter, Dina Mahmoud, Ailsa M. S. Rose, Suzanne Elcombe, Helen Lachmann, Catherine Cargo, et al. 2022. “Allogeneic Haematopoietic Stem Cell Transplantation for VEXAS Syndrome: UK Experience.” *British Journal of Haematology* 199 (5): 777–81.

Amatangelo, Michael D., Lynn Quek, Alan Shih, Eytan M. Stein, Mikhail Roshal, Muriel D. David, Benoit Marteyn, et al. 2017. “Enasidenib Induces Acute Myeloid Leukemia Cell Differentiation to Promote Clinical Response.” *Blood* 130 (6): 732–41.

Andreatta, Massimo, and Santiago J. Carmona. 2021. “UCell: Robust and Scalable Single-Cell Gene Signature Scoring.” *Computational and Structural Biotechnology Journal* 19 (June): 3796–98.

Arber, Daniel A., Attilio Orazi, Robert P. Hasserjian, Michael J. Borowitz, Katherine R. Calvo, Hans-Michael Kvasnicka, Sa A. Wang, et al. 2022. “International Consensus Classification of Myeloid Neoplasms and Acute Leukemias: Integrating Morphologic, Clinical, and Genomic Data.” *Blood* 140 (11): 1200–1228.

Armstrong, Scott A., Jane E. Staunton, Lewis B. Silverman, Rob Pieters, Monique L. den Boer, Mark D. Minden, Stephen E. Sallan, Eric S. Lander, Todd R. Golub, and Stanley J. Korsmeyer. 2002. “MLL Translocations Specify a Distinct Gene Expression Profile That Distinguishes a Unique Leukemia.” *Nature Genetics* 30 (1): 41–47.

“A Study of Enasidenib in People With Clonal Cytopenia of Undetermined Significance.” n.d. Accessed March 1, 2023. https://clinicaltrials.gov/ct2/show/NCT05102370.

Avagyan, S., J. E. Henninger, W. P. Mannherz, M. Mistry, J. Yoon, S. Yang, M. C. Weber, J. L. Moore, and L. I. Zon. 2021. “Resistance to Inflammation Underlies Enhanced Fitness in Clonal Hematopoiesis.” *Science* 374 (6568): 768–72.

Baccin, Chiara, Jude Al-Sabah, Lars Velten, Patrick M. Helbling, Florian Grünschläger, Pablo Hernández-Malmierca, César Nombela-Arrieta, Lars M. Steinmetz, Andreas Trumpp, and Simon Haas. 2020. “Combined Single-Cell and Spatial Transcriptomics Reveal the Molecular, Cellular and Spatial Bone Marrow Niche Organization.” *Nature Cell Biology* 22 (1): 38–48.

Ball, Brian J., Meier Hsu, Sean M. Devlin, Maria Arcila, Mikhail Roshal, Yanming Zhang, Chris A. Famulare, et al. 2021. “The Prognosis and Durable Clearance of RAS Mutations in Patients with Acute Myeloid Leukemia Receiving Induction Chemotherapy.” *American Journal of Hematology* 96 (5): E171–75.

Barba, Thomas, Yvan Jamilloux, Cécile-Audrey Durel, Estelle Bourbon, Fanélie Mestrallet, Pierre Sujobert, and Arnaud Hot. 2021. “VEXAS Syndrome in a Woman.” *Rheumatology* 60 (11): e402–3.

Barreyro, Laura, Timothy M. Chlon, and Daniel T. Starczynowski. 2018. “Chronic Immune Response Dysregulation in MDS Pathogenesis.” *Blood* 132 (15): 1553–60.

Baryawno, Ninib, Dariusz Przybylski, Monika S. Kowalczyk, Youmna Kfoury, Nicolas Severe, Karin Gustafsson, Konstantinos D. Kokkaliaris, et al. 2019. “A Cellular Taxonomy of the Bone Marrow Stroma in Homeostasis and Leukemia.” *Cell* 177 (7): 1915–32.e16.

Beck, David B., Dale L. Bodian, Vandan Shah, Uyenlinh L. Mirshahi, Jung Kim, Yi Ding, Samuel J. Magaziner, et al. 2023. “Estimated Prevalence and Clinical Manifestations of UBA1 Variants Associated With VEXAS Syndrome in a Clinical Population.” *JAMA: The Journal of the American Medical Association* 329 (4): 318–24.

Beck, David B., Marcela A. Ferrada, Keith A. Sikora, Amanda K. Ombrello, Jason C. Collins, Wuhong Pei, Nicholas Balanda, et al. 2020. “Somatic Mutations in UBA1 and Severe Adult-Onset Autoinflammatory Disease.” *The New England Journal of Medicine* 383 (27): 2628–38.

Beneyto-Calabuig, Sergi, Anne Kathrin Ludwig, Jonas-Alexander Kniffka, Chelsea Szu-Tu, Christian Rohde, Magdalena Antes, Alexander Waclawiczek, et al. 2022. “Clonally Resolved Single-Cell Multi-Omics Identifies Routes of Cellular Differentiation in Acute Myeloid Leukemia.” *bioRxiv*. https://doi.org/10.1101/2022.08.29.505648.

Bernard, Elsa, Yasuhito Nannya, Robert P. Hasserjian, Sean M. Devlin, Heinz Tuechler, Juan S. Medina-Martinez, Tetsuichi Yoshizato, et al. 2020. “Implications of TP53 Allelic State for Genome Stability, Clinical Presentation and Outcomes in Myelodysplastic Syndromes.” *Nature Medicine* 26 (10): 1549–56.

Bernard, Elsa, Heinz Tuechler, Peter L. Greenberg, Robert P. Hasserjian, Juan E. Arango Ossa, Yasuhito Nannya, Sean M. Devlin, et al. 2022. “Molecular International Prognostic Scoring System for Myelodysplastic Syndromes.” *NEJM Evidence* 1 (7). https://doi.org/10.1056/evidoa2200008.

Bhagwat, Neha, Ross L. Levine, and Priya Koppikar. 2013. “Sensitivity and Resistance of JAK2 Inhibitors to Myeloproliferative Neoplasms.” *International Journal of Hematology* 97 (6): 695–702.

Boddu, Prajwal C., and Amer M. Zeidan. 2019. “Myeloid Disorders after Autoimmune Disease.” *Best Practice & Research. Clinical Haematology* 32 (1): 74–88.

Bolton, Kelly L., Ryan N. Ptashkin, Teng Gao, Lior Braunstein, Sean M. Devlin, Daniel Kelly, Minal Patel, et al. 2020. “Cancer Therapy Shapes the Fitness Landscape of Clonal Hematopoiesis.” *Nature Genetics* 52 (11): 1219–26.

Botton, Stéphane de, Joseph M. Brandwein, Andrew H. Wei, Arnaud Pigneux, Bruno Quesnel, Xavier Thomas, Ollivier Legrand, et al. 2021. “Improved Survival with Enasidenib versus Standard of Care in Relapsed/refractory Acute Myeloid Leukemia Associated with IDH2 Mutations Using Historical Data and Propensity Score Matching Analysis.” *Cancer Medicine* 10 (18): 6336–43.

Bourbon, Estelle, Maël Heiblig, Mathieu Gerfaud Valentin, Thomas Barba, Cécile-Audrey Durel, Jean Christophe Lega, Fiorenza Barraco, Pascal Sève, Yvan Jamilloux, and Pierre Sujobert. 2021. “Therapeutic Options in VEXAS Syndrome: Insights from a Retrospective Series.” *Blood* 137 (26): 3682–84.

Bralten, Linda B. C., Nanne K. Kloosterhof, Rutger Balvers, Andrea Sacchetti, Lariesa Lapre, Martine Lamfers, Sieger Leenstra, et al. 2011. “IDH1 R132H Decreases Proliferation of Glioma Cell Lines in Vitro and in Vivo.” *Annals of Neurology* 69 (3): 455–63.

Braun, Thorsten, and Pierre Fenaux. 2013. “Myelodysplastic Syndromes (MDS) and Autoimmune Disorders (AD): Cause or Consequence?” *Best Practice & Research. Clinical Haematology* 26 (4): 327–36.

Butler, Andrew, Paul Hoffman, Peter Smibert, Efthymia Papalexi, and Rahul Satija.

2018. “Integrating Single-Cell Transcriptomic Data across Different Conditions, Technologies, and Species.” *Nature Biotechnology* 36 (5): 411–20.
Chen, Chong, Yu Liu, Chao Lu, Justin R. Cross, John P. Morris 4th, Aditya S. Shroff, Patrick S. Ward, James E. Bradner, Craig Thompson, and Scott W. Lowe. 2013. “Cancer-Associated IDH2 Mutants Drive an Acute Myeloid Leukemia That Is Susceptible to Brd4 Inhibition.” *Genes & Development* 27 (18): 1974–85.
Choe, Sung, Hongfang Wang, Courtney D. DiNardo, Eytan M. Stein, Stéphane de Botton, Gail J. Roboz, Jessica K. Altman, et al. 2020. “Molecular Mechanisms Mediating Relapse Following Ivosidenib Monotherapy in IDH1-Mutant Relapsed or Refractory AML.” *Blood Advances* 4 (9): 1894–1905.
Christopher, Matthew J., Allegra A. Petti, Michael P. Rettig, Christopher A. Miller, Ezhilarasi Chendamarai, Eric J. Duncavage, Jeffery M. Klco, et al. 2018. “Immune Escape of Relapsed AML Cells after Allogeneic Transplantation.” *The New England Journal of Medicine* 379 (24): 2330–41.
Cibulskis, Kristian, Michael S. Lawrence, Scott L. Carter, Andrey Sivachenko, David Jaffe, Carrie Sougnez, Stacey Gabriel, Matthew Meyerson, Eric S. Lander, and Gad Getz. 2013. “Sensitive Detection of Somatic Point Mutations in Impure and Heterogeneous Cancer Samples.” *Nature Biotechnology* 31 (3): 213–19.
Clark, Owen, Katharine Yen, and Ingo K. Mellinghoff. 2016. “Molecular Pathways: Isocitrate Dehydrogenase Mutations in CancerIsocitrate Dehydrogenase Mutations in Cancer.” *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research* 22 (8): 1837–42.
Collins, Jason C., Nicholas Balanda, Samuel J. Magaziner, Maya English, Daniela Ospina Cardona, Mrinal M. Patnaik, Benjamin Terrier, Olivier Kosmider, Achim Werner, and David B. Beck. 2022. “Novel Disease-Causing Mutations in UBA1 Reveal Disease Mechanisms in Bone Marrow Failure and Inflammation.” *Blood* 140 (Supplement 1): 2914–15.
Comont, Thibault, Mael Heiblig, Etienne Rivière, Louis Terriou, Julien Rossignol, Didier Bouscary, Virginie Rieu, et al. 2022. “Azacitidine for Patients with Vacuoles, E1 Enzyme, X-Linked, Autoinflammatory, Somatic Syndrome (VEXAS) and Myelodysplastic Syndrome: Data from the French VEXAS Registry.” *British Journal of Haematology* 196 (4): 969–74.
Demaree, Benjamin, Cyrille L. Delley, Harish N. Vasudevan, Cheryl A. C. Peretz, David Ruff, Catherine C. Smith, and Adam R. Abate. 2021. “Joint Profiling of DNA and Proteins in Single Cells to Dissect Genotype-Phenotype Associations in Leukemia.” *Nature Communications* 12 (1): 1583.
Desai, Pinkal, Nuria Mencia-Trinchant, Oleksandr Savenkov, Michael S. Simon, Gloria Cheang, Sangmin Lee, Michael Samuel, et al. 2018. “Somatic Mutations Precede Acute Myeloid Leukemia Years before Diagnosis.” *Nature Medicine* 24 (7): 1015–23.
Diarra, Ava, Nicolas Duployez, Elise Fournier, Claude Preudhomme, Valérie Coiteux, Leonardo Magro, Bruno Quesnel, et al. 2022. “Successful Allogeneic Hematopoietic Stem Cell Transplantation in Patients with VEXAS Syndrome: A 2-Center Experience.” *Blood Advances* 6 (3): 998–1003.
Dillon, Laura W., Gege Gui, Niveditha Ravindra, Zoë Wong, Georgia Andrew, Devdeep Mukherjee, Scott L. Zeger, et al. 2022. “Pre-MEASURE:*FLT3*-ITD and mutated*NPM1*measurable Residual Disease before Allogeneic Transplant in Adults with AML in First Remission.” *bioRxiv*. https://doi.org/10.1101/2022.10.21.22281322.
Dillon, Richard, Robert Hills, Sylvie Freeman, Nicola Potter, Jelena Jovanovic, Adam Ivey, Anju Shankar Kanda, et al. 2020. “Molecular MRD Status and Outcome after

Transplantation in NPM1-Mutated AML." *Blood* 135 (9): 680–88.
DiNardo, Courtney D., Andre C. Schuh, Eytan M. Stein, Pau Montesinos, Andrew H. Wei, Stéphane de Botton, Amer M. Zeidan, et al. 2021. "Enasidenib plus Azacitidine versus Azacitidine Alone in Patients with Newly Diagnosed, Mutant-IDH2 Acute Myeloid Leukaemia (AG221-AML-005): A Single-Arm, Phase 1b and Randomised, Phase 2 Trial." *The Lancet Oncology* 22 (11): 1597–1608.
DiNardo, Courtney D., Eytan M. Stein, Stéphane de Botton, Gail J. Roboz, Jessica K. Altman, Alice S. Mims, Ronan Swords, et al. 2018. "Durable Remissions with Ivosidenib in IDH1-Mutated Relapsed or Refractory AML." *The New England Journal of Medicine* 378 (25): 2386–98.
Dufva, Olli, Petri Pölönen, Oscar Brück, Mikko A. I. Keränen, Jay Klievink, Juha Mehtonen, Jani Huuhtanen, et al. 2020. "Immunogenomic Landscape of Hematological Malignancies." *Cancer Cell* 38 (3): 424–28.
Duminuco, Andrea, Calogero Vetro, Uros Markovic, Francesco Di Raimondo, and Giuseppe A. M. Palumbo. 2022. "VEXAS-like Syndrome: A Potential New Entity?" *Annals of Hematology* 101 (5): 1125–28.
Erba, Harry P., Amir T. Fathi, Ghayas C. Issa, Jessica K. Altman, Pau Montesinos, Mrinal M. Patnaik, James M. Foran, et al. 2022. "Update on a Phase 1/2 First-in-Human Study of the Menin-KMT2A (MLL) Inhibitor Ziftomenib (KO-539) in Patients with Relapsed or Refractory Acute Myeloid Leukemia." *Blood*. https://doi.org/10.1182/blood-2022-167412.
Fan, Jean, Hae-Ock Lee, Soohyun Lee, Da-Eun Ryu, Semin Lee, Catherine Xue, Seok Jin Kim, et al. 2018. "Linking Transcriptional and Genetic Tumor Heterogeneity through Allele Analysis of Single-Cell RNA-Seq Data." *Genome Research* 28 (8): 1217–27.
Fernandez, Jenna A., Bhavisha A. Patel, Fernanda Gutierrez-Rodrigues, Yael Kusne, Terra Lasho, Christy Finke, Matthew Koster, et al. 2022. "DNMT3A/TET2 Mutant Clonal Hematopoiesis in Vexas Syndrome Results in DNA Hypomethylation and Transcriptional Activation of WT1 and MPL Oncogenic Pathways." *Blood* 140 (Supplement 1): 11410–11.
Figueroa, Maria E., Omar Abdel-Wahab, Chao Lu, Patrick S. Ward, Jay Patel, Alan Shih, Yushan Li, et al. 2010. "Leukemic IDH1 and IDH2 Mutations Result in a Hypermethylation Phenotype, Disrupt TET2 Function, and Impair Hematopoietic Differentiation." *Cancer Cell* 18 (6): 553–67.
Fiskus, Warren, Steffen Boettcher, Naval Daver, Christopher P. Mill, Koji Sasaki, Christine E. Birdwell, John A. Davis, et al. 2022. "Effective Menin Inhibitor-Based Combinations against AML with MLL Rearrangement or NPM1 Mutation (NPM1c)." *Blood Cancer Journal* 12 (1): 5.
Forbes, Simon A., David Beare, Harry Boutselakis, Sally Bamford, Nidhi Bindal, John Tate, Charlotte G. Cole, et al. 2017. "COSMIC: Somatic Cancer Genetics at High-Resolution." *Nucleic Acids Research* 45 (D1): D777–83.
Forte, Dorian, Valentina Salvestrini, Giulia Corradi, Lara Rossi, Lucia Catani, Roberto M. Lemoli, Michele Cavo, and Antonio Curti. 2017. "The Tissue Inhibitor of Metalloproteinases-1 (TIMP-1) Promotes Survival and Migration of Acute Myeloid Leukemia Cells through CD63/PI3K/Akt/p21 Signaling." *Oncotarget* 8 (2): 2261–74.
Fu, Xudong, Randall M. Chin, Laurent Vergnes, Heejun Hwang, Gang Deng, Yanpeng Xing, Melody Y. Pai, et al. 2015. "2-Hydroxyglutarate Inhibits ATP Synthase and mTOR Signaling." *Cell Metabolism*. https://doi.org/10.1016/j.cmet.2015.06.009.
Gaiti, Federico, Ronan Chaligne, Hongcang Gu, Ryan M. Brand, Steven Kothen-Hill, Rafael C. Schulman, Kirill Grigorev, et al. 2019. "Epigenetic Evolution and Lineage Histories of Chronic Lymphocytic Leukaemia." *Nature* 569 (7757): 576–80.

Galán-Díez, Marta, Florence Borot, Abdullah Mahmood Ali, Junfei Zhao, Eva Gil-Iturbe, Xiaochuan Shan, Na Luo, et al. 2022. “Subversion of Serotonin Receptor Signaling in Osteoblasts by Kynurenine Drives Acute Myeloid Leukemia.” *Cancer Discovery*. https://doi.org/10.1158/2159-8290.cd-21-0692.

Galen, Peter van, Volker Hovestadt, Marc H. Wadsworth Ii, Travis K. Hughes, Gabriel K. Griffin, Sofia Battaglia, Julia A. Verga, et al. 2019. “Single-Cell RNA-Seq Reveals AML Hierarchies Relevant to Disease Progression and Immunity.” *Cell* 176 (6): 1265–81.e24.

Gao, Teng, Ruslan Soldatov, Hirak Sarkar, Adam Kurkiewicz, Evan Biederstedt, Po-Ru Loh, and Peter V. Kharchenko. 2022. “Haplotype-Aware Analysis of Somatic Copy Number Variations from Single-Cell Transcriptomes.” *Nature Biotechnology*, September. https://doi.org/10.1038/s41587-022-01468-y.

Gbyli, Rana, Yuanbin Song, Wei Liu, Yimeng Gao, Giulia Biancon, Namrata S. Chandhok, Xiaman Wang, et al. 2022. “In Vivo Anti-Tumor Effect of PARP Inhibition in IDH1/2 Mutant MDS/AML Resistant to Targeted Inhibitors of Mutant IDH1/2.” *Leukemia* 36 (5): 1313–23.

“Genomic and Epigenomic Landscapes of Adult De Novo Acute Myeloid Leukemia.” 2013. *The New England Journal of Medicine* 368 (22): 2059–74.

Genovese, Giulio, Siddhartha Jaiswal, Benjamin L. Ebert, and Steven A. McCarroll. 2015. “Clonal Hematopoiesis and Blood-Cancer Risk.” *The New England Journal of Medicine*.

Georgin-Lavialle, S., B. Terrier, A. F. Guedon, M. Heiblig, T. Comont, E. Lazaro, V. Lacombe, et al. 2022. “Further Characterization of Clinical and Laboratory Features in VEXAS Syndrome: Large-Scale Analysis of a Multicentre Case Series of 116 French Patients.” *The British Journal of Dermatology* 186 (3): 564–74.

Greenman, Christopher, Philip Stephens, Raffaella Smith, Gillian L. Dalgliesh, Christopher Hunter, Graham Bignell, Helen Davies, et al. 2007. “Patterns of Somatic Mutation in Human Cancer Genomes.” *Nature* 446 (7132): 153–58.

Gruber, Emily, Joan So, Alexander C. Lewis, Rheana Franich, Rachel Cole, Luciano G. Martelotto, Amy J. Rogers, et al. 2022. “Inhibition of Mutant IDH1 Promotes Cycling of Acute Myeloid Leukemia Stem Cells.” *Cell Reports* 40 (7): 111182.

Haferlach, T., Y. Nagata, V. Grossmann, Y. Okuno, U. Bacher, G. Nagae, S. Schnittger, et al. 2014. “Landscape of Genetic Lesions in 944 Patients with Myelodysplastic Syndromes.” *Leukemia*. https://doi.org/10.1038/leu.2013.336.

Harding, James J., Maeve A. Lowery, Alan H. Shih, Juan M. Schvartzman, Shengqi Hou, Christopher Famulare, Minal Patel, et al. 2018. “Isoform Switching as a Mechanism of Acquired Resistance to Mutant Isocitrate Dehydrogenase Inhibition.” *Cancer Discovery*. https://doi.org/10.1158/2159-8290.cd-18-0877.

Hecker, Judith S., Luise Hartmann, Jennifer Rivière, Michèle C. Buck, Mark van der Garde, Maja Rothenberg-Thurley, Luise Fischer, et al. 2021. “CHIP and Hips: Clonal Hematopoiesis Is Common in Patients Undergoing Hip Arthroplasty and Is Associated with Autoimmune Disease.” *Blood* 138 (18): 1727–32.

Heiblig, Maël, Marcela A. Ferrada, Matthew T. Koster, Thomas Barba, Mathieu Gerfaud-Valentin, Arsène Mékinian, Henrique Coelho, et al. 2022. “Ruxolitinib Is More Effective than Other JAK Inhibitors to Treat VEXAS Syndrome: A Retrospective Multicenter Study.” *Blood*. https://doi.org/10.1182/blood.2022016642.

Hernández-Malmierca, Pablo, Dominik Vonficht, Alexandra Schnell, Hannah J. Uckelmann, Alina Bollhagen, Mohamed A. A. Mahmoud, Sophie-Luise Landua, et al. 2022. “Antigen Presentation Safeguards the Integrity of the Hematopoietic Stem Cell Pool.” *Cell Stem Cell* 29 (5): 760–75.e10.

Hollanda, A. de, A. Beucher, D. Henrion, A. Ghali, C. Lavigne, H. Lévesque, M.

Hamidou, J. F. Subra, N. Ifrah, and C. Belizna. 2011. “Systemic and Immune Manifestations in Myelodysplasia: A Multicenter Retrospective Study.” *Arthritis Care & Research* 63 (8): 1188–94.

Intlekofer, Andrew M., Alan H. Shih, Bo Wang, Abbas Nazir, Ariën S. Rustenburg, Steven K. Albanese, Minal Patel, et al. 2018. “Acquired Resistance to IDH Inhibition through Trans or Cis Dimer-Interface Mutations.” *Nature* 559 (7712): 125–29.

Issa, Ghayas C., Ibrahim Aldoss, John F. DiPersio, Branko Cuglievan, Richard M. Stone, Martha L. Arellano, Michael J. Thirman, et al. 2022. “The Menin Inhibitor SNDX-5613 (revumenib) Leads to Durable Responses in Patients (Pts) with *KMT2A*-Rearranged or *NPM1* Mutant AML: Updated Results of a Phase (Ph) 1 Study.” *Blood*. https://doi.org/10.1182/blood-2022-164849.

Ivey, Adam, Robert K. Hills, Michael A. Simpson, Jelena V. Jovanovic, Amanda Gilkes, Angela Grech, Yashma Patel, et al. 2016. “Assessment of Minimal Residual Disease in Standard-Risk AML.” *The New England Journal of Medicine* 374 (5): 422–33.

“Ivosidenib for Patients With Clonal Cytopenia of Undetermined Significance and Mutations in IDH1.” n.d. Accessed March 1, 2023. https://clinicaltrials.gov/ct2/show/NCT05030441.

Jaiswal, Siddhartha, and Benjamin L. Ebert. 2019. “Clonal Hematopoiesis in Human Aging and Disease.” *Science* 366 (6465). https://doi.org/10.1126/science.aan4673.

Jaiswal, Siddhartha, Pierre Fontanillas, Jason Flannick, Alisa Manning, Peter V. Grauman, Brenton G. Mar, R. Coleman Lindsley, et al. 2014. “Age-Related Clonal Hematopoiesis Associated with Adverse Outcomes.” *The New England Journal of Medicine* 371 (26): 2488–98.

Jones, David, Keiran M. Raine, Helen Davies, Patrick S. Tarpey, Adam P. Butler, Jon W. Teague, Serena Nik-Zainal, and Peter J. Campbell. 2016. “cgpCaVEManWrapper: Simple Execution of CaVEMan in Order to Detect Somatic Single Nucleotide Variants in NGS Data.” *Current Protocols in Bioinformatics / Editoral Board, Andreas D. Baxevanis ... [et Al.]* 56 (December): 15.10.1–15.10.18.

Kandoth, Cyriac, Michael D. McLellan, Fabio Vandin, Kai Ye, Beifang Niu, Charles Lu, Mingchao Xie, et al. 2013. “Mutational Landscape and Significance across 12 Major Cancer Types.” *Nature* 502 (7471): 333–39.

Kantarjian, Hagop, Tapan Kadia, Courtney DiNardo, Naval Daver, Gautam Borthakur, Elias Jabbour, Guillermo Garcia-Manero, Marina Konopleva, and Farhad Ravandi. 2021. “Acute Myeloid Leukemia: Current Progress and Future Directions.” *Blood Cancer Journal* 11 (2): 41.

Karamitros, Dimitris, Bilyana Stoilova, Zahra Aboukhalil, Fiona Hamey, Andreas Reinisch, Marina Samitsch, Lynn Quek, et al. 2018. “Single-Cell Analysis Reveals the Continuum of Human Lympho-Myeloid Progenitor Cells.” *Nature Immunology* 19 (1): 85–97.

Kats, Lev M., Markus Reschke, Riccardo Taulli, Olga Pozdnyakova, Kerri Burgess, Parul Bhargava, Kimberly Straley, et al. 2014. “Proto-Oncogenic Role of Mutant IDH2 in Leukemia Initiation and Maintenance.” *Cell Stem Cell*. https://doi.org/10.1016/j.stem.2013.12.016.

Kats, L. M., S. J. Vervoort, R. Cole, A. J. Rogers, G. P. Gregory, E. Vidacs, J. Li, R. Nagaraja, K. E. Yen, and R. W. Johnstone. 2017. “A Pharmacogenomic Approach Validates AG-221 as an Effective and on-Target Therapy in IDH2 Mutant AML.” *Leukemia* 31 (6): 1466–70.

Khoury, Joseph D., Eric Solary, Oussama Abla, Yassmine Akkari, Rita Alaggio, Jane F. Apperley, Rafael Bejar, et al. 2022. “The 5th Edition of the World Health Organization Classification of Haematolymphoid Tumours: Myeloid and

Histiocytic/Dendritic Neoplasms." *Leukemia* 36 (7): 1703–19.
Komorowska, Karolina, Alexander Doyle, Martin Wahlestedt, Agatheeswaran Subramaniam, Shubhranshu Debnath, Jun Chen, Shamit Soneji, et al. 2017. "Hepatic Leukemia Factor Maintains Quiescence of Hematopoietic Stem Cells and Protects the Stem Cell Pool during Regeneration." *Cell Reports* 21 (12): 3514–23.
Lasorella, Anna, Robert Benezra, and Antonio Iavarone. 2014. "The ID Proteins: Master Regulators of Cancer Stem Cells and Tumour Aggressiveness." *Nature Reviews. Cancer* 14 (2): 77–91.
Lasry, Audrey, Bettina Nadorp, Maarten Fornerod, Deedra Nicolet, Huiyun Wu, Christopher J. Walker, Zhengxi Sun, et al. 2022. "An Inflammatory State Remodels the Immune Microenvironment and Improves Risk Stratification in Acute Myeloid Leukemia." *Nature Cancer*. https://doi.org/10.1038/s43018-022-00480-0.
Liberzon, Arthur, Chet Birger, Helga Thorvaldsdóttir, Mahmoud Ghandi, Jill P. Mesirov, and Pablo Tamayo. 2015. "The Molecular Signatures Database (MSigDB) Hallmark Gene Set Collection." *Cell Systems* 1 (6): 417–25.
Loschi, Michael, Christian Roux, Isabelle Sudaka, Corinne Ferrero-Vacher, Alice Marceau-Renaut, Nicolas Duployez, Thierry Passeron, and Thomas Cluzeau. 2022. "Allogeneic Stem Cell Transplantation as a Curative Therapeutic Approach for VEXAS Syndrome: A Case Report." *Bone Marrow Transplantation* 57 (2): 315–18.
Lu, Chao, Patrick S. Ward, Gurpreet S. Kapoor, Dan Rohle, Sevin Turcan, Omar Abdel-Wahab, Christopher R. Edwards, et al. 2012. "IDH Mutation Impairs Histone Demethylation and Results in a Block to Cell Differentiation." *Nature* 483 (7390): 474–78.
Ludwig, Leif S., Caleb A. Lareau, Jacob C. Ulirsch, Elena Christian, Christoph Muus, Lauren H. Li, Karin Pelka, et al. 2019. "Lineage Tracing in Humans Enabled by Mitochondrial Mutations and Single-Cell Genomics." *Cell*. https://doi.org/10.1016/j.cell.2019.01.022.
Lu, Yi-Chien, Chad Sanada, Juliana Xavier-Ferrucio, Lin Wang, Ping-Xia Zhang, H. Leighton Grimes, Meenakshi Venkatasubramanian, et al. 2018. "The Molecular Signature of Megakaryocyte-Erythroid Progenitors Reveals a Role for the Cell Cycle in Fate Specification." *Cell Reports* 25 (11): 3229.
Mangaonkar, Abhishek A., Kimberly J. Langer, Terra L. Lasho, Christy Finke, Mark R. Litzow, William J. Hogan, Mithun V. Shah, et al. 2023. "Reduced Intensity Conditioning Allogeneic Hematopoietic Stem Cell Transplantation in VEXAS Syndrome: Data from a Prospective Series of Patients." *American Journal of Hematology* 98 (2): E28–31.
Marnell, Christopher S., Alexander Bick, and Pradeep Natarajan. 2021. "Clonal Hematopoiesis of Indeterminate Potential (CHIP): Linking Somatic Mutations, Hematopoiesis, Chronic Inflammation and Cardiovascular Disease." *Journal of Molecular and Cellular Cardiology* 161 (December): 98–105.
Martincorena, Iñigo, and Peter J. Campbell. 2015. "Somatic Mutation in Cancer and Normal Cells." *Science* 349 (6255): 1483–89.
Martincorena, Iñigo, Joanna C. Fowler, Agnieszka Wabik, Andrew R. J. Lawson, Federico Abascal, Michael W. J. Hall, Alex Cagan, et al. 2018. "Somatic Mutant Clones Colonize the Human Esophagus with Age." *Science* 362 (6417): 911–17.
Martincorena, Iñigo, Amit Roshan, Moritz Gerstung, Peter Ellis, Peter Van Loo, Stuart McLaren, David C. Wedge, et al. 2015. "Tumor Evolution. High Burden and Pervasive Positive Selection of Somatic Mutations in Normal Human Skin." *Science* 348 (6237): 880–86.
Medina-Martínez, Juan S., Juan E. Arango-Ossa, Max F. Levine, Yangyu Zhou, Gunes Gundem, Andrew L. Kung, and Elli Papaemmanuil. 2020. "Isabl Platform, a Digital

Biobank for Processing Multimodal Patient Data." *BMC Bioinformatics* 21 (1): 549.
Mekinian, Arsène, Eric Grignano, Thorsten Braun, Olivier Decaux, Eric Liozon, Nathalie Costedoat-Chalumeau, Jean-Emmanuel Kahn, et al. 2016. "Systemic Inflammatory and Autoimmune Manifestations Associated with Myelodysplastic Syndromes and Chronic Myelomonocytic Leukaemia: A French Multicentre Retrospective Study." *Rheumatology* 55 (2): 291–300.
Mekinian, Arsene, Lin Pierre Zhao, Sylvie Chevret, Kristell Desseaux, Laurent Pascal, Thibaut Comont, Alexandre Maria, et al. 2022. "A Phase II Prospective Trial of Azacitidine in Steroid-Dependent or Refractory Systemic Autoimmune/inflammatory Disorders and VEXAS Syndrome Associated with MDS and CMML." *Leukemia* 36 (11): 2739–42.
Miles, Linde A., Robert L. Bowman, Tiffany R. Merlinsky, Isabelle S. Csete, Aik T. Ooi, Robert Durruthy-Durruthy, Michael Bowman, et al. 2020. "Single-Cell Mutation Analysis of Clonal Evolution in Myeloid Malignancies." *Nature* 587 (7834): 477–82.
Montesinos, Pau, Christian Recher, Susana Vives, Ewa Zarzycka, Jianxiang Wang, Giambattista Bertani, Michael Heuser, et al. 2022. "Ivosidenib and Azacitidine in IDH1-Mutated Acute Myeloid Leukemia." *The New England Journal of Medicine* 386 (16): 1519–31.
Montoro, Julia, Laura Gallur, Brayan Merchán, Antonieta Molero, Elisa Roldán, Ferrán Martínez-Valle, Guillermo Villacampa, et al. 2018. "Autoimmune Disorders Are Common in Myelodysplastic Syndrome Patients and Confer an Adverse Impact on Outcomes." *Annals of Hematology* 97 (8): 1349–56.
Moran-Crusio, Kelly, Linsey Reavie, Alan Shih, Omar Abdel-Wahab, Delphine Ndiaye-Lobry, Camille Lobry, Maria E. Figueroa, et al. 2011. "Tet2 Loss Leads to Increased Hematopoietic Stem Cell Self-Renewal and Myeloid Transformation." *Cancer Cell* 20 (1): 11–24.
Morita, Kiyomi, Feng Wang, Katharina Jahn, Tianyuan Hu, Tomoyuki Tanaka, Yuya Sasaki, Jack Kuipers, et al. 2020. "Clonal Evolution of Acute Myeloid Leukemia Revealed by High-Throughput Single-Cell Genomics." *Nature Communications* 11 (1): 5327.
Myers, Robert M., Franco Izzo, Sanjay Kottapalli, Tamara Prieto, Andrew Dunbar, Robert L. Bowman, Eleni P. Mimitou, et al. 2022. "Integrated Single-Cell Genotyping and Chromatin Accessibility Charts JAK2V617F Human Hematopoietic Differentiation." *bioRxiv*. https://doi.org/10.1101/2022.05.11.491515.
Mylonas, E., M. Janin, O. Bawa, P. Opolon, M. David, C. Quivoron, O. A. Bernard, C. Ottolenghi, S. DeBotton, and V. Penard-Lacronique. 2014. "Isocitrate Dehydrogenase (IDH)2 R140Q Mutation Induces Myeloid and Lymphoid Neoplasms in Mice." *Leukemia* 28 (6): 1343–46.
Nam, Anna S., Kyu-Tae Kim, Ronan Chaligne, Franco Izzo, Chelston Ang, Justin Taylor, Robert M. Myers, et al. 2019. "Somatic Mutations and Cell Identity Linked by Genotyping of Transcriptomes." *Nature* 571 (7765): 355–60.
Oganesyan, Artem, Vincent Jachiet, Francois Chasset, Pierre Hirsch, Mehdi Hage-Sleiman, Bettina Fabiani, Paul Duriez, et al. 2021. "VEXAS Syndrome: Still Expanding the Clinical Phenotype." *Rheumatology* 60 (9): e321–23.
Ok, Chi Young, Sanam Loghavi, Dawen Sui, Peng Wei, Rashmi Kanagal-Shamanna, C. Cameron Yin, Zhuang Zuo, et al. 2019. "Persistent *IDH1/2* Mutations in Remission Can Predict Relapse in Patients with Acute Myeloid Leukemia." *Haematologica*. https://doi.org/10.3324/haematol.2018.191148.
Papaemmanuil, Elli, Moritz Gerstung, Lars Bullinger, Verena I. Gaidzik, Peter Paschka, Nicola D. Roberts, Nicola E. Potter, et al. 2016. "Genomic Classification and Prognosis in Acute Myeloid Leukemia." *The New England Journal of Medicine* 374

(23): 2209–21.
Paschka, Peter, Richard F. Schlenk, Verena I. Gaidzik, Marianne Habdank, Jan Krönke, Lars Bullinger, Daniela Späth, et al. 2010. “IDH1 and IDH2 Mutations Are Frequent Genetic Alterations in Acute Myeloid Leukemia and Confer Adverse Prognosis in Cytogenetically Normal Acute Myeloid Leukemia with NPM1 Mutation without FLT3 Internal Tandem Duplication.” *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology* 28 (22): 3636–43.
Patel, Bhavisha A., Fernanda Gutierrez-Rodrigues, Yael Kusne, Jenna A. Fernandez, Terra Lasho, Ruba Shalhoub, Xiaoyang Ma, et al. 2022. “Clonal Hematopoiesis in Vexas Syndrome.” *Blood* 140 (Supplement 1): 5745–46.
Pellin, Danilo, Mariana Loperfido, Cristina Baricordi, Samuel L. Wolock, Annita Montepeloso, Olga K. Weinberg, Alessandra Biffi, Allon M. Klein, and Luca Biasco. 2019. “A Comprehensive Single Cell Transcriptional Landscape of Human Hematopoietic Progenitors.” *Nature Communications* 10 (1): 2395.
Perl, Alexander E., Giovanni Martinelli, Jorge E. Cortes, Andreas Neubauer, Ellin Berman, Stefania Paolini, Pau Montesinos, et al. 2019. “Gilteritinib or Chemotherapy for Relapsed or Refractory FLT3-Mutated AML.” *The New England Journal of Medicine* 381 (18): 1728–40.
Petti, Allegra A., Stephen R. Williams, Christopher A. Miller, Ian T. Fiddes, Sridhar N. Srivatsan, David Y. Chen, Catrina C. Fronick, Robert S. Fulton, Deanna M. Church, and Timothy J. Ley. 2019. “A General Approach for Detecting Expressed Mutations in AML Cells Using Single Cell RNA-Sequencing.” *Nature Communications* 10 (1): 3660.
Pollyea, Daniel A., Martin S. Tallman, Stéphane de Botton, Hagop M. Kantarjian, Robert Collins, Anthony S. Stein, Mark G. Frattini, et al. 2019. “Enasidenib, an Inhibitor of Mutant IDH2 Proteins, Induces Durable Remissions in Older Patients with Newly Diagnosed Acute Myeloid Leukemia.” *Leukemia* 33 (11): 2575–84.
Popescu, Dorin-Mirel, Rachel A. Botting, Emily Stephenson, Kile Green, Simone Webb, Laura Jardine, Emily F. Calderbank, et al. 2019. “Decoding Human Fetal Liver Haematopoiesis.” *Nature* 574 (7778): 365–71.
Poulter, James A., Jason C. Collins, Catherine Cargo, Ruth M. De Tute, Paul Evans, Daniela Ospina Cardona, David T. Bowen, et al. 2021. “Novel Somatic Mutations in UBA1 as a Cause of VEXAS Syndrome.” *Blood*.
Psaila, Bethan, Nikolaos Barkas, Deena Iskander, Anindita Roy, Stacie Anderson, Neil Ashley, Valentina S. Caputo, et al. 2016. “Single-Cell Profiling of Human Megakaryocyte-Erythroid Progenitors Identifies Distinct Megakaryocyte and Erythroid Differentiation Pathways.” *Genome Biology* 17 (May): 83.
Quek, Lynn, Muriel D. David, Alison Kennedy, Marlen Metzner, Michael Amatangelo, Alan Shih, Bilyana Stoilova, et al. 2018. “Clonal Heterogeneity of Acute Myeloid Leukemia Treated with the IDH2 Inhibitor Enasidenib.” *Nature Medicine* 24 (8): 1167–77.
Raaijmakers, Marc H. G. P., Maud Hermans, Anna Aalbers, Melissa Rijken, Virgil A. S. H. Dalm, Paul van Daele, and Peter J. M. Valk. 2021. “Azacytidine Treatment for VEXAS Syndrome.” *HemaSphere* 5 (12): e661.
Ricard, Laure, Pierre Hirsch, Laëtitia Largeaud, Caroline Deswarte, Vincent Jachiet, Mohamad Mohty, Sébastien Rivière, et al. 2020. “Clonal Haematopoiesis Is Increased in Early Onset in Systemic Sclerosis.” *Rheumatology* 59 (11): 3499–3504.
Roboz, Gail J., Courtney D. DiNardo, Eytan M. Stein, Stéphane de Botton, Alice S. Mims, Gabrielle T. Prince, Jessica K. Altman, et al. 2020. “Ivosidenib Induces Deep Durable Remissions in Patients with Newly Diagnosed IDH1-Mutant Acute Myeloid

Leukemia." *Blood* 135 (7): 463–71.
Rodriguez-Meira, Alba, Gemma Buck, Sally-Ann Clark, Benjamin J. Povinelli, Veronica Alcolea, Eleni Louka, Simon McGowan, et al. 2019. "Unravelling Intratumoral Heterogeneity through High-Sensitivity Single-Cell Mutational Analysis and Parallel RNA Sequencing." *Molecular Cell* 73 (6): 1292–1305.e8.
Sakuma, Maki, Piers Blombery, Manja Meggendorfer, Claudia Haferlach, Markus Lindauer, Uwe M. Martens, Wolfgang Kern, Torsten Haferlach, and Wencke Walter. 2023. "Novel Causative Variants of VEXAS in UBA1 Detected through Whole Genome Transcriptome Sequencing in a Large Cohort of Hematological Malignancies." *Leukemia*, February. https://doi.org/10.1038/s41375-023-01857-5.
Sano, Soichi, Kosei Oshima, Ying Wang, Susan MacLauchlan, Yasufumi Katanasaka, Miho Sano, María A. Zuriaga, et al. 2018. "Tet2-Mediated Clonal Hematopoiesis Accelerates Heart Failure Through a Mechanism Involving the IL-1β/NLRP3 Inflammasome." *Journal of the American College of Cardiology*. https://doi.org/10.1016/j.jacc.2017.12.037.
Sasaki, Masato, Christiane B. Knobbe, Momoe Itsumi, Andrew J. Elia, Isaac S. Harris, Iok In Christine Chio, Rob A. Cairns, et al. 2012. "D-2-Hydroxyglutarate Produced by Mutant IDH1 Perturbs Collagen Maturation and Basement Membrane Function." *Genes & Development* 26 (18): 2038–49.
Sasaki, Masato, Christiane B. Knobbe, Joshua C. Munger, Evan F. Lind, Dirk Brenner, Anne Brüstle, Isaac S. Harris, et al. 2012. "IDH1(R132H) Mutation Increases Murine Haematopoietic Progenitors and Alters Epigenetics." *Nature* 488 (7413): 656–59.
Saunders, Christopher T., Wendy S. W. Wong, Sajani Swamy, Jennifer Becq, Lisa J. Murray, and R. Keira Cheetham. 2012. "Strelka: Accurate Somatic Small-Variant Calling from Sequenced Tumor-Normal Sample Pairs." *Bioinformatics* 28 (14): 1811–17.
Shahid, Sanam, Nicholas Ceglia, Jean-Benoit Le Luduec, Andrew McPherson, Barbara Spitzer, Viktoria Bojilova, M. Kazim Panjwani, et al. 2022. "Immune Escape after Allogeneic Hematopoietic Cell Transplantation in Pediatric Patients with Acute Myeloid Leukemia." *Blood* 140 (Supplement 1): 10209–10.
Shlush, Liran I., Sasan Zandi, Amanda Mitchell, Weihsu Claire Chen, Joseph M. Brandwein, Vikas Gupta, James A. Kennedy, et al. 2014. "Identification of Pre-Leukaemic Haematopoietic Stem Cells in Acute Leukaemia." *Nature* 506 (7488): 328–33.
Stein, Eytan M., Courtney D. DiNardo, Amir T. Fathi, Alice S. Mims, Keith W. Pratz, Michael R. Savona, Anthony S. Stein, et al. 2021. "Ivosidenib or Enasidenib Combined with Intensive Chemotherapy in Patients with Newly Diagnosed AML: A Phase 1 Study." *Blood* 137 (13): 1792–1803.
Stein, Eytan M., Courtney D. DiNardo, Daniel A. Pollyea, Amir T. Fathi, Gail J. Roboz, Jessica K. Altman, Richard M. Stone, et al. 2017. "Enasidenib in Mutant IDH2 Relapsed or Refractory Acute Myeloid Leukemia." *Blood* 130 (6): 722–31.
Stiburkova, Blanka, Katerina Pavelcova, Monika Belickova, Samuel J. Magaziner, Jason C. Collins, Achim Werner, David B. Beck, et al. 2023. "Novel Somatic UBA1 Variant in a Patient with VEXAS Syndrome." *Arthritis & Rheumatology (Hoboken, N.J.)*, February. https://doi.org/10.1002/art.42471.
Stoeckius, Marlon, Christoph Hafemeister, William Stephenson, Brian Houck-Loomis, Pratip K. Chattopadhyay, Harold Swerdlow, Rahul Satija, and Peter Smibert. 2017. "Simultaneous Epitope and Transcriptome Measurement in Single Cells." *Nature Methods* 14 (9): 865–68.
Stone, Richard M., Sumithra J. Mandrekar, Ben L. Sanford, Kristina Laumann, Susan Geyer, Clara D. Bloomfield, Christian Thiede, et al. 2017. "Midostaurin plus

Chemotherapy for Acute Myeloid Leukemia with a FLT3 Mutation." *The New England Journal of Medicine* 377 (5): 454–64.
Stuani, Lucille, Marie Sabatier, Estelle Saland, Guillaume Cognet, Nathalie Poupin, Claudie Bosc, Florence A. Castelli, et al. 2021. "Mitochondrial Metabolism Supports Resistance to IDH Mutant Inhibitors in Acute Myeloid Leukemia." *The Journal of Experimental Medicine* 218 (5). https://doi.org/10.1084/jem.20200924.
Stubbins, Ryan J., Eric McGinnis, Bhupinder Johal, Luke Yc Chen, Lorena Wilson, Daniela Ospina Cardona, and Thomas J. Nevill. 2022. "VEXAS Syndrome in a Female Patient with Constitutional 45,X (Turner Syndrome)." *Haematologica* 107 (4): 1011–13.
Sulkowski, Parker L., Christopher D. Corso, Nathaniel D. Robinson, Susan E. Scanlon, Karin R. Purshouse, Hanwen Bai, Yanfeng Liu, et al. 2017. "2-Hydroxyglutarate Produced by Neomorphic IDH Mutations Suppresses Homologous Recombination and Induces PARP Inhibitor Sensitivity." *Science Translational Medicine* 9 (375). https://doi.org/10.1126/scitranslmed.aal2463.
Su, Rui, Lei Dong, Chenying Li, Sigrid Nachtergaele, Mark Wunderlich, Ying Qing, Xiaolan Deng, et al. 2018. "R-2HG Exhibits Anti-Tumor Activity by Targeting FTO/m6A/MYC/CEBPA Signaling." *Cell* 172 (1-2): 90–105.e23.
Tazi, Yanis, Juan E. Arango-Ossa, Yangyu Zhou, Elsa Bernard, Ian Thomas, Amanda Gilkes, Sylvie Freeman, et al. 2022. "Unified Classification and Risk-Stratification in Acute Myeloid Leukemia." *Nature Communications* 13 (1): 4622.
Tickle, T., I. Tirosh, C. Georgescu, M. Brown, and B. Haas. n.d. "inferCNV of the Trinity CTAT Project." *Broad Institute of MIT and Harvard*.
Tirosh, Itay, Benjamin Izar, Sanjay M. Prakadan, Marc H. Wadsworth 2nd, Daniel Treacy, John J. Trombetta, Asaf Rotem, et al. 2016. "Dissecting the Multicellular Ecosystem of Metastatic Melanoma by Single-Cell RNA-Seq." *Science* 352 (6282): 189–96.
Toffalori, Cristina, Laura Zito, Valentina Gambacorta, Michela Riba, Giacomo Oliveira, Gabriele Bucci, Matteo Barcella, et al. 2019. "Immune Signature Drives Leukemia Escape and Relapse after Hematopoietic Cell Transplantation." *Nature Medicine* 25 (4): 603–11.
Triana, Sergio, Dominik Vonficht, Lea Jopp-Saile, Simon Raffel, Raphael Lutz, Daniel Leonce, Magdalena Antes, et al. 2021. "Single-Cell Proteo-Genomic Reference Maps of the Hematopoietic System Enable the Purification and Massive Profiling of Precisely Defined Cell States." *Nature Immunology* 22 (12): 1577–89.
Tsoucas, Daphne, Rui Dong, Haide Chen, Qian Zhu, Guoji Guo, and Guo-Cheng Yuan. 2019. "Accurate Estimation of Cell-Type Composition from Gene Expression Data." *Nature Communications*. https://doi.org/10.1038/s41467-019-10802-z.
Velten, Lars, Simon F. Haas, Simon Raffel, Sandra Blaszkiewicz, Saiful Islam, Bianca P. Hennig, Christoph Hirche, et al. 2017. "Human Haematopoietic Stem Cell Lineage Commitment Is a Continuous Process." *Nature Cell Biology* 19 (4): 271–81.
Velten, Lars, Benjamin A. Story, Pablo Hernández-Malmierca, Simon Raffel, Daniel R. Leonce, Jennifer Milbank, Malte Paulsen, et al. 2021. "Identification of Leukemic and Pre-Leukemic Stem Cells by Clonal Tracking from Single-Cell Transcriptomics." *Nature Communications* 12 (1): 1366.
Wang, Fang, Jeremy Travins, Byron DeLaBarre, Virginie Penard-Lacronique, Stefanie Schalm, Erica Hansen, Kimberly Straley, et al. 2013. "Targeted Inhibition of Mutant IDH2 in Leukemia Cells Induces Cellular Differentiation." *Science* 340 (6132): 622–26.
Wang, Feng, Kiyomi Morita, Courtney D. DiNardo, Ken Furudate, Tomoyuki Tanaka, Yuanqing Yan, Keyur P. Patel, et al. 2021. "Leukemia Stemness and Co-Occurring

Mutations Drive Resistance to IDH Inhibitors in Acute Myeloid Leukemia." *Nature Communications* 12 (1): 2607.
Ward, Patrick S., Jay Patel, David R. Wise, Omar Abdel-Wahab, Bryson D. Bennett, Hilary A. Coller, Justin R. Cross, et al. 2010. "The Common Feature of Leukemia-Associated IDH1 and IDH2 Mutations Is a Neomorphic Enzyme Activity Converting Alpha-Ketoglutarate to 2-Hydroxyglutarate." *Cancer Cell* 17 (3): 225–34.
Watson, Caroline J., A. L. Papula, Gladys Y. P. Poon, Wing H. Wong, Andrew L. Young, Todd E. Druley, Daniel S. Fisher, and Jamie R. Blundell. 2020. "The Evolutionary Dynamics and Fitness Landscape of Clonal Hematopoiesis." *Science* 367 (6485): 1449–54.
Wiseman, D. H., E. L. Williams, D. P. Wilks, H. Sun Leong, T. D. D. Somerville, M. W. Dennis, E. A. Struys, A. Bakkali, G. S. Salomons, and T. C. P. Somervaille. 2016. "Frequent Reconstitution of IDH2R140Q Mutant Clonal Multilineage Hematopoiesis Following Chemotherapy for Acute Myeloid Leukemia." *Leukemia* 30 (9): 1946–50.
Xu, Jin, Kevin Nuno, Ulrike M. Litzenburger, Yanyan Qi, M. Ryan Corces, Ravindra Majeti, and Howard Y. Chang. 2019. "Single-Cell Lineage Tracing by Endogenous Mutations Enriched in Transposase Accessible Mitochondrial DNA." *eLife*. https://doi.org/10.7554/elife.45105.
Xu, Wei, Hui Yang, Ying Liu, Ying Yang, Ping Wang, Se-Hee Kim, Shinsuke Ito, et al. 2011. "Oncometabolite 2-Hydroxyglutarate Is a Competitive Inhibitor of α-Ketoglutarate-Dependent Dioxygenases." *Cancer Cell* 19 (1): 17–30.
Yeaton, Anna, Geraldine Cayanan, Sanam Loghavi, Igor Dolgalev, Emmett M. Leddin, Christian E. Loo, Hedieh Torabifard, et al. 2022. "The Impact of Inflammation-Induced Tumor Plasticity during Myeloid Transformation." *Cancer Discovery* 12 (10): 2392–2413.
Ye, Kai, Marcel H. Schulz, Quan Long, Rolf Apweiler, and Zemin Ning. 2009. "Pindel: A Pattern Growth Approach to Detect Break Points of Large Deletions and Medium Sized Insertions from Paired-End Short Reads." *Bioinformatics* 25 (21): 2865–71.
Yen, K., J. Travins, F. Wang, M. D. David, E. Artin, and K. Straley. 2017. "AG-221, a First-in-Class Therapy Targeting Acute Myeloid Leukemia Harboring Oncogenic IDH2 MutationsAG-221 Therapy for IDH2-Mutant AML." *Cancer Discovery*. https://aacrjournals.org/cancerdiscovery/article-abstract/7/5/478/6259.
Yizhak, Keren, François Aguet, Jaegil Kim, Julian M. Hess, Kirsten Kübler, Jonna Grimsby, Ruslana Frazer, et al. 2019. "RNA Sequence Analysis Reveals Macroscopic Somatic Clonal Expansion across Normal Tissues." *Science* 364 (6444). https://doi.org/10.1126/science.aaw0726.
Yokoyama, Akira, Nobuyuki Kakiuchi, Tetsuichi Yoshizato, Yasuhito Nannya, Hiromichi Suzuki, Yasuhide Takeuchi, Yusuke Shiozawa, et al. 2019. "Age-Related Remodelling of Oesophageal Epithelia by Mutated Cancer Drivers." *Nature* 565 (7739): 312–17.
Yoon, James J., Hongzhi Miao, Trupta Purohit, Dong Chen, Tomasz Cierpicki, and Jolanta Grembecka. 2022. "Combination of Menin and IDH Mutant Inhibition in Acute Myeloid Leukemia." *Blood* 140 (Supplement 1): 5946–5946.
Yoshimi, Akihide, Kuan-Ting Lin, Daniel H. Wiseman, Mohammad Alinoor Rahman, Alessandro Pastore, Bo Wang, Stanley Chun-Wei Lee, et al. 2019. "Coordinated Alterations in RNA Splicing and Epigenetic Regulation Drive Leukaemogenesis." *Nature* 574 (7777): 273–77.
Young, Andrew L., Grant A. Challen, Brenda M. Birmann, and Todd E. Druley. 2016. "Clonal Haematopoiesis Harbouring AML-Associated Mutations Is Ubiquitous in Healthy Adults." *Nature Communications* 7 (August): 12484.
Young, Andrew L., R. Spencer Tong, Brenda M. Birmann, and Todd E. Druley. 2019.

"Clonal Hematopoiesis and Risk of Acute Myeloid Leukemia." *Haematologica* 104 (12): 2410–17.
Zeng, Andy G. X., Suraj Bansal, Liqing Jin, Amanda Mitchell, Weihsu Claire Chen, Hussein A. Abbas, Michelle Chan-Seng-Yue, et al. 2022. "A Cellular Hierarchy Framework for Understanding Heterogeneity and Predicting Drug Response in Acute Myeloid Leukemia." *Nature Medicine* 28 (6): 1212–23.
Zeng, Peiting, Wenhua Lu, Jingyu Tian, Shuang Qiao, Jiangjiang Li, Christophe Glorieux, Shijun Wen, Hui Zhang, Yiqing Li, and Peng Huang. 2022. "Reductive TCA Cycle Catalyzed by Wild-Type IDH2 Promotes Acute Myeloid Leukemia and Is a Metabolic Vulnerability for Potential Targeted Therapy." *Journal of Hematology & Oncology* 15 (1): 30.
Zhang, Xiaotian, Xinyu Wang, Xue Qing David Wang, Jianzhong Su, Nagireddy Putluri, Ting Zhou, Ying Qu, et al. 2020. "Dnmt3a Loss and Idh2 Neomorphic Mutations Mutually Potentiate Malignant Hematopoiesis." *Blood* 135 (11): 845–56.
Zhao, Lin-Pierre, Maxime Boy, Célia Azoulay, Emmanuelle Clappier, Marie Sébert, Ludivine Amable, Jihene Klibi, et al. 2021. "Genomic Landscape of MDS/CMML Associated with Systemic Inflammatory and Autoimmune Disease." *Leukemia* 35 (9): 2720–24.
Zhao, Lin-Pierre, Berenice Schell, Marie Sébert, Rathana Kim, Pierre Lemaire, Maxime Boy, Stephanie Mathis, et al. 2021. "Prevalence of UBA1 Mutations in MDS/CMML Patients with Systemic Inflammatory and Auto-Immune Disease." *Leukemia* 35 (9): 2731–33.
Zheng, Grace X. Y., Jessica M. Terry, Phillip Belgrader, Paul Ryvkin, Zachary W. Bent, Ryan Wilson, Solongo B. Ziraldo, et al. 2017. "Massively Parallel Digital Transcriptional Profiling of Single Cells." *Nature Communications* 8 (January): 14049.
Zheng, Shiwei, Efthymia Papalexi, Andrew Butler, William Stephenson, and Rahul Satija. 2018. "Molecular Transitions in Early Progenitors during Human Cord Blood Hematopoiesis." *Molecular Systems Biology* 14 (3): e8041.