Spring 5-15-2019

The Role of Smc3 in Mouse Embryonic and Adult Hematopoiesis

Tianjiao Wang
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST. LOUIS

Division of Biology and Biomedical Sciences
Molecular Cell Biology

Dissertation Examination Committee:
John Welch, Chair
Grant Challen
Timothy Ley
Matthew Walter
Katherine Weilbaecher

The Role of Smc3 in Mouse Embryonic and Adult Hematopoiesis

by

Tianjiao Wang

A dissertation presented to
The Graduate School
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

May 2019
St. Louis, Missouri

Table of Contents

List of Figures ..... **iv**

List of Tables ..... **vi**

Acknowledgements ..... **vii**

Abstract of the Dissertation ..... **xi**

Chapter 1: Introduction **1**

1.1. Acute Myeloid Leukemia **2**

1.2 Cohesin in Cancer **16**

Figure Legends **25**

Figures **26**

Tables **28**

References **30**

Chapter 2: *Smc3* is required for mouse embryonic and adult hematopoiesis **42**

Abstract **43**

Introduction **43**

Methods **45**

Results **50**

Discussion **56**

Figure Legends **62**

Figures **67**

References **79**

Chapter 3: Exome analysis of treatment-related AML after APL suggests secondary evolution 84

Figure Legends 90

Figures 91

Tables 92

References 94

List of Figures

Chapter 1

Figure 1.1. Eight functional categories of genes that are frequently mutated in AML..... 26

Figure 1.2. The cohesin complex..... 27

Chapter 2

Figure 2.1. Generation of *Smc3* conditional deficient mice and allele validation 67

Figure 2.2. Embryonic hematopoietic *Smc3* deletion 68

Figure 2.3. Homozygous somatic *Smc3* deletion 69

Figure 2.4. Hematopoeitic *Smc3* haploinsufficiency 70

Figure 2.5. Competitive transplantation of *Smc3* haploinsufficient bone marrow cells 71

Figure 2.6. Effect of *Dnmt3a* haploinsufficiency on competitive disadvantage in *Smc3* haploinsufficient BM cells 72

Figure 2.7. Splicing analysis of exon 3 to exon 5 in wild-type and *Smc3*$^{fl/+}$*/Vav1-Cre*$^{+/-}$ KL cells 73

Figure 2.8. Representative plot of intracellular flow cytometry data (Figure 2.1.D-E) 74

Figure 2.9. Analysis of homozygous somatic *Smc3* deletion..... 75

Figure 2.10. Analysis of germline heterozygous *Smc3* deletion 76

Figure 2.11. Immunophenotypic analysis of colonies in serial replating assay *ex vivo* (Figure 2.4.C-E) 77

Figure 2.12 Competitive transplantation of *ERT2-Cre*$^{+/-}$ bone marrow cells..... 78

Chapter 3

Figure 3.1. Exome analysis of patient 10DD-1029 91

List of Tables

Chapter 1

Table 1.1. Core subunits and regulatory proteins of the cohesin complex 28

Table 1.2. Studies of cohesin mutations in hematopoiesis 29

Chapter 2

Chapter 3

Table 3.1. Clinical data of the patient 92

Table 3.2. APL patients with cytogenetic abnormalities in chromsomes 5 or 7 and evaluable subclonal architecture 93

Acknowledgements

First and foremost, I would like to thank my thesis advisor, Dr. John Welch. This research could not have been completed without his continued guidance, support, and patience. I learned so much from John's enthusiasm and dedication to improving treatment options for leukemia patients. He has been extremely encouraging over the years, especially in the last year of my graduate studies when I encountered difficulties. Thank you, John for believing in me when I had doubts and for cheering me all the way to the finish line.

I would like to express my gratitude to my wonderful thesis committee, Dr. Timothy Ley, Dr. Matthew Walter, Dr. Katherine Weilbeacher, Dr. Grant Challen. Tim is the best chair anyone could ask for. Thank you, Tim for all the great scientific discussion, kind advice, and for always keeping us on track. I rotated with Matt and began my journey on the floor in his lab. Thank you, Matt for introducing me to hematopoietic malignancies research and for providing insights whenever I have questions. Also, thank you for letting me use your printer all these years and for inviting me to your lab potlucks, cake sessions, and BBQ. Kathy was a huge support when I got scooped. Thank you, Kathy for teaching me to treat the incidence as a learning experience and for bringing your delightful enthusiasm for science to every committee meeting. I often run into Grant either in the mouse house or on my way to the mouse house. Thank you, Grant for being a role model of working hard and for giving me helpful suggestions on experiments.

Many thanks to members of the Welch lab. Thank you, Brandi for helping me with the project and for becoming my close friend. I miss the time we spend in and outside of lab together. Thank you for letting me be a part of Bradin's life. It was a surreal experience welcoming him into the

world. Thank you, Orsola, my Italian sister. I wish you had come to the lab earlier and we would have more time together. It was a true pleasure getting to know you and being your friend. I love our coffee chats. Thank you for shinning on me when the days were dark and for always caring about me. Thank you, Gayla for keeping our lab organized and for helping me with cloning Smc3. I have learned so much from your work ethic and your superb note-keeping skills. Thank you, Debbie and Conner for taking care of the mice over the years. I would like to also thank my past students Marti and David for the opportunity to mentor and learn from them.

I have been so lucky to be in a nurturing environment for my thesis research. The 6th and 7th floor of the Southwest tower is such a collaborative community. I want to thank Nikki for helping me with the cDNA library prep, Drs. Dave Spencer and Chris Miller for analyzing sequencing data, Julie for injecting mice for all my transplant experiments. I thank Bill, Dan, Matt at the Siteman Flow Cytometry Core for their expertise and help with my project. I am grateful for the instrumental support and guidance provided by Stacy and the MCB directors Drs. Jason Weber and Heather True. I am also thankful for the Siteman Cancer Center Cancer Biology Pathway funding that helped to make this work possible.

I would like to thank all the friends I have met during graduate school. Special thanks to Peter for being my person, whom I can always talk to and count on, and for going through the PhD years together. I also thank Brian, Ling, Tanzir, Monique, KK, Cara, Neal, Hamza for help with my experiments, words of wisdom, and great memories. Thank you to my friends from the entering class of 2012 - Kyle, Linxuan, Bhavna, Melissa, Sarem, Terin for making this journey so memorable.

I would like to thank my families away from home, the Capozzola's and the Serlin's. Thank you, Kate, Shannon, Mr & Mrs. Capozzola, Grandma and Grandpa Toomey for taking me in as your sister, daughter, and granddaughter and loving me for the past 10 years. Thank you, Robert for all the wonderful conversations and career help. Guillemette, Felix, Lukas, Navit, and Marec, I love all of you guys. Words cannot express my gratitude to my dear friends, Mo, Jeff, Fujun, Sicong, Huakang, and Qianwen. Thank you for the late nights, early mornings, and the countless messages. Thank you for all your love and support.

Finally, I would like to thank my family. Heartfelt thanks to my grandparents for being my biggest cheerleaders and for giving me strength. Thank you, Dad for the witty comfort and the unconditional support. Thank you, Mom for all the sacrifices you have made to help me achieve my goals, for taking care of me and keeping me company. Thank you, Dumpling, my dear cat for finding me on the street 2 years ago and following me home. You have been the source of my happiness. Thank you for always being with me.

Tianjiao Jephne Wang

Washington University in St. Louis

May 2019

Dedicated to my parents and grandparents

This work is impossible without your love and unwavering support

And to my beloved cat, Dumpling

Who is always with me, for better, or for worse

The Role of Smc3 in Mouse Embryonic and Adult Hematopoiesis

by

Tianjiao Wang

Doctor of Philosophy in Biology in Biomedical Sciences

Molecular Cell Biology

Washington University in St. Louis, 2019

Dr. John Welch, Chair

Acute myeloid leukemia (AML) is a heterogeneous disease, characterized by recurrent genetic mutations. Mutations in the cohesin complex are one of the 8 functional categories of mutations in AML. *SMC3* encodes a subunit of the cohesin complex, which has important roles in chromosome segregation, genome instability, and gene expression. In the first chapter of the dissertation, we discuss the genetics of AML, normal functions of the cohesin complex, and the interplay between cohesin mutations and myeloid malignancies.

SMC3 is recurrently mutated in AML and other myeloid malignancies. In the second chapter of the dissertation, we compare the consequences of *Smc3* deficient and haploinsufficient mouse models to determine whether the heterozygous missense mutations in *SMC3* might have dominant-negative effects or phenocopy loss-of-function effects. We found that homozygous deletion of *Smc3* during embryogenesis or in adult mice resulted in hematopoietic failure. *SMC3* missense mutations are therefore unlikely to be associated with simple dominant negative phenotypes due to incompatibility with hematopoiesis. *Smc3* haploinsufficiency, in contrast, was

tolerated during embryonic and adult hematopoiesis. Under steady-state conditions, *Smc3* haploinsufficiency did not alter colony forming capacity *ex vivo* and led to modest transcriptional and chromatin accessibility changes in Lin-cKit+ progenitor cells. However, following tamoxifen-induced deletion in competitive transplantations, we observed a significant hematopoietic competitive disadvantage in *Smc3* haploinsufficient bone marrow cells across myeloid and lymphoid lineages and within the stem/progenitor compartments. The competitive disadvantage was not affected by different conditions of hematopoietic stresses, but was partially abrogated by concurrent *Dnmt3a* haploinsufficiency, suggesting that antecedent mutations may be the prerequisites to realize the leukemogenic potential of *Smc3* mutations.

In the third chapter of the dissertation, we present a case of an older women that initially appeared to be treatment-related AML following non-cytotoxic all-trans retinoic acid (ATRA)/arsenic trioxide (ATO) therapy for acute promyelocytic leukemia (APL), but upon further analysis found to be more consistent with secondary AML. Exome sequencing revealed a *TET2*-mutated dominant clonal process that preceded the APL diagnosis, persisted, and gave rise to an AML-associated new subclone with a *NPM1* mutation. Review of additional cytogenetic abnormalities observed in APL patients showed that cytogenetic abnormalities commonly occur as subclones of the APL clone, although one rare case with del(7) independent of the APL clone was identified. These results demonstrated that APL may emerge within the context of clonal hematopoiesis and caution must be exercised when interpreting the development of tAML after ATRA/ATO therapy, especially in older patients.

Chapter 1:

Introduction

1.1 Acute Myeloid Leukemia (AML)

1.1.1 Disease statistics

Acute myeloid leukemia is an aggressive myeloid neoplasm characterized by accumulation of myeloblasts in the blood or bone marrow.[1] Proliferating immature myeloblasts impair the development of normal hematopoiesis, leading to severe infections, cytopenias, anemia, immune compromise, and death.[2] AML is the most common acute leukemia in adults, with 19,520 estimated new cases in 2018, accounting for 1.1% of all new cancer cases in the US.[3, 4] AML is slightly more common among men than women, and approximately 0.5% of the population will be affected at some point during their lifetime based on 2013-2015 data.[4, 5] Although AML can occur in any age group, AML is primarily a disease of the elderly, with a median age at diagnosis of 68 years.[5, 6] Advances in the treatment of AML have significantly improved the outcomes for younger adult patients, with 5-year survivals of 35 to 40% among those who are 60 years of age or younger.[7] However, prognosis in older patients, who account for the majority of new cases, remains dismal, with 2-year survivals of only 5 to 15% among patients who are older than 60 years of age, as much as 70% of the elderly will die within 1 year of diagnosis.[7, 8] Across all age groups, the 5-year overall survival of AML is 27.4%, with an estimated 10,670 deaths in 2018, consisting of 1.8% of all cancer deaths in the US.[4]

1.1.2 Genetics

1.1.2.1 AML with recurrent genetic abnormalities

AML is a heterogeneous disease. The cytogenetic heterogeneity of AML has been recognized for more than three decades. Based on karyotype analysis, AML with recurrent genetic

abnormalities can be divided into two subtypes: (1) AML with chromosomal aneuploidies; (2) AML with balanced genomic rearrangements.[9]

AML with chromosomal aneuploidies

Over 60% of cases in the subgroup of AML with chromosomal aneuploidies have at least 3 chromosomal events, of which the most frequent are -5/5q, -7/7q, -12/12p, -17/17p, and +8/8q. Approximately 50% of patients with deletions in chromosomes 5, 7, 12, or 17 have *TP53* mutations, and these are more commonly observed in older patients; the median age of patients with chromosomal aneuploidy and *TP53* mutations is 58 years vs. 49 years with aneuploidy alone.[9] Patients with both complex karyotype and *TP53* mutations have significantly inferior prognosis to the poor overall outcomes associated with either subset alone, but recent data by groups at Washington University in St. Louis suggest that this unique subgroup of AML may respond favorably to hypomethylating agent, decitabine.[10]

AML with balanced genomic rearrangements

AML with balanced genomic rearrangements tend to present at a younger age and have, on average, 1 genomic rearrangement and lower overall number of acquired mutations, most frequently concurrent with activating mutations *FLT3*- internal tandem duplication (ITD), *KIT*, *NRAS*, tyrosine or serine-threonine kinases, and protein tyrosine phosphatases.[11] There are at least 7 distinct subtypes of recurrent genomic rearrangements in AML, each defining a clinicopathologic entity.[12] These translocations and inversions are considered leukemia-initiating and are almost uniformly present in patients who subsequently relapse.

The most common translocation fusion gene is *PML-RARA*, defined by t(15;17)(q22;q21), which occurs in 5-13% of patients and is characteristic of acute promyelocytic leukemia (APL). *FLT3*-

ITD and *WT1* mutations co-occur with *PML-RARA* in approximately 35% and 15% of APL cases, respectively. APL patients with *PML-RARA* that are *FLT3* negative are associated with favorable outcomes when treated with combinational chemotherapy that includes all-*trans*-retinoic acid (ATRA).[13] Outcomes in APL patients treated with chemotherapy alone were historically dismal, demonstrating the adaptive relevance of mutation: treatment interactions.

RUNX1-RUNXIT1 AML, defined by t(8;21)(q22;q22.1), occurs in 1-6% of patients, is associated with good risk following treatment with high dose cytarabine. *KIT* mutations co-occur with t(8;21) in approximately 25% of *RUNX1-RUNXIT1* AML, and these patients have inferior outcomes compared to *KIT* wild type patients. [13, 14]

The *CBFB-MYH11* fusion results from inv(16)(p13.1q22) and occurs in 1-6% of AML patients. *CBFB-MYH11* AML also has favorable prognosis in the absence of *KIT* mutations. *NRAS* mutations co-occur with *CBFB-MYH11* in approximately 40% of AML cases. The less frequent genomic rearrangements, affecting about 1% or less of AML patients include: *MLLT3-KMT2A*, defined by t (9;11)(p21.3;q23.3); *DEK-NUP214*, defined by t(6;9)(p23;q34.1); *GATA2*, *MECOM*, defined by inv(3)(q21.3q26.2), and *RBM15-MKL1* ,defined by (t1;22)(p13.3;q13.3).[9, 13]

1.1.2.2 AML with gene mutations

Over the past 15 years, advances in next-generation sequencing (NGS) have tremendously increased our knowledge of the molecular heterogeneity of AML. AML was the first primary cancer to be studied by massively-parallel sequencing technologies.[15] In 2008, the first AML genome was published in a landmark study done by groups at Washington University in St. Louis.[16] Subsequent studies have identified numerous novel recurrent somatic mutations with

biologic, prognostic, and therapeutic relevance and have demonstrated that AML is a complex and dynamic disease. Emerging data with the use of NGS are revolutionizing our view of the spectrum and frequency of mutations, their distinct patterns of cooperativity and mutual exclusivity, their subclonal architecture, the epigenetic landscape of the disease, and the clonal evolution during AML.[17]

Clonal Evolution

Studies have shown that most cases of AML are characterized by clonal heterogeneity at the time of diagnosis, with more than half of the patients exhibiting at least one subclone in addition to a founding clone.[11] Data from clonal evolution studies provide support for a model that mutations in genes involved in epigenetic regulation (specifically genes involved in the regulation of DNA methylation and chromatin modifications, most commonly *DNMT3A*, *TET2*, *IDH1*, *IDH2*, and *ASXL1*) are present in preleukemic hematopoietic stem cells (HSCs) and occur early in the evolution of AML, preceding secondary leukemogenic events such as mutations in nucelophosmin (*NPM1*) or signaling genes (*FLT3*, *RAS*).[13, 17] Furthermore, the epigenetic modifying genes are frequently found to be mutated in elderly individuals along with clonal expansion of hematopoiesis that confers and increased risk for the development of hematologic cancers.[13] Such ancestral preleukemic stem cells are capable of multilineage differentiation. For example, preleukemic *DNMT3A*-mutant HSCs were shown to have a multilineage repopulation advantage over wild type HSCs and were detected in samples collected from patients who were in morphologic complete remission, indicating their potential to be resistant and survive chemotherapy.[18] Thus, preleukemic hematopoietic clones can persist over time, survive chemotherapy, expand during remission, and eventually leading to relapse and the various patterns of clonal composition that occur at relapse may contribute to resistance to therapy.[19]

Clonal Hematopoiesis

Recent studies of large population-based cohorts show that clonal hematopoiesis with recurrent mutations in epigenetic regulators *DNMT3A, TET2,* and *ASXL1* (and less frequently in splicing factor genes *SRSF2, SF3B1* and in the genotoxic sensor *TP53*) increases as people age and confers an increased risk of hematologic cancer and death.[20, 21, 22, 23, 24] Expanded clones containing these somatic mutations can be identified in the blood or bone marrow of patients without evidence of overt hematologic malignancy and decades before the development of AML. This defines a new entity, termed either "age-related clonal hematopoiesis" (ARCH) or "clonal hematopoiesis of indeterminate potential (CHIP)", which has been identified in approximately 10% of patients 70-80 years old.[22, 23] A recent study conducted by groups at Washington University in St. Louis using bar-coded sequencing found a higher incidence of ARCH if the threshold of detection is lowered to 0.5%.[25] The incidence of CHIP increases with age, predisposes patients to AML and other hematologic malignancies, including myelodysplastic syndromes (MDS), and the transformation rate of CHIP into a hematologic disease is about 0.5-1% per year.[26]

The Cancer Genome Atlas (TCGA) Project

The Cancer Genome Atlas Research Network analyzed the mutational profiling of 200 patients with de novo AML by either whole-genome (n=50) or whole-exome (n=150) sequencing, along with RNA and microRNA expression and DNA methylation analysis.[11] Significantly mutated genes in AML were organized into 8 functional categories, summarized in Figure 1.1[17]. (1) Mutations in *NPM1*, encoding a multifunctional nucleo-cytoplasmic shuttling protein, resulting in the aberrant cytoplasmic localization of NPM1 and NPM1-interacting proteins; (2) Mutations in signaling genes such as kinases *FLT3, KIT,* or RAS family members *KRAS, NRAS* that confer

a proliferative advantage through the RAS-RAF, JAK-STAT, and PI3K-AKT signaling pathways; (3) Mutations in myeloid transcription factors such as *RUNX1* and *CEBPA*, leading to transcriptional deregulation and impaired hematopoietic differentiation; (4) Mutations in tumor-suppressor genes such as TP53 and WT1 that result in transcriptional deregulation and impaired degradation through the mouse double minute 2 homologue (MDM2) and the phosphatase and tensin homologue (PTEN); (5) Mutations in DNA methylation-associated genes *DNMT3A* and *TET2* that deregulate DNA methylation patterns and lead to transcriptional deregulation of leukemia-associated gens or in *IDH1* and *IDH2* that act through the 2-hydroxyglutarate (2-HG) oncometabolite production and impact DNA methylation via impairment of TET2; (6) Mutations in chromatin-modifying genes such as *AXL1* and *PHF6*, leading to deregulation of chromatin modification, for instance methylation of histones H3 and H2A; (7) Mutations in spliceosome-complex genes such as *SRSF2*, *SF3B1*, *U2AF1*, and *ZRSR2* that are involved in impaired spliceosome function and deregulated RNA processing; (8) Mutations in cohesin-complex genes such as *SMC3*, *STAG2*, and *RAD21* that may impair accurate chromosome segregation and transcriptional regulation.[17]

NPM1 mutations

NPM1 encodes a phosphoprotein that normally shuttles between the nucleus and the cytoplasm and plays a role in in epigenetic control, ribosomal protein assembly, and regulation of p53 tumor suppressor pathway.[27] *NPM1* mutations are the most common genetic mutations in AML, found in approximately 30% of all AML and 45-60% of AML with normal karyotype.[17, 28] They are mutually exclusive to other genomic rearrangements and frequently co-exist with *DNMT3A* (approximately 50%), *FLT3*-ITD (approximately 40%), *NRAS* (approximately 20%), cohesin genes *SMC3*, *SMC1A*, *RAD21* (approximately 20%), *TET2* (approximately 15%), *IDH1*

(approximately 15%), *IDH2*^{R140} (approximately 15%) mutations, and *PTPN11* (approximately 15%).[13] Mutations in DNA hydroxymethylation genes (*DNMT3A*, *TET2*, *IDH1*, and *IDH2*) typically represent the first acquired event and are present in the founding clone while *NPM1* is acquired as a secondary event during leukemogenesis, together with mutations in *FLT3, NRAS*, and *PTPN11*. In younger patients (<60 years old), *NPM1* mutations in cytogenetically normal AML without *FLT3*-ITD mutations portend a favorable prognosis.[29] However, patients with concomitant mutations in *NPM1*, *FLT3*-ITD, and *DNMT3A*, which represent the most frequent triple genotype in AML, have significantly shorter event-free survival and inferior overall survival.[9]

Mutations in signaling genes

FLT3 encodes a receptor tyrosine kinase involved in hematopoiesis. There are two common mutations that occur in *FLT3*: ITD in the juxtamembrane domain and a point mutation of the tyrosine kinase domain (TKD), both mutations lead to constitutive activation. Approximately 20% of all AML cases harbor a *FLT3*-ITD mutation, which is associated with an unfavorable prognosis and the mutation is more common in cytogenetically normal AML, accounting for approximately 30% of these cases.[30] The frequency of *FLT3*-ITD mutations decreases with older age and *FLT3* mutations are associated with *NPM1* mutations.[13] There is variability in the size of the *FLT3*-ITD, ranging from a few base pairs to over 1000 base pairs, the number of *FLT3*-ITD mutations, approximately 14-25% of *FLT3*-ITD positive patients will have more than one *FLT3*-ITD mutation.[31, 32] Sequencing of *FLT3*-ITD reveals that the sequence and site of the mutations are variable: in fact, only about two-thirds of the *FLT3* mutations are true tandem duplications while the remaining are insertions or complex duplications and insertions; approximately 30% of *FLT3*-ITD occur outside the juxtamembrane domain and instead occur in the TKD, usually in the

β1 sheet.[32, 33, 34] The less common *FLT3*-TKD mutation is found in approximately 10% of AML.[28]

KIT encodes a receptor tyrosine kinase that plays important roles in proliferation, differentiation, and cell survival. The ligand for KIT is stem cell factor (SCF). Binding of SCF to the extracellular domain of KIT induces receptor dimerization and activation of downstream signaling pathways that are involved in mediating pro-growth and pro-survival signals within the cell, including the MAPK signaling pathway (RAS-RAF-MEK-ERK), the PI3K pathway (PI3K-AKT-mTOR), and the STAT3 pathway.[35] *KIT* mutations are gain-of-function mutations that occur in less than 5% of all AML cases and are higher, 25-35% of cases in core-binding factor leukemia.[2, 30] *KIT* mutations occur primarily in exon 17 and affect the activation loop of the kinase domain, resulting in improved proliferation and survival of leukemic cells.[36] *KIT* mutations confer unfavorable prognosis in AML with t(8;21), *RUNX1-RUNXIT1* AML.[13]

KRAS and *NRAS* belong to the RAS GTPase family of genes. *KRAS* mutations are less common in adults, found in only 2% of cases vs. 9% of cases in children.[37, 38] *NRAS* mutations occur in approximately 15% of AML cases in adults and children.[13] The concurrent mutations of *NRAS* are *NPM1* and biallelic *CEPBA*. RAS mutations do not appear to have a clear impact on outcome except for $NRAS^{G12/G13}$, which confers superior outcomes in presence of *NPM1* and *DNMT3A* mutations.[39]

Mutations in myeloid transcription factors

RUNX1 encodes the alpha subunit of the heterodimer core binding factor, which is involved in transcription.[30] Somatic *RUNX1* mutations occur in 5-20% of AML and the incidence increases with older age.[2] They co-segregate with mutations in *SRSF2* (approximately 25%), *ASXL1*

(approximately 20%), *KMT2A* (15-20%), *IDH2*^{R140} (approximately 12%).[13] They are mutually exclusive with *NPM1*, biallelic *CEBPA*, and AML with recurrent cytogenetic abnormalities.[9] *RUNX1* mutations are associated with male sex, inferior outcome, and secondary AML evolving from MDS.[40] Germline *RUNX1* mutations are found in the autosomal dominant familial platelet disorder, conferring a predisposition to AML.[41]

CEBPA encodes a transcription factor involved in granulocytes differentiation. *CEBPA* mutations are found in approximately 10% of AML and are more common in cytogenetically normal AML or with 9q deletions.[42] The incidence of *CEBPA* mutations declines with older age. Approximately 2/3 of *CEBPA* mutations may be biallelic, which usually include one N-terminus and one C-terminus mutation, leading to null expression of *CEBPA*, and the rest are monoallelic, which can be truncating N-terminal mutations resulting in a shortened CEBPA with a dominant negative effect or C-terminal mutations that decrease dimerization or DNA binding.[43, 44, 45, 46] Biallelic *CEBPA* mutations co-occur with *NRAS* (approximately 30%), *GATA2* (approximately 30%), *WT1* (approximately 20%), *CSF3R* (approximately 20%), and 9q- (approximately 15%), and confer a favorable prognosis.[13]

Mutations in tumor-suppressor genes

TP53 is a tumor suppressor gene and frequently referred to as the "guardian of the genome" that regulates the cell cycle in response to cellular stresses. *TP53* mutations occur in 5-20% of adult AML and approximately 1% of pediatric AML.[13, 37] The incidence of *TP53* mutations significantly increases with older age. *TP53* mutations are predominantly detected in AML with complex karyotype (56-78% of cases) and are associated with very poor outcome in AML as in other cancers.[17, 47]

WT1 encodes a transcription factor important for normal cellular development and cell survival that appears to play a tumor suppressor role in renal tissues, but an oncogenic role in leukemia.[48] WT1 mutations can be found in 4-11% of AML cases and are linked with poor outcome in AML with a normal karyotype.[30]

Mutations in DNA-methylation-associated genes

DNMT3A encodes a DNA methyltransferase involved in the epigenetic regulation of the genome through methylation. *DNMT3A* mutations are quite common in AML, occurring in approximately 20% of patients and frequently co-occur with *NPM1, FLT3*-ITD, *IDH1, IDH2*, and *SMC3* mutations.[17, 49] The most common mutation is a substitution of arginine at position 882 (R882).[50] DNMT3A with heterozygous R882H mutation forms stable heterodimers with wild type DNMT3A, disrupting the ability of the wild type DNMT3A protein to form active tetramers and leading to a hypomorphic effect on the methyltransferase activity of the enzyme and also a dominant negative effect on the wild type DNMT3A.[51, 52, 53] The incidence of *DNMT3A* mutations increases with older age. They are associated with CHIP and secondary AML evolving from MDS and are early events in leukemogenesis. The frequency of *DNMT3A* R882 mutations is less than one-third of CHIP *DNMT3A* mutations, but more than two-thirds of AML *DNMT3A* mutations. *DNMT3A* mutations have moderate adverse effect on outcome, which can be overcome by high doses anthracycline chemotherapy.[47, 38]

TET2 encodes an epigenetic modifier that converts methylcytosine to 5-hydroxymethylcytosine and is also involved in myelopoiesis. *TET2* mutations are found in 7-25% of adult AML and 1.5-4% of pediatric AML and are early events in leukemogenesis. Mutations in *TET2* are highly

variable, including nonsense mutations, deletions, missense mutations, and splice-site mutations, which all appear to cause loss-of-function and decrease hydroxymethylation of DNA.[54] *NPM1* mutations and *TET2* mutations statistically co-occur with *FLT3*-ITD and -TKD aberrations.[55] In contrast, *IDH* mutations seldom co-exist with *TET2* mutations possibly because 2-HG inhibits the activity of *TET2* (see below).[55, 56] The incidence of *TET2* mutations in AML increases with older age and *TET2* mutations have been found in healthy elderly individuals with CHIP.[20]

IDH1 and *IDH2* are genes involved in metabolism and may also play an epigenetic role in histone and DNA methylation.[57] Mutations in *IDH1* and *IDH2* occur at the active isocitrate binding site, which alters the enzymatic activity and leads to the generation of a novel oncometabolite, 2-HG.[58] *IDH* mutations statistically co-occur with *NPM1* mutations (except for $IDH2^{R172}$).[59] They are associated with CHIP in healthy elderly individuals (although much less commonly than *DNMT3A, TET2, ASXL1,* and *TP53* mutations) and are early events in leukemogenesis.[13] *IDH1* mutations affect the arginine at position 132 or 170 (R132 or R170) and can be found in 7-14% of adult AML cases, but only 1% of pediatric AML.[17, 37] These mutations are mutually exclusive and exclusive of the *IDH2* mutations. *IDH1* mutations are associated with unfavorable outcome.[60] *IDH2* mutations affect the arginine at position 140 or 172 (R140 or R172) and occur in 8-19% of adult AML, but only 1-2% of pediatric cases.[17, 61] The incidence of $IDH2^{R140}$ mutation increases with older age and has been shown to have a favorable prognosis in intermediate risk AML with *NPM1* mutations.[38]

Mutations in chromatin-modifying genes

ASXL1 encodes a chromatin binding protein, which regulates gene transcription in localized areas via modifying chromatin structure. *ASXL1* mutations are frequently found in MDS and AML, with a frequency of 5-15%, but appear to be enriched in secondary AML and intermediate

risk AML.[62] *ASXL1* mutations are associated with male sex and CHIP in healthy elderly people and they also increase with older age, more prevalent in patients over 60 years old and quite rare (approximately 1%) in children.[23, 63] Frequent concomitant mutations are *RUNX1* (approximately 20%), *IDH2*^{R140} (approximately 13%), and *SRSF2*.[13] *ASXL1* mutations are early events in leukemogenesis, with most studies showing they are predictive of inferior outcome, particularly genotypes *ASXL1*^{mut}*/RUNX1*^{mut} and *ASXL1*^{mut}*/SRSF2*^{mut}.[17]

PHF6 is an X-linked gene that appears to be a highly dynamic chromatin adaptor protein that interacts with a growing number of partners (nucleosome remodeling and deacetylation complex, PAF1, UBF) to regulate transcription.[64] Germline loss-of-function mutations in *PHF6* are the cause of the Börjeson-Forssman-Lehmann X-linked intellectual disability syndrome.[65] Somatic *PHF6* mutations occur in 2-3% of adult AML and are more frequent in males than females.[66, 67] They are associated with adverse prognosis in intermediate risk AML patients who are negative for *FLT3*-ITD.[38]

Mutations in spliceosome-complex genes

Mutations in splicing factor genes *SRSF2*, *SF3B1*, *U2AF1*, and *ZRSR2* lead to impaired spliceosome function and deregulated RNA processing resulting in aberrant splicing patterns. Mutations in spliceosome-complex genes account for 14% of AML patients in the TCGA cohort.[11] They are associated with CHIP in healthy elderly persons and poor outcome, shown by a few studies on clinical significance.[13] Moreover, mutations of splicing factors occur in high frequencies in MDS. Refractory anemia with ringed sideroblasts (RARS) is a subtype of MDS characterized by the accumulation of erythroid precursor cells and 15% or more ring sideroblasts in the bone marrow. *SF3B1* is highly mutated in RARS, whereas *U2AF1* mutations are not linked with ringed sideroblasts and RARS.[68]

SRSF2 mutations are also found in chronic myelomonocytic leukemia (CMML) and confer an increased risk of transformation to acute leukemia. Mutations can occur in multiple domains, although the most recurrent mutations affect the RNA recognition motif and arginine/serine-rich protein interaction domain of the protein.[69, 70] Functional studies have shown that the P95 SRSF2 mutations have an altered RNA-binding activity resulting in mis-splicing of many important genes including *EZH2*.[71]

SF3B1 is the most commonly mutated spliceosomal gene in hematological cancers, including MDS, with almost half of *SF3B1* mutations in lysine 700. Heterozygous *SF3B1* mutations are mostly missense substitutions in addition to hotspots in the HEAT repeat domains.[71]

U2AF1 is frequently mutated in codons S34 and Q157 in approximately 11% of MDS patients. Heterozygous insertions and deletions have also been reported.[72, 73] *U2AF1* mutations appear to interfere with 3' splice site binding function of the protein, leading to aberrant alternative splicing of numerous U2-dependent introns potentially and constructing an entirely novel transcriptome specific to MDS.[74]

ZRSR2 mutations are distributed throughout the gene in MDS patients, interrupting the coding capacity by creating in-frame stop codons and therefore suggestive of loss-of-function phenotypes.[72] Knockdown of ZRSR2 revealed a distinct splicing defect pattern of the U12-dependent introns, affecting a large number of U12-type intron-containing genes that play a significant role in MAPK signaling pathways and E2F transcription activities, and impaired *in vitro* erythroid differentiation while promoted myeloid differentiation of cord blood-derived CD34+ cells, which supports MDS phenotype.[75]

Mutations in cohesin-complex genes

Mutations in cohesin complex genes *SMC3*, *STAG2*, *RAD21*, and *SMC1A* may cause defects in chromatid cohesion or impact transcriptional regulation. Cohesin mutations occurred in about 10% of non-M3 AML cases and were identified in 13% of AML patients in the TCGA cohort.[11, 49] Cohesin mutations frequently co-occur with *NPM1* mutations and *RUNX1-RUNXIT1*. Other common mutations concurrent with cohesin mutations in AML include *RAS*, *RUNX1*, *TET2*, *ASXL1*, and *EZH2*.[11, 39, 76] Cohesin mutations are not only found in AML, but also in other myeloid malignancies such as CMML, chronic myeloid leukemia (CML), myeloproliferative neoplasms (MPNs), and MDS.[77, 78] Notably, more than 50% of patients with Down syndrome-associated acute megakaryocytic leukemia (DS-AMKL) have mutations in *STAG2*, which can co-occur with mutations in *RAS*, *ASXL1*, *EZH2*, *JAK2*, and *JAK3*.[79]

SMC3 and *RAD21* mutations are nearly universally heterozygous and the majority of *SMC3* mutations are missense mutations. Intriguingly, *SMC3* mutations frequently co-occur with *DNMT3A* mutations, one of the most commonly mutated genes in AML. *STAG2* and *SMC1A* are encoded on the X chromosome, and therefore mutations would be thought to result in null alleles.[80] Additionally, cohesin mutations tend to be mutually exclusive, implying that either they may not be tolerated by a cell when co-occurring or alteration in one component may be sufficient to disrupt the entire complex.[49, 76] Although cohesin mutations are often observed as early subclonal events during leukemia development, conceivably facilitating disease initiation, they are not observed in CHIP; thus, they are unlikely to be the initiating event.[21, 81, 82] In most AML cases, cohesin mutations are not associated with karyotypic abnormalities, suggesting cohesin mutations contribute to leukemogenesis through alternative pathways other than inducing genomic instability.[11, 76]

This thesis focuses on understanding the contribution of cohesin mutations, particularly *SMC3* mutations to the pathogenesis of AML.

1.2 Cohesin in Cancer

1.2.1 Roles of cohesin

The cohesin complex consists of four core subunits, structural maintenance of chromosomes (SMC) proteins SMC1A and SMC3, RAD21, and STAG.[83] In mammals, there are two related STAG proteins, STAG1 or STAG2. Both SMC proteins are rod-shaped proteins containing ATP-binding cassette (ABC)-like ATPase motifs and are characterized by a globular hinge domain flanked by two alpha-helical domains, which fold back on themselves at the hinge, forming a long antiparallel coiled coil arm that brings the N- and C-termini together. SMC1A and SMC3 form a V-shaped heterodimer at the hinge domains. At the distal end of the two coiled coil arms, the N- and C-termini of each SMC protein form an ATPase head domain.[84] The kleisin family protein RAD21 physically connects the ATPase heads of SMC1A and SMC3, thus forming a tripartite ring-like structure, with an internal diameter of about 40nm.[85] The STAG subunit interacts with RAD21 and further stabilizes the cohesin ring. In addition to the four core subunits, cohesin loaders (Scc2/NIPBL, Scc4/MAU2), cohesin regulators (PDS5, SORORIN, and WAPL), cohesin protector (SGOL), and cohesin modifiers (ESCO and HDAC8) also bind to or modify the cohesin complex (Figure 1.2).[86]

The cohesin complex is highly conserved through evolution with homologs in yeast, fruit flies, and mammals (Table 1.1[84]).[87] Among the several models have been proposed to depict how the cohesin complex associates with chromatin, the one-ring "embrace" model and the two-ring "handcuff" model are supported by experimental data.[88, 89] The one-ring model suggests that the

cohesin ring embraces two chromatins until their segregation.[88] The two-ring model describes each cohesin ring entraps one chromatin and cohesion is mediated by interactions between the two cohesin rings.[89] The canonical role of the cohesin complex is to ensure proper segregation of chromosomes during mitosis and meiosis. In addition to its essential role in sister chromatids cohesion, cohesin contributes to genome maintenance and functions by involving in DNA damage repair and gene expression.[90]

1.2.1.1 Cohesin functions in chromosome segregation

Cohesins are loaded to the chromatins at the G1/S phase in yeast and at telophase in mammalian cells by loading complex NIPBL-MAU2.[91, 92, 93, 94, 95] During DNA replication at S phase, each cohesin ring embraces one of the sister chromatids. After DNA replication at the S phase, acetylation of SMC3 by cohesin acetyltransferases ESCO1 and ESCO2 establishes stable cohesion between the newly replicated sister chromatids. PDS5 and SORORIN form a complex to maintain the cohesion throughout the G2 phase until prophase when SORORIN is phosphorylated and destabilized.[96, 97, 98, 99] The removal of cohesins are facilitated by the formation of PDS5-WAPL complex. At prophase, cohesins on the chromosomal arms are removed by the phosphorylation of RAD21 and STAG1/2 by PLK1.[100] Centromeric cohesion is protected by SGOL1 until Separase gets activated and cleaves RAD21 at anaphase and therefore separating the sister chromatids.[101, 102] The dissociated cohesins can be recycled after the acetyl groups are removed from SMC3 proteins by cohesin deacetylases HDAC8.[103, 104] In meiosis, a similar biphasic removal of cohesin occurs, with RAD21 replaced by REC8.[105]

The main roles of cohesin during cell cycle are to keep sister chromatids together and to provide resistance when sister chromatids are pulled by microtubules towards the opposing spindle poles, thus ensuring accurate separation of sister chromatids during the transition from metaphase to

anaphase.[106] Failure in the formation and maintenance of sister chromatid cohesion results in premature chromosome segregation, which is thought to be a major pathway to aneuploidy, a characteristic observed in many human cancers.[107]

1.2.1.2 Cohesin functions in genome instability

Cellular DNA is exposed to single and double strand breaks (DSBs) through multiple endogenous and exogenous mechanisms. Cells respond to DNA damage by activation of DNA-damage checkpoints that halt cell cycle progression until the damaged DNA is repaired. If the damage cannot be repaired properly, cells may undergo apoptosis. Eukaryote cells have two distinct mechanisms to repair DSBs, the homologous recombination (HR) between sister chromatids in the S and G2 phases and the non-homologous end joining (NHEJ), involving re-ligation of broken DNA, which occurs throughout the cell cycle.[84]

The function of cohesin in DNA damage repair is evolutionarily conserved from yeast to humans.[108, 109, 110] Rad21 was cloned originally by complementing the γ-radiation sensitivity in fission yeast with a function in DSB repair, before its role in sister chromatid cohesion was identified.[111] In response to laser-induced DNA damage in human cells, cohesins are recruited to the DSB site and *de novo* cohesion, named damage-induced cohesion (DI-cohesion), is established.[112] Besides cohesins, factors that are required to load cohesins to chromatin, establish cohesion, and maintain cohesion are all needed for DNA damage repair. Defects in the cohesin-loading complex NIPBL-MAU2, cohesin acetyltransferase ESCO, or maintenance factor SORORIN block the accumulation of cohesins at DSBs and prevent DNA damage repair, suggesting the presence of cohesins on chromatin is not sufficient to mediate DNA repair and instead, additional cohesion is required.[113, 114, 115] DI-cohesion may help to structurally stabilize chromosomes whose DNA backbone has been fragmented by DSBs and to provide the proximity

between the damaged sister chromatid and the template, allowing HR to occur. DNA damage-induced phosphorylation and acetylation on SMC3 were found to be important for genome-wide DI-cohesion caused by DSB in the G2 phase and DSB repair.[112]

Moreover, cohesins are required to activate checkpoints when DSBs occur.[116] In *C. elegans*, when SCC2, a component of the cohesin loading complex, is mutated, cohesins cannot be loaded onto chromatin in meiosis, resulting in failure of both checkpoint activation and DNA damage repair.[117] This shows the importance of cohesin recruitment to the damaged chromatin. The checkpoint role of cohesins is independent of its function in sister chromatid cohesion because cohesins are required for the phosphorylation and activation of Chk2 although no sister chromatid cohesion occurs in the G1 phase. As evidence, depletion of SORORIN, a protein essential for the generation and maintenance of sister chromatid cohesion, leads to checkpoint activation but DSB repair failure.[116]

1.2.1.3 Cohesin functions in gene expression

The first evidence that cohesin factors regulate gene expression and development came from the studies of *Drosophila cut* and *Ultrabithorax* genes: heterozygous *Nipped-B* mutants showed reduced *cut* expression, whereas loss of Smc1, Rad21, or SA led to increased *cut* expression.[118, 119] Cohesins also facilitate expression of *c-myc*, a function conserved across *Drosophila*, zebrafish, mouse, and humans and cohesin depletion reduces *myc* transcription.[120] Furthermore, cohesins present in non-cycling and even post-mitotic cell in higher eukaryotes.[121] Accumulating evidence implies an important non-canonical role of cohesin in regulating gene expression, which is independent of cohesins' role in cell division.[83] In non-dividing mouse thymocytes, genetic deletion of cohesin resulted in reduced transcription and rearrangements at the T cell receptor, thereby affecting thymocyte differentiation.[122]

Cohesins have been shown to mediate long-range transcriptional regulation by controlling the spatial conformation of chromatin at multiple gene loci that are important for normal development and differentiation.[123] Studies revealed two distinct types of cohesin sites: sites that coincide with the binding of CTCF (CCCTC-binding factor) vs. sites that map to active enhancers and promoters and are usually cell-type specific. The CTCF-dependent interaction of cohesins with insulator blocks enhancer activity and disrupts distal enhancer-promoter interactions required for gene activation.[121] Moreover, cohesin has a CTCF-independent role in tissue-specific transcriptional regulation.[124] ChIP-Seq data suggest that cohesins co-localize with master regulators in several tissues, such as liver-specific transcription factors in HepG2 cells and estrogen receptor α in MCF-7 cells.[124, 125] Cohesin also co-localize with transcriptional coactivators, such as mediator to facilitate chromatin looping between the enhancer and promoter of some pluripotency genes (e.g POU5F1) in mouse embryonic stem cells.[126] The cohesin complex lacks a definitive DNA-binding domain. Therefore, DNA localization appears to be facilitated through binding to CTCF and transcription factors, thus forming a regulatory network for transcriptional programs of specific cell type.[121]

Cohesins play an essential role in the maintenance of pluripotency. Depletion of cohesins blocks self-renewal, induces spontaneous differentiation, and interferes with reprogramming of fibroblasts to pluripotent cells.[127] Mutations in core components of the cohesin complex can cause developmental defects in a number of species. For instance, heterozygous mutations in cohesin loader NIPBL or less frequently, in cohesin subunits SMC1A and SMC3 result in Cornelia de Lange syndrome (CdLS), a neurodevelopmental disorder with upper extremity malformations.[90]

1.2.2 Cohesin deregulation in cancer

Mutations of cohesins have been found in many cancers including leukemias[19, 49, 128], colorectal carcinomas[129], ovarian carcinomas[11, 130], glioblastoma, melanomas, and Ewing's sarcomas[131]. The first somatic mutations of cohesin in cancer were reported in 2008 when heterozygous missense mutations in *SMC1A, SMC3, STAG3* (a component of meiotic cohesin) and *NIPBL* were identified in aneuploid colorectal cancers.[129] In 2010, deletions of *RAD21* in a CML and deletions of *STAG2* in an AML were reported.[132] In 2011, *STAG2* mutations were reported to result in cohesion defects and aneuploidy in glioblastoma cell lines, melanomas, and Ewing's sarcomas.[131] *STAG2* is the most frequently mutated gene of the cohesin complex. Because *STAG2* is located on the X chromosome, only a single mutational event is required to inactivate it in both males and females (due to X inactivation). *STAG2* mutations are considered loss-of-function mutations because: 1) the majority of mutations are truncating, 2) truncating mutations are present in early exons, resulting in a very short protein, 3) in many cases a truncated STAG2 protein is absent, likely due to nonsense-mediated decay of the mutant *STAG2* mRNA.[133, 134] In 2013, three studies reported frequent somatic mutations of *STAG2* in bladder cancer.[135, 136, 137] In addition, SMC1A has been shown to be overexpressed in gliomas and reducing its levels inhibits glioma cell growth *in vitro*.[138] Upregulation of ESCO2 and WAPL is associated with tumor progression in melanomas and cervical cancer, respectively.[139, 140] Overexpression of Separase is sufficient to induce tumorigenesis in mammary epithelial cells in a *TP53*-mutant background.[141] Pan-cancer analysis of the TCGA data found that the cohesin complex was recurrently mutated across 12 cancer types and identified the cohesin complex as one of the 16 significantly mutated subnetworks.[142]

1.2.2.1 Cohesin mutations in myeloid malignancies

Cohesin mutations in AML

Recurrent mutations in all four members of the cohesin complex, *SMC1A*, *SMC3*, *RAD21*, and *STAG2* were first identified in M1 AML cases. They co-occurred with *NPM1*, *RUNX1*, *DNMT3A*, or *TET2* mutations in 17/19 cases, indicating cooperation with other leukemogenic pathways.[49] Subsequently, the TCGA data confirmed and extended these results, identifying a cumulative cohesin mutation frequency of 13% (26/200).[11] The frequency of cohesin mutations was assessed in both *de novo* and secondary AML. One study showed a higher frequency of cohesin mutations in *de novo* AML (13%, 16/120 samples studied) than in secondary AML (8%, 3/37 samples studied).[77] In contrast, another study showed a higher frequency of cohesin mutations in secondary AML (20%, 30/149) than in *de novo* AML (11%, 32/301).[78] In both studies, the most frequently mutated cohesin gene was *STAG2*, followed by *SMC3* and *RAD21*.[77, 78] Cohesin mutations are nearly always mutually exclusive and are mostly found in samples with normal karyotypes. Based on allelic burden analysis, cohesin mutations are often, but not always, observed as early event during leukemogenesis.[76]

Cohesin mutations in MDS

STAG2 mutations were identified in MDS samples from patients whose disease later progressed to secondary AML.[143] Subsequent studies showed that 7% (10/150) of MDS samples harbor *STAG2* mutations and 8% (18/224) of MDS samples harbor cohesin mutations, the majority of which were *STAG2* mutations, with lower mutation frequencies in *SMC3* and *RAD21*.[77, 144] Cohesin mutations were also identified in 17% of high-risk MDS samples and 11% of low-risk samples, respectively.[78]

Cohesin mutations in DS-AMKL

Down syndrome is associated with trisomy 21, and individuals with trisomy 21 are more susceptible to hematologic abnormalities. Up to 10% of children with Down syndrome will

present with transient abnormal myelopoiesis at birth, a necessary predecessor to DS-AMKL. Virtually all DS-AMKL patients have an inactivating mutation in *GATA1*, which results in exclusive expression of a shorter isoform, named GATA1s.[80]

Cohesin mutations are prominent in DS-AMKL and are predicted to be heterozygous, loss-of-function, and early events during leukemia development. Deep sequencing revealed that 53% of the DS-AMKL samples had acquired cohesin mutations that were not in the self-limiting pre-leukemic transient abnormal myelopoiesis, suggesting cohesin haploinsufficiency may drive oncogenic transformation and progression to DS-AMKL.[79] In addition to the presence of trisomy 21 and GATA1s, cohesin mutations likely cooperate with chromosome 21 genes such as *RUNX1*, *ERG*, and *ETS2* to promote the development of DS-AMKL.[80, 145] Furthermore, mutations in *CTCF* occur in approximately 20% of DS-AMKL and are not mutually exclusive to cohesin mutations.[80] Cohesins and CTCF interact to regulate chromatin architecture, and thus mutations in either could have non-overlapping effects on genomic structure and induce global changes on gene expression.

Phenotypic consequences of cohesin mutations on hematopoiesis

Four recent studies sought to elucidate the phenotypic consequences of loss of cohesin and cohesin mutations on hematopoiesis in mouse and human models (Table 1.2[146]).[147, 148, 149, 150]

Viny et al. showed a dose-dependent role for Smc3 in regulating hematopoietic stem and progenitor cell (HSPC) function and chromatin structure. Biallelic loss of *Smc3* in mice led to bone marrow aplasia with premature sister chromatid separation, revealing an absolute requirement for cohesin in HSPC function; whereas, Smc3 haploinsufficiency increased self-renewal *in vitro* and *in vivo*. Furthermore, *Smc3* haploinsufficiency reduced expression of

transcription factors and lineage commitment-associated genes and cooperated with *FLT3-ITD* mutation to induce AML *in vivo*.[147]

Mullenders et al. generated a series of inducible shRNA mouse models targeting each of the four cohesin subunits. Knockdown of cohesin resulted in gain of replating capacity of mouse HSPCs and altered hematopoiesis with skewing towards myeloid differentiation. Upregulation of genes involved in myeloid differentiation and increased chromatin accessibility around those genes were also observed. In addition, aged cohesin knockdown mice developed a clinical picture closely resembling MPNs, implying that cohesin mutations can occur as an early event in leukemogenesis and facilitate the potential development of a myeloid malignancy.[148]

Complementary work in cohesin mutant human HSPCs showed that depletion of cohesin subunits increased replating capacity *in vitro* and led to myeloid-skewed differentiation, consistent with phenotypes seen in mouse models. Mazumdar et al. found that introduction of cohesin mutants into AML cell lines and primary human cord blood HSPCs resulted in a differentiation block with an increased frequency of CD34+ cord blood progenitors. Cohesin mutants augmented the serial replating capability of human HSPCs *in vitro* and elevated chromatin accessibility and predicted transcription factor binding for HSPC regulators including RUNX1, GATA2, and ERG, measured by ATAC-Seq and ChIP-Seq.[149]

Similarly, Galeev et al. identified several members of the cohesin complex SMC3, RAD21, STAG1/2 in an RNAi screen as critical modifiers of self-renewal and differentiation in human HSPCs. They showed that cohesin deficiency induced HSC-specific gene programs and the reconstitution potential of cohesin-deficient HSPCs was increased in primary and secondary transplantation studies.[150]

Figure Legends

Figure 1.1. Eight functional categories of genes that are frequently mutated in AML.[17]

Mutations in signaling genes such as *FLT3* (upper left box). Mutations in tumor-suppressor genes such as *TP53* (upper middle box). Mutations in DNA-methylation-associated genes such as *DNMT3A*, *TET2*, *IDH1*, and *IDH2* (upper right box). Mutations in myeloid transcription factors such as *RUNX1* (center left box). Mutations in cohesin-complex genes such as *STAG2* and *RAD21* (center middle box). Mutations in chromatin-modifying genes such as *ASXL1* and *PHF6* (center right box). Mutations in *NPM1* (lower left box). Mutations in spliceosome-complex genes such as *SRSF2*, *SF3B1*, *U2AF1*, and *ZRSR2* (lower right box).

Figure 1.2. The cohesin complex.[86]

Cohesin is a ring-shaped complex, composed of four core subunits SMC1A, SMC3, RAD21, and STAG1/2. SMC1A and SMC3 form intramolecular antiparallel coiled coils and fold back on themselves, creating a hinge domain at one end and an ATPase head at the other. SMC1A and SMC3 dimerize at the hinge domains and their ATPase heads are bound by RAD21. STAG1/2 interacts with the central region of RAD21. PDS5, SORORIN, and WAPL are regulatory proteins of cohesin.

Figure 1.1. Eight functional categories of genes that are frequently mutated in AML.[17]

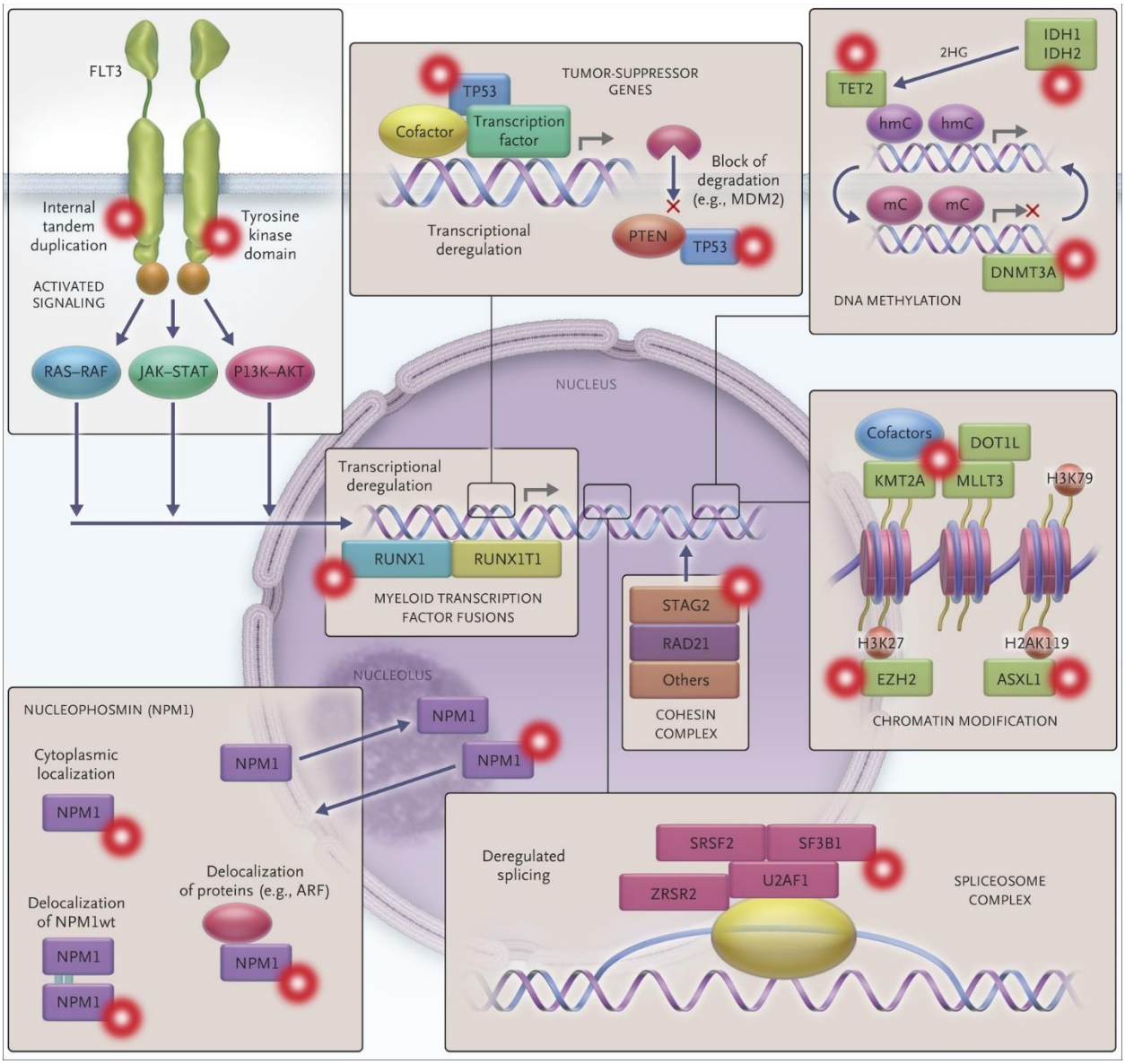


Figure 1.2. The cohesin complex.[86]

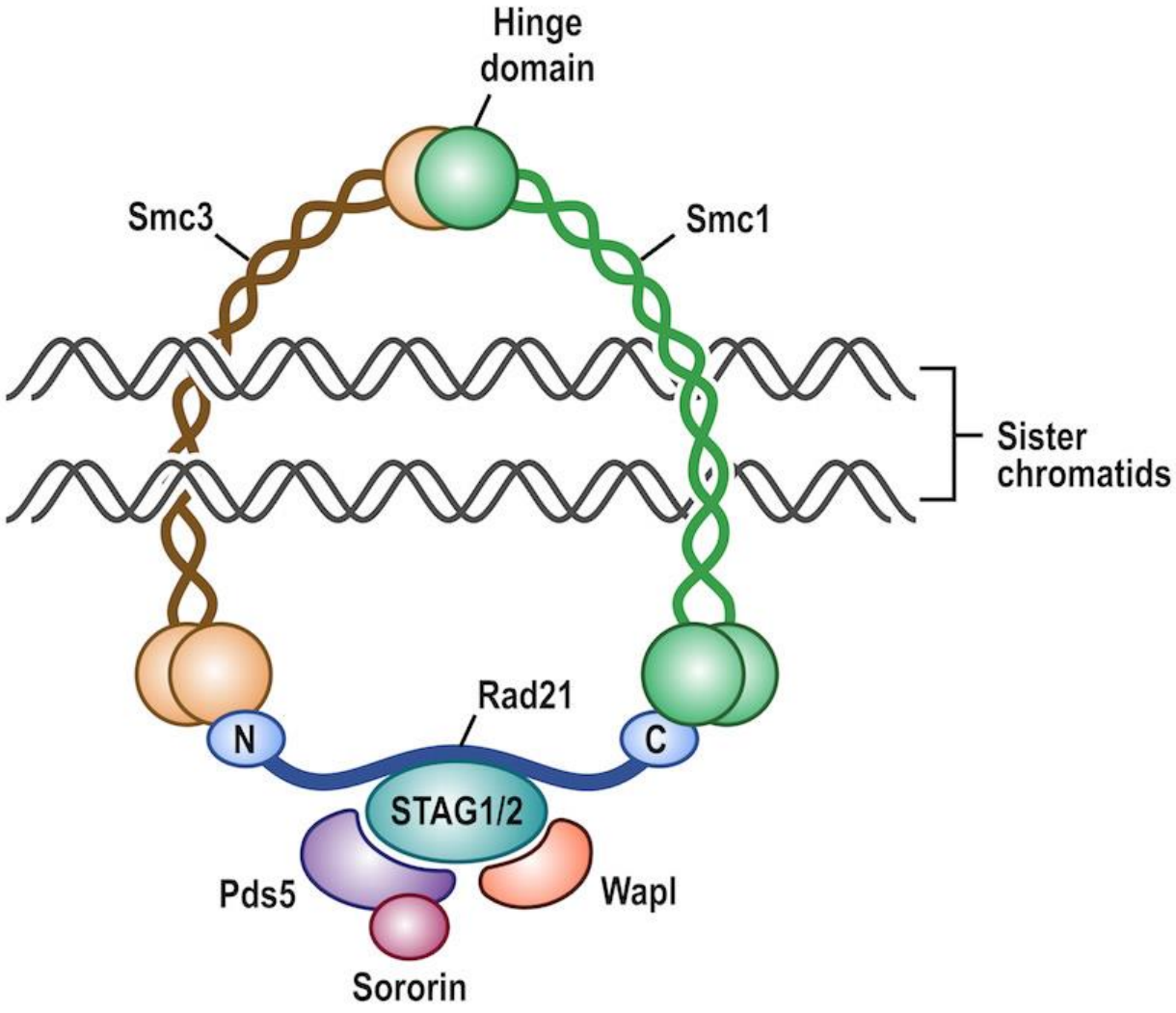


Table 1.1. Core subunits and regulatory proteins of the cohesin complex.[84]

Mammals	***D. melanogaster***	***S. cerevisiae***	***S. pombe***	**Function**
SMC1A	Smc1	Smc1	Psm1	Core subunit (mitosis)
SMC1B				Core subunit (meiosis)
SMC3	Smc3	Smc3	Psm3	Core subunit
RAD21	Rad21/Vtd	Scc1/Mcd1	Rad21	Core subunit (mitosis)
REC8	C(2)M	Rec8	Rec8	Core subunit (meiosis)
STAG1/SA1	SA	Scc3	Psc3	Core subunit (mitosis)
STAG2/SA2	SA2			Core subunit (mitosis)
STAG3/SA3	/		Rec11	Core subunit (meiosis)
NIPBL/SCC2	Nipped-B	Scc2	Mis4	Cohesin loading
MAU2/SCC4	Scc4	Scc4	Ssl3	Cohesin loading
ESCO1	Eco/Deco	Eco1/Ctf7	Eso1	Cohesion establishment
ESCO2	San			
PDS5A	Pds5	Pds5	Pds5	Cohesion maintenance
PDS5B				
WAPL	Wapl	Wpl1/Rad61	Wpl1	Cohesion maintenance
SORORIN/CDCA5	Dalmatian	/	/	Cohesion maintenance
HDAC8	/	Hos1	/	Cohesin dacetylase
Shugosin1/SGOL1	Mei-S332	Sgo1	Sgo1	Cohesin protection
Separase	Sse1	Esp1	Separase	Cohesin removal
Polo like Kinase 1 (PLK1)	Polo	Cdc5	Plk1	Cohesin removal

Table 1.2. Studies of cohesin mutations in hematopoiesis.[146]

Model system	Approach	Conclusions
Mouse model	*Smc3*biallelic and haploinsufficient conditional knockout	Increased replating, enrichment of HSPC gene signature, chromatiin accessibiligy changes, dose dependent
Mouse model	shRNA knockdown of cohesin subunits	Increased replating, enrichment of HSPC gene signature, chromatiin accessibiligy changes, MPN-like phenotype in aged mice
Human cord blood (HSPCs)	Lentiviral transduction of cohesin mutants or shRNA knockdown	Increased replating, enrichment of HSPC gene signature, chromatiin accessibiligy changes
Human cord blood (HSPCs)	RNAi screen	Increased replating, enrichment of HSPC gene signature, increased secondary transplant engraftment

References

1. Hasserjian RP. Acute myeloid leukemia: advances in diagnosis and classification. *Int J Lab Hematol*. 2013;35(3):358-366. doi:10.1111/ijlh.12081

2. Short NJ, Rytting ME, Cortes JE. Acute myeloid leukaemia. *Lancet Lond Engl*. 2018;392(10147):593-606. doi:10.1016/S0140-6736(18)31041-9

3. Yamamoto JF, Goodman MT. Patterns of leukemia incidence in the United States by subtype and demographic characteristics, 1997-2002. *Cancer Causes Control CCC*. 2008;19(4):379-390. doi:10.1007/s10552-007-9097-2

4. Acute Myeloid Leukemia - Cancer Stat Facts. SEER. https://seer.cancer.gov/statfacts/html/amyl.html. Accessed December 10, 2018.

5. Key Statistics for Acute Myeloid Leukemia (AML). https://www.cancer.org/cancer/acute-myeloid-leukemia/about/key-statistics.html. Accessed December 10, 2018.

6. Patel JP, Levine RL. How do novel molecular genetic markers influence treatment decisions in acute myeloid leukemia? *Hematol Am Soc Hematol Educ Program*. 2012;2012:28-34. doi:10.1182/asheducation-2012.1.28

7. Döhner H, Estey EH, Amadori S, et al. Diagnosis and management of acute myeloid leukemia in adults: recommendations from an international expert panel, on behalf of the European LeukemiaNet. *Blood*. 2010;115(3):453-474. doi:10.1182/blood-2009-07-235358

8. Meyers J, Yu Y, Kaye JA, Davis KL. Medicare fee-for-service enrollees with primary acute myeloid leukemia: an analysis of treatment patterns, survival, and healthcare resource utilization and costs. *Appl Health Econ Health Policy*. 2013;11(3):275-286. doi:10.1007/s40258-013-0032-2

9. Moarii M, Papaemmanuil E. Classification and risk assessment in AML: integrating cytogenetics and molecular profiling. *Hematol Am Soc Hematol Educ Program*. 2017;2017(1):37-44. doi:10.1182/asheducation-2017.1.37

10. Welch JS, Petti AA, Miller CA, et al. TP53 and Decitabine in Acute Myeloid Leukemia and Myelodysplastic Syndromes. *N Engl J Med*. 2016;375(21):2023-2036. doi:10.1056/NEJMoa1605949

11. Cancer Genome Atlas Research Network, Ley TJ, Miller C, et al. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N Engl J Med*. 2013;368(22):2059-2074. doi:10.1056/NEJMoa1301689

12. Arber DA, Orazi A, Hasserjian R, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood*. 2016;127(20):2391-2405. doi:10.1182/blood-2016-03-643544

13. Bullinger L, Döhner K, Döhner H. Genomics of Acute Myeloid Leukemia Diagnosis and Pathways. *J Clin Oncol Off J Am Soc Clin Oncol*. 2017;35(9):934-946. doi:10.1200/JCO.2016.71.2208

14. Grimwade D, Hills RK, Moorman AV, et al. Refinement of cytogenetic classification in acute myeloid leukemia: determination of prognostic significance of rare recurring chromosomal abnormalities among 5876 younger adult patients treated in the United Kingdom Medical Research Council trials. *Blood*. 2010;116(3):354-365. doi:10.1182/blood-2009-11-254441

15. Shivarov V, Bullinger L. Expression profiling of leukemia patients: key lessons and future directions. *Exp Hematol*. 2014;42(8):651-660. doi:10.1016/j.exphem.2014.04.006

16. Ley TJ, Mardis ER, Ding L, et al. DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome. *Nature*. 2008;456(7218):66-72. doi:10.1038/nature07485

17. Döhner H, Weisdorf DJ, Bloomfield CD. Acute Myeloid Leukemia. *N Engl J Med*. 2015;373(12):1136-1152. doi:10.1056/NEJMra1406184

18. Shlush LI, Zandi S, Mitchell A, et al. Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia. *Nature*. 2014;506(7488):328-333. doi:10.1038/nature13038

19. Ding L, Ley TJ, Larson DE, et al. Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature*. 2012;481(7382):506-510. doi:10.1038/nature10738

20. Busque L, Patel JP, Figueroa ME, et al. Recurrent somatic TET2 mutations in normal elderly individuals with clonal hematopoiesis. *Nat Genet*. 2012;44(11):1179-1181. doi:10.1038/ng.2413

21. Xie M, Lu C, Wang J, et al. Age-related mutations associated with clonal hematopoietic expansion and malignancies. *Nat Med*. 2014;20(12):1472-1478. doi:10.1038/nm.3733

22. Genovese G, Kähler AK, Handsaker RE, et al. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N Engl J Med*. 2014;371(26):2477-2487. doi:10.1056/NEJMoa1409405

23. Jaiswal S, Fontanillas P, Flannick J, et al. Age-related clonal hematopoiesis associated with adverse outcomes. *N Engl J Med*. 2014;371(26):2488-2498. doi:10.1056/NEJMoa1408617

24. McKerrell T, Park N, Moreno T, et al. Leukemia-associated somatic mutations drive distinct patterns of age-related clonal hemopoiesis. *Cell Rep*. 2015;10(8):1239-1245. doi:10.1016/j.celrep.2015.02.005

25. Wong TN, Miller CA, Jotte MRM, et al. Cellular stressors contribute to the expansion of hematopoietic clones of varying leukemic potential. *Nat Commun*. 2018;9(1):455. doi:10.1038/s41467-018-02858-0

26. Steensma DP, Bejar R, Jaiswal S, et al. Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes. *Blood*. 2015;126(1):9-16. doi:10.1182/blood-2015-03-631747

27. Federici L, Falini B. Nucleophosmin mutations in acute myeloid leukemia: a tale of protein unfolding and mislocalization. *Protein Sci Publ Protein Soc*. 2013;22(5):545-556. doi:10.1002/pro.2240

28. Ofran Y, Rowe JM. Genetic profiling in acute myeloid leukaemia--where are we and what is its role in patient management. *Br J Haematol*. 2013;160(3):303-320. doi:10.1111/bjh.12135

29. Schlenk RF, Döhner K, Krauter J, et al. Mutations and treatment outcome in cytogenetically normal acute myeloid leukemia. *N Engl J Med*. 2008;358(18):1909-1918. doi:10.1056/NEJMoa074306

30. Yohe S. Molecular Genetic Markers in Acute Myeloid Leukemia. *J Clin Med*. 2015;4(3):460-478. doi:10.3390/jcm4030460

31. Kayser S, Schlenk RF, Londono MC, et al. Insertion of FLT3 internal tandem duplication in the tyrosine kinase domain-1 is associated with resistance to chemotherapy and inferior outcome. *Blood*. 2009;114(12):2386-2392. doi:10.1182/blood-2009-03-209999

32. Schnittger S, Bacher U, Haferlach C, Alpermann T, Kern W, Haferlach T. Diversity of the juxtamembrane and TKD1 mutations (exons 13-15) in the FLT3 gene with regards to mutant load, sequence, length, localization, and correlation with biological data. *Genes Chromosomes Cancer*. 2012;51(10):910-924. doi:10.1002/gcc.21975

33. Breitenbuecher F, Schnittger S, Grundler R, et al. Identification of a novel type of ITD mutations located in nonjuxtamembrane domains of the FLT3 tyrosine kinase receptor. *Blood*. 2009;113(17):4074-4077. doi:10.1182/blood-2007-11-125476

34. Blau O, Berenstein R, Sindram A, Blau IW. Molecular analysis of different FLT3-ITD mutations in acute myeloid leukemia. *Leuk Lymphoma*. 2013;54(1):145-152. doi:10.3109/10428194.2012.704999

35. KIT Exon 8 Mutation in Acute Myeloid Leukemia - My Cancer Genome. https://www.mycancergenome.org/content/disease/acute-myeloid-leukemia/kit/273/. Accessed January 19, 2019.

36. Ashman LK, Griffith R. Therapeutic targeting of c-KIT in cancer. *Expert Opin Investig Drugs*. 2013;22(1):103-115. doi:10.1517/13543784.2013.740010

37. Liang D-C, Liu H-C, Yang C-P, et al. Cooperating gene mutations in childhood acute myeloid leukemia with special reference on mutations of ASXL1, TET2, IDH1, IDH2, and DNMT3A. *Blood*. 2013;121(15):2988-2995. doi:10.1182/blood-2012-06-436782

38. Patel JP, Gönen M, Figueroa ME, et al. Prognostic relevance of integrated genetic profiling in acute myeloid leukemia. *N Engl J Med*. 2012;366(12):1079-1089. doi:10.1056/NEJMoa1112304

39. Papaemmanuil E, Gerstung M, Bullinger L, et al. Genomic Classification and Prognosis in Acute Myeloid Leukemia. *N Engl J Med*. 2016;374(23):2209-2221. doi:10.1056/NEJMoa1516192

40. Gaidzik VI, Teleanu V, Papaemmanuil E, et al. RUNX1 mutations in acute myeloid leukemia are associated with distinct clinico-pathologic and genetic features. *Leukemia*. 2016;30(11):2282. doi:10.1038/leu.2016.207

41. Preudhomme C, Renneville A, Bourdon V, et al. High frequency of RUNX1 biallelic alteration in acute myeloid leukemia secondary to familial platelet disorder. *Blood*. 2009;113(22):5583-5587. doi:10.1182/blood-2008-07-168260

42. Martelli MP, Sportoletti P, Tiacci E, Martelli MF, Falini B. Mutational landscape of AML with normal cytogenetics: biological and clinical implications. *Blood Rev*. 2013;27(1):13-22. doi:10.1016/j.blre.2012.11.001

43. Pabst T, Mueller BU, Zhang P, et al. Dominant-negative mutations of CEBPA, encoding CCAAT/enhancer binding protein-alpha (C/EBPalpha), in acute myeloid leukemia. *Nat Genet*. 2001;27(3):263-270. doi:10.1038/85820

44. Nerlov C. C/EBPalpha mutations in acute myeloid leukaemias. *Nat Rev Cancer*. 2004;4(5):394-400. doi:10.1038/nrc1363

45. Mueller BU, Pabst T. C/EBPalpha and the pathophysiology of acute myeloid leukemia. *Curr Opin Hematol*. 2006;13(1):7-14.

46. Wouters BJ, Löwenberg B, Erpelinck-Verschueren CAJ, van Putten WLJ, Valk PJM, Delwel R. Double CEBPA mutations, but not single CEBPA mutations, define a subgroup of acute myeloid leukemia with a distinctive gene expression profile that is uniquely associated with a favorable outcome. *Blood*. 2009;113(13):3088-3091. doi:10.1182/blood-2008-09-179895

47. Kihara R, Nagata Y, Kiyoi H, et al. Comprehensive analysis of genetic alterations and their prognostic impacts in adult acute myeloid leukemia patients. *Leukemia*. 2014;28(8):1586-1595. doi:10.1038/leu.2014.55

48. Yang L, Han Y, Suarez Saiz F, Saurez Saiz F, Minden MD. A tumor suppressor and oncogene: the WT1 story. *Leukemia*. 2007;21(5):868-876. doi:10.1038/sj.leu.2404624

49. Welch JS, Ley TJ, Link DC, et al. The origin and evolution of mutations in acute myeloid leukemia. *Cell*. 2012;150(2):264-278. doi:10.1016/j.cell.2012.06.023

50. Ibrahem L, Mahfouz R, Elhelw L, Abdsalam EM, Soliman R. Prognostic significance of DNMT3A mutations in patients with acute myeloid leukemia. *Blood Cells Mol Dis*. 2015;54(1):84-89. doi:10.1016/j.bcmd.2014.07.015

51. Russler-Germain DA, Spencer DH, Young MA, et al. The R882H DNMT3A mutation associated with AML dominantly inhibits wild-type DNMT3A by blocking its ability to form active tetramers. *Cancer Cell*. 2014;25(4):442-454. doi:10.1016/j.ccr.2014.02.010

52. Yamashita Y, Yuan J, Suetake I, et al. Array-based genomic resequencing of human leukemia. *Oncogene*. 2010;29(25):3723-3731. doi:10.1038/onc.2010.117

53. Holz-Schietinger C, Matje DM, Reich NO. Mutations in DNA methyltransferase (DNMT3A) observed in acute myeloid leukemia patients disrupt processive methylation. *J Biol Chem*. 2012;287(37):30941-30951. doi:10.1074/jbc.M112.366625

54. Aslanyan MG, Kroeze LI, Langemeijer SMC, et al. Clinical and biological impact of TET2 mutations and expression in younger adult AML patients treated within the EORTC/GIMEMA AML-12 clinical trial. *Ann Hematol*. 2014;93(8):1401-1412. doi:10.1007/s00277-014-2055-7

55. Gaidzik VI, Paschka P, Späth D, et al. TET2 mutations in acute myeloid leukemia (AML): results from a comprehensive genetic and clinical analysis of the AML study group. *J Clin Oncol Off J Am Soc Clin Oncol*. 2012;30(12):1350-1357. doi:10.1200/JCO.2011.39.2886

56. Tian X, Xu Y, Yin J, et al. TET2 gene mutation is unfavorable prognostic factor in cytogenetically normal acute myeloid leukemia patients with NPM1+ and FLT3-ITD - mutations. *Int J Hematol*. 2014;100(1):96-104. doi:10.1007/s12185-014-1595-x

57. Lu C, Ward PS, Kapoor GS, et al. IDH mutation impairs histone demethylation and results in a block to cell differentiation. *Nature*. 2012;483(7390):474-478. doi:10.1038/nature10860

58. Ward PS, Cross JR, Lu C, et al. Identification of additional IDH mutations associated with oncometabolite R(-)-2-hydroxyglutarate production. *Oncogene*. 2012;31(19):2491-2498. doi:10.1038/onc.2011.416

59. Paschka P, Schlenk RF, Gaidzik VI, et al. 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. *J Clin Oncol Off J Am Soc Clin Oncol*. 2010;28(22):3636-3643. doi:10.1200/JCO.2010.28.3762

60. Abbas S, Lugthart S, Kavelaars FG, et al. Acquired mutations in the genes encoding IDH1 and IDH2 both are recurrent aberrations in acute myeloid leukemia: prevalence and prognostic value. *Blood*. 2010;116(12):2122-2126. doi:10.1182/blood-2009-11-250878

61. Ho PA, Kutny MA, Alonzo TA, et al. Leukemic mutations in the methylation-associated genes DNMT3A and IDH2 are rare events in pediatric AML: a report from the Children's Oncology Group. *Pediatr Blood Cancer*. 2011;57(2):204-209. doi:10.1002/pbc.23179

62. Lindsley RC, Mar BG, Mazzola E, et al. Acute myeloid leukemia ontogeny is defined by distinct somatic mutations. *Blood*. 2015;125(9):1367-1376. doi:10.1182/blood-2014-11-610543

63. El-Sharkawi D, Ali A, Evans CM, et al. ASXL1 mutations are infrequent in young patients with primary acute myeloid leukemia and their detection has a limited role in therapeutic risk stratification. *Leuk Lymphoma*. 2014;55(6):1326-1331. doi:10.3109/10428194.2013.833332

64. Todd MAM, Ivanochko D, Picketts DJ. PHF6 Degrees of Separation: The Multifaceted Roles of a Chromatin Adaptor Protein. *Genes*. 2015;6(2):325-352. doi:10.3390/genes6020325

65. Lower KM, Turner G, Kerr BA, et al. Mutations in PHF6 are associated with Börjeson-Forssman-Lehmann syndrome. *Nat Genet*. 2002;32(4):661-665. doi:10.1038/ng1040

66. Van Vlierberghe P, Patel J, Abdel-Wahab O, et al. PHF6 mutations in adult acute myeloid leukemia. *Leukemia*. 2011;25(1):130-134. doi:10.1038/leu.2010.247

67. Yoo NJ, Kim YR, Lee SH. Somatic mutation of PHF6 gene in T-cell acute lymphoblatic leukemia, acute myelogenous leukemia and hepatocellular carcinoma. *Acta Oncol Stockh Swed*. 2012;51(1):107-111. doi:10.3109/0284186X.2011.592148

68. Shukla GC, Singh J. Mutations of RNA splicing factors in hematological malignancies. *Cancer Lett*. 2017;409:1-8. doi:10.1016/j.canlet.2017.08.042

69. Papaemmanuil E, Gerstung M, Malcovati L, et al. Clinical and biological implications of driver mutations in myelodysplastic syndromes. *Blood*. 2013;122(22):3616-3627; quiz 3699. doi:10.1182/blood-2013-08-518886

70. Makishima H, Visconte V, Sakaguchi H, et al. Mutations in the spliceosome machinery, a novel and ubiquitous pathway in leukemogenesis. *Blood*. 2012;119(14):3203-3210. doi:10.1182/blood-2011-12-399774

71. Kim E, Ilagan JO, Liang Y, et al. SRSF2 Mutations Contribute to Myelodysplasia by Mutant-Specific Effects on Exon Recognition. *Cancer Cell*. 2015;27(5):617-630. doi:10.1016/j.ccell.2015.04.006

72. Yoshida K, Sanada M, Shiraishi Y, et al. Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature*. 2011;478(7367):64-69. doi:10.1038/nature10496

73. Przychodzen B, Jerez A, Guinta K, et al. Patterns of missplicing due to somatic U2AF1 mutations in myeloid neoplasms. *Blood*. 2013;122(6):999-1006. doi:10.1182/blood-2013-01-480970

74. Shirai CL, Ley JN, White BS, et al. Mutant U2AF1 Expression Alters Hematopoiesis and Pre-mRNA Splicing In Vivo. *Cancer Cell*. 2015;27(5):631-643. doi:10.1016/j.ccell.2015.04.008

75. Madan V, Kanojia D, Li J, et al. Aberrant splicing of U12-type introns is the hallmark of ZRSR2 mutant myelodysplastic syndrome. *Nat Commun*. 2015;6:6042. doi:10.1038/ncomms7042

76. Thol F, Bollin R, Gehlhaar M, et al. Mutations in the cohesin complex in acute myeloid leukemia: clinical and prognostic implications. *Blood*. 2014;123(6):914-920. doi:10.1182/blood-2013-07-518746

77. Kon A, Shih L-Y, Minamino M, et al. Recurrent mutations in multiple components of the cohesin complex in myeloid neoplasms. *Nat Genet*. 2013;45(10):1232-1237. doi:10.1038/ng.2731

78. Thota S, Viny AD, Makishima H, et al. Genetic alterations of the cohesin complex genes in myeloid malignancies. *Blood*. 2014;124(11):1790-1798. doi:10.1182/blood-2014-04-567057

79. Yoshida K, Toki T, Okuno Y, et al. The landscape of somatic mutations in Down syndrome-related myeloid disorders. *Nat Genet*. 2013;45(11):1293-1299. doi:10.1038/ng.2759

80. Fisher JB, McNulty M, Burke MJ, Crispino JD, Rao S. Cohesin Mutations in Myeloid Malignancies. *Trends Cancer*. 2017;3(4):282-293. doi:10.1016/j.trecan.2017.02.006

81. Conese M, Liso A. Cohesin complex is a major player on the stage of leukemogenesis. *Stem Cell Investig*. 2016;3:18. doi:10.21037/sci.2016.05.04

82. Jaiswal S, Natarajan P, Silver AJ, et al. Clonal Hematopoiesis and Risk of Atherosclerotic Cardiovascular Disease. *N Engl J Med*. 2017;377(2):111-121. doi:10.1056/NEJMoa1701719

83. Faure AJ, Schmidt D, Watt S, et al. Cohesin regulates tissue-specific expression by stabilizing highly occupied cis-regulatory modules. *Genome Res*. 2012;22(11):2163-2175. doi:10.1101/gr.136507.111

84. Bhardwaj S, Gullerova M. Cohesin: It's Functions and Relevance to Cancer. *J Genet Syndr Gene Ther*. 2013;04(05). doi:10.4172/2157-7412.1000148

85. Nasmyth K, Haering CH. Cohesin: its roles and mechanisms. *Annu Rev Genet*. 2009;43:525-558. doi:10.1146/annurev-genet-102108-134233

86. Research | Solomon Lab. https://solomonlab.ucsf.edu/research. Accessed January 11, 2019.

87. Peters J-M. The many functions of cohesin--different rings to rule them all? *EMBO J*. 2012;31(9):2061-2063. doi:10.1038/emboj.2012.90

88. Nasmyth K. How might cohesin hold sister chromatids together? *Philos Trans R Soc Lond B Biol Sci*. 2005;360(1455):483-496. doi:10.1098/rstb.2004.1604

89. Zhang N, Kuznetsov SG, Sharan SK, Li K, Rao PH, Pati D. A handcuff model for the cohesin complex. *J Cell Biol*. 2008;183(6):1019-1031. doi:10.1083/jcb.200801157

90. Remeseiro S, Cuadrado A, Gómez-López G, Pisano DG, Losada A. A unique role of cohesin-SA1 in gene regulation and development. *EMBO J*. 2012;31(9):2090-2102. doi:10.1038/emboj.2012.60

91. Ciosk R, Shirayama M, Shevchenko A, et al. Cohesin's binding to chromosomes depends on a separate complex consisting of Scc2 and Scc4 proteins. *Mol Cell*. 2000;5(2):243-254.

92. Rollins RA, Korom M, Aulner N, Martens A, Dorsett D. Drosophila nipped-B protein supports sister chromatid cohesion and opposes the stromalin/Scc3 cohesion factor to facilitate long-range activation of the cut gene. *Mol Cell Biol*. 2004;24(8):3100-3111.

93. Bernard P, Drogat J, Maure J-F, et al. A screen for cohesion mutants uncovers Ssl3, the fission yeast counterpart of the cohesin loading factor Scc4. *Curr Biol CB*. 2006;16(9):875-881. doi:10.1016/j.cub.2006.03.037

94. Watrin E, Schleiffer A, Tanaka K, Eisenhaber F, Nasmyth K, Peters J-M. Human Scc4 is required for cohesin binding to chromatin, sister-chromatid cohesion, and mitotic progression. *Curr Biol CB*. 2006;16(9):863-874. doi:10.1016/j.cub.2006.03.049

95. Seitan VC, Banks P, Laval S, et al. Metazoan Scc4 homologs link sister chromatid cohesion to cell and axon migration guidance. *PLoS Biol*. 2006;4(8):e242. doi:10.1371/journal.pbio.0040242

96. Losada A, Yokochi T, Hirano T. Functional contribution of Pds5 to cohesin-mediated cohesion in human cells and Xenopus egg extracts. *J Cell Sci*. 2005;118(Pt 10):2133-2141. doi:10.1242/jcs.02355

97. Zhang J, Shi X, Li Y, et al. Acetylation of Smc3 by Eco1 is required for S phase sister chromatid cohesion in both human and yeast. *Mol Cell*. 2008;31(1):143-151. doi:10.1016/j.molcel.2008.06.006

98. Rowland BD, Roig MB, Nishino T, et al. Building sister chromatid cohesion: smc3 acetylation counteracts an antiestablishment activity. *Mol Cell*. 2009;33(6):763-774. doi:10.1016/j.molcel.2009.02.028

99. Nishiyama T, Ladurner R, Schmitz J, et al. Sororin mediates sister chromatid cohesion by antagonizing Wapl. *Cell*. 2010;143(5):737-749. doi:10.1016/j.cell.2010.10.031

100. Sumara I, Vorlaufer E, Gieffers C, Peters BH, Peters JM. Characterization of vertebrate cohesin complexes and their regulation in prophase. *J Cell Biol*. 2000;151(4):749-762.

101. Kitajima TS, Sakuno T, Ishiguro K, et al. Shugoshin collaborates with protein phosphatase 2A to protect cohesin. *Nature*. 2006;441(7089):46-52. doi:10.1038/nature04663

102. Hauf S, Waizenegger IC, Peters JM. Cohesin cleavage by separase required for anaphase and cytokinesis in human cells. *Science*. 2001;293(5533):1320-1323. doi:10.1126/science.1061376

103. Borges V, Lehane C, Lopez-Serra L, et al. Hos1 deacetylates Smc3 to close the cohesin acetylation cycle. *Mol Cell*. 2010;39(5):677-688. doi:10.1016/j.molcel.2010.08.009

104. Deardorff MA, Bando M, Nakato R, et al. HDAC8 mutations in Cornelia de Lange syndrome affect the cohesin acetylation cycle. *Nature*. 2012;489(7415):313-317. doi:10.1038/nature11316

105. Watanabe Y, Nurse P. Cohesin Rec8 is required for reductional chromosome segregation at meiosis. *Nature*. 1999;400(6743):461-464. doi:10.1038/22774

106. Nasmyth K. Disseminating the genome: joining, resolving, and separating sister chromatids during mitosis and meiosis. *Annu Rev Genet*. 2001;35:673-745. doi:10.1146/annurev.genet.35.102401.091334

107. Panigrahi AK, Pati D. Road to the crossroads of life and death: linking sister chromatid cohesion and separation to aneuploidy, apoptosis and cancer. *Crit Rev Oncol Hematol*. 2009;72(3):181-193. doi:10.1016/j.critrevonc.2008.12.002

108. Sjögren C, Nasmyth K. Sister chromatid cohesion is required for postreplicative double-strand break repair in Saccharomyces cerevisiae. *Curr Biol CB*. 2001;11(12):991-995.

109. Schär P, Fäsi M, Jessberger R. SMC1 coordinates DNA double-strand break repair pathways. *Nucleic Acids Res*. 2004;32(13):3921-3929. doi:10.1093/nar/gkh716

110. Atienza JM, Roth RB, Rosette C, et al. Suppression of RAD21 gene expression decreases cell growth and enhances cytotoxicity of etoposide and bleomycin in human breast cancer cells. *Mol Cancer Ther*. 2005;4(3):361-368. doi:10.1158/1535-7163.MCT-04-0241

111. Birkenbihl RP, Subramani S. Cloning and characterization of rad21 an essential gene of Schizosaccharomyces pombe involved in DNA double-strand-break repair. *Nucleic Acids Res*. 1992;20(24):6605-6611.

112. Kim B-J, Li Y, Zhang J, et al. Genome-wide reinforcement of cohesin binding at pre-existing cohesin sites in response to ionizing radiation in human cells. *J Biol Chem*. 2010;285(30):22784-22792. doi:10.1074/jbc.M110.134577

113. Schmitz J, Watrin E, Lénárt P, Mechtler K, Peters J-M. Sororin is required for stable binding of cohesin to chromatin and for sister chromatid cohesion in interphase. *Curr Biol CB*. 2007;17(7):630-636. doi:10.1016/j.cub.2007.02.029

114. Ström L, Karlsson C, Lindroos HB, et al. Postreplicative formation of cohesion is required for repair and induced by a single DNA break. *Science*. 2007;317(5835):242-245. doi:10.1126/science.1140649

115. Unal E, Heidinger-Pauli JM, Kim W, et al. A molecular determinant for the establishment of sister chromatid cohesion. *Science*. 2008;321(5888):566-569. doi:10.1126/science.1157880

116. Watrin E, Peters J-M. The cohesin complex is required for the DNA damage-induced G2/M checkpoint in mammalian cells. *EMBO J*. 2009;28(17):2625-2635. doi:10.1038/emboj.2009.202

117. Lightfoot J, Testori S, Barroso C, Martinez-Perez E. Loading of meiotic cohesin by SCC-2 is required for early processing of DSBs and for the DNA damage checkpoint. *Curr Biol CB*. 2011;21(17):1421-1430. doi:10.1016/j.cub.2011.07.007

118. Rollins RA, Morcillo P, Dorsett D. Nipped-B, a Drosophila homologue of chromosomal adherins, participates in activation by remote enhancers in the cut and Ultrabithorax genes. *Genetics*. 1999;152(2):577-593.

119. Dorsett D, Eissenberg JC, Misulovin Z, Martens A, Redding B, McKim K. Effects of sister chromatid cohesion proteins on cut gene expression during wing development in Drosophila. *Dev Camb Engl*. 2005;132(21):4743-4753. doi:10.1242/dev.02064

120. Rhodes JM, Bentley FK, Print CG, et al. Positive regulation of c-Myc by cohesin is direct, and evolutionarily conserved. *Dev Biol*. 2010;344(2):637-649. doi:10.1016/j.ydbio.2010.05.493

121. Merkenschlager M, Odom DT. CTCF and cohesin: linking gene regulatory elements with their targets. *Cell*. 2013;152(6):1285-1297. doi:10.1016/j.cell.2013.02.029

122. Seitan VC, Hao B, Tachibana-Konwalski K, et al. A role for cohesin in T-cell-receptor rearrangement and thymocyte differentiation. *Nature*. 2011;476(7361):467-471. doi:10.1038/nature10312

123. Seitan VC, Merkenschlager M. Cohesin and chromatin organisation. *Curr Opin Genet Dev*. 2012;22(2):93-100. doi:10.1016/j.gde.2011.11.003

124. Schmidt D, Schwalie PC, Ross-Innes CS, et al. A CTCF-independent role for cohesin in tissue-specific transcription. *Genome Res*. 2010;20(5):578-588. doi:10.1101/gr.100479.109

125. Fullwood MJ, Liu MH, Pan YF, et al. An oestrogen-receptor-alpha-bound human chromatin interactome. *Nature*. 2009;462(7269):58-64. doi:10.1038/nature08497

126. Kagey MH, Newman JJ, Bilodeau S, et al. Mediator and cohesin connect gene expression and chromatin architecture. *Nature*. 2010;467(7314):430-435. doi:10.1038/nature09380

127. Zhang H, Jiao W, Sun L, et al. Intrachromosomal looping is required for activation of endogenous pluripotency genes during reprogramming. *Cell Stem Cell*. 2013;13(1):30-35. doi:10.1016/j.stem.2013.05.012

128. Walter MJ, Payton JE, Ries RE, et al. Acquired copy number alterations in adult acute myeloid leukemia genomes. *Proc Natl Acad Sci U S A*. 2009;106(31):12950-12955. doi:10.1073/pnas.0903091106

129. Barber TD, McManus K, Yuen KWY, et al. Chromatid cohesion defects may underlie chromosome instability in human colorectal cancers. *Proc Natl Acad Sci U S A*. 2008;105(9):3443-3448. doi:10.1073/pnas.0712384105

130. Gorringe KL, Ramakrishna M, Williams LH, et al. Are there any more ovarian tumor suppressor genes? A new perspective using ultra high-resolution copy number and loss of heterozygosity analysis. *Genes Chromosomes Cancer*. 2009;48(10):931-942. doi:10.1002/gcc.20694

131. Solomon DA, Kim T, Diaz-Martinez LA, et al. Mutational inactivation of STAG2 causes aneuploidy in human cancer. *Science*. 2011;333(6045):1039-1043. doi:10.1126/science.1203619

132. Rocquain J, Gelsi-Boyer V, Adélaïde J, et al. Alteration of cohesin genes in myeloid diseases. *Am J Hematol*. 2010;85(9):717-719. doi:10.1002/ajh.21798

133. Kim J-S, He X, Orr B, et al. Intact Cohesion, Anaphase, and Chromosome Segregation in Human Cells Harboring Tumor-Derived Mutations in STAG2. *PLoS Genet*. 2016;12(2):e1005865. doi:10.1371/journal.pgen.1005865

134. Hill VK, Kim J-S, Waldman T. Cohesin mutations in human cancer. *Biochim Biophys Acta*. 2016;1866(1):1-11. doi:10.1016/j.bbcan.2016.05.002

135. Solomon DA, Kim J-S, Bondaruk J, et al. Frequent truncating mutations of STAG2 in bladder cancer. *Nat Genet*. 2013;45(12):1428-1430. doi:10.1038/ng.2800

136. Balbás-Martínez C, Sagrera A, Carrillo-de-Santa-Pau E, et al. Recurrent inactivation of STAG2 in bladder cancer is not associated with aneuploidy. *Nat Genet*. 2013;45(12):1464-1469. doi:10.1038/ng.2799

137. Guo G, Sun X, Chen C, et al. Whole-genome and whole-exome sequencing of bladder cancer identifies frequent alterations in genes involved in sister chromatid cohesion and segregation. *Nat Genet*. 2013;45(12):1459-1463. doi:10.1038/ng.2798

138. Ma Z, Lin M, Li K, et al. Knocking down SMC1A inhibits growth and leads to G2/M arrest in human glioma cells. *Int J Clin Exp Pathol*. 2013;6(5):862-869.

139. Ryu B, Kim DS, Deluca AM, Alani RM. Comprehensive expression profiling of tumor cell lines identifies molecular signatures of melanoma progression. *PloS One*. 2007;2(7):e594. doi:10.1371/journal.pone.0000594

140. Oikawa K, Ohbayashi T, Kiyono T, et al. Expression of a novel human gene, human wings apart-like (hWAPL), is associated with cervical carcinogenesis and tumor progression. *Cancer Res*. 2004;64(10):3545-3549. doi:10.1158/0008-5472.CAN-03-3822

141. Meyer R, Fofanov V, Panigrahi A, Merchant F, Zhang N, Pati D. Overexpression and mislocalization of the chromosomal segregation protein separase in multiple human cancers. *Clin Cancer Res Off J Am Assoc Cancer Res*. 2009;15(8):2703-2710. doi:10.1158/1078-0432.CCR-08-2454

142. Leiserson MDM, Vandin F, Wu H-T, et al. Pan-cancer network analysis identifies combinations of rare somatic mutations across pathways and protein complexes. *Nat Genet*. 2015;47(2):106-114. doi:10.1038/ng.3168

143. Walter MJ, Shen D, Ding L, et al. Clonal architecture of secondary acute myeloid leukemia. *N Engl J Med*. 2012;366(12):1090-1098. doi:10.1056/NEJMoa1106968

144. Walter MJ, Shen D, Shao J, et al. Clonal diversity of recurrently mutated genes in myelodysplastic syndromes. *Leukemia*. 2013;27(6):1275-1282. doi:10.1038/leu.2013.58

145. Malinge S, Bliss-Moreau M, Kirsammer G, et al. Increased dosage of the chromosome 21 ortholog Dyrk1a promotes megakaryoblastic leukemia in a murine model of Down syndrome. *J Clin Invest*. 2012;122(3):948-962. doi:10.1172/JCI60455

146. Mazumdar C, Majeti R. The role of mutations in the cohesin complex in acute myeloid leukemia. *Int J Hematol*. 2017;105(1):31-36. doi:10.1007/s12185-016-2119-7

147. Viny AD, Ott CJ, Spitzer B, et al. Dose-dependent role of the cohesin complex in normal and malignant hematopoiesis. *J Exp Med*. 2015;212(11):1819-1832. doi:10.1084/jem.20151317

148. Mullenders J, Aranda-Orgilles B, Lhoumaud P, et al. Cohesin loss alters adult hematopoietic stem cell homeostasis, leading to myeloproliferative neoplasms. *J Exp Med*. 2015;212(11):1833-1850. doi:10.1084/jem.20151323

149. Mazumdar C, Shen Y, Xavy S, et al. Leukemia-Associated Cohesin Mutants Dominantly Enforce Stem Cell Programs and Impair Human Hematopoietic Progenitor Differentiation. *Cell Stem Cell*. 2015;17(6):675-688. doi:10.1016/j.stem.2015.09.017

150. Galeev R, Baudet A, Kumar P, et al. Genome-wide RNAi Screen Identifies Cohesin Genes as Modifiers of Renewal and Differentiation in Human HSCs. *Cell Rep*. 2016;14(12):2988-3000. doi:10.1016/j.celrep.2016.02.082

Chapter 2:
Smc3 is required for mouse embryonic and adult hematopoiesis

Abstract

SMC3 encodes a subunit of the cohesin complex that has canonical roles in regulating sister chromatids segregation during mitosis and meiosis. Recurrent heterozygous mutations in *SMC3* have been reported in acute myeloid leukemia (AML) and other myeloid malignancies. In this study, we investigated whether the missense mutations in *SMC3* might have dominant-negative effects or phenocopy loss-of-function effects by comparing the consequences of *Smc3* deficient and haploinsufficient mouse models. We found that homozygous deletion of *Smc3* during embryogenesis or in adult mice led to hematopoietic failure, suggesting that *SMC3* missense mutations are unlikely to be associated with simple dominant negative phenotypes. In contrast, *Smc3* haploinsufficiency was tolerated during embryonic and adult hematopoiesis. Under steady-state conditions, *Smc3* haploinsufficiency did not alter colony forming in methylcellulose, only modestly decreased mature myeloid cell populations, and led to limited expression changes and chromatin alteration in Lin-cKit+ bone marrow cells. However, following transplantation, engraftment, and subsequent deletion, we observed a hematopoietic competitive disadvantage across myeloid and lymphoid lineages and within the stem/progenitor compartments. This disadvantage was not affected by hematopoietic stresses but was partially abrogated by concurrent *Dnmt3a* haploinsufficiency, suggesting that antecedent mutations may be required to optimize the leukemogenic potential of *Smc3* mutations.

Introduction

Acute myeloid leukemia (AML) is an aggressive hematopoietic malignancy, characterized by the accumulation of myeloblasts in the blood or bone marrow (BM) with maturation arrest and retained self-renewal.[1] Tremendous progress has been made in identifying recurrent gene

mutations in AML, yet we are still in the early stages of understanding the mechanisms through which these genetic alterations contribute to the onset of the disease.[2]

Recurring mutations in the cohesin complex occur in four core components, *SMC3, SMC1A*, *RAD21*, and *STAG2*, and have been identified in AML and other myeloid malignancies.[3, 4, 5] Over 50% of patients with Down syndrome-associated acute megakaryocytic leukemia (DS-AMKL) have cohesin mutations, specifically in *STAG2*.[6] Somatic cohesin mutations have also been observed in a variety of solid cancers, including colorectal carcinoma, ovarian carcinoma, glioblastoma, bladder carcinoma, and Ewing's sarcoma.[7, 8, 9, 10, 11, 12] Additionally, germline mutations of the cohesin complex are causally related to developmental disorders, particularly cohesinopathies such as Cornelia de Lange syndrome (CdLS).[13, 14]

SMC3 and *RAD21* mutations are nearly universally heterozygous, whereas mutations in *SMC1A* and *STAG2* may be hemizygous because they are X-linked. Cohesin mutations also tend to be mutually exclusive, implying that alteration in one component may be sufficient to disrupt the entire complex or alternatively, they may not be tolerated by a cell when co-occurring.[15, 16] Cohesin mutations are often observed as early subclonal events in AML, conceivably facilitating disease initiation, although they are not observed in cases of clonal hematopoiesis of indeterminate potential (CHIP), suggesting they are unlikely to be the initiating event.[15, 17, 18, 19]

The majority of *SMC3* mutations are missense mutations; only one-third of *SMC3* mutations are nonsense or splice-site variants. The missense mutations are scattered across all domains, although a few recurrently mutated nucleotides have been observed (R381Q, R661P). This pattern suggests that many of these mutations may result in simple loss-of-function consequences, although novel dominant negative activities cannot be dismissed within the hot-spot variants. Intriguingly, *DNMT3A* mutations, one of the most commonly mutated genes in

AML, frequently coincided with *SMC3* mutations, suggesting there may be leukemogenic interactions between these mutations.[5, 15, 16, 20]

In yeast and cell line-based studies, cohesin has been shown to play essential roles in sister chromatid segregation during cell cycle, DNA damage repair, transcriptional regulation via chromatin looping, and maintenance of chromatin architecture.[21, 22, 23, 24] Notably, AML patients who harbor cohesin mutations typically have normal karyotype, indicating that hematopoietic cohesin mutations do not lead directly to chromosomal instability.[25, 16]

To define the hematopoietic consequences of *SMC3* mutations and to determine whether these could reflect dominant negative or loss of function phenotypes, we characterized the *in vivo* effects of *Smc3* deficiency and *Smc3* haploinsufficiency on murine hematopoiesis using conditionally deleted strategies. In contrast to our expectations that these leukemia-associated mutations would lead to expansions of hematopoietic stem cell populations or augmented self-renewal, we observed a competitive disadvantage in *Smc3* deficient and haploinsufficient BM cells *in vivo* without an associated increase in maturation-arrested stem cells.

Methods

Animal Studies

Smc3^{trap} mice were obtained from the European Conditional Mouse Mutagenesis Program (EUCOMM) (Smc3<tm1a(EUCOMM)Wtsi>, MGI:4434007). To generate *Smc3*^{fl} mice, the gene-trap was removed by crossing *Smc3*^{trap} mice with Flp deleter mice (B6.129S4-Gt(ROSA)26Sortm2(FLP*)Sor/J), and subsequently outbreeding the Flp allele with C57BL/6J intercrosses. We generated *Smc3* conditional deficient mice by breeding the *Smc3*$^{fl/fl}$ mice with *Vav1-Cre* (B6.Cg-Commd10Tg(Vav1-icre)A2Kio/J), *ERT2-Cre* (B6.Cg-Tg(cre/Esr1)5Amc/J),

and *CMV-Cre* (B6.C-Tg(CMV-cre)1Cgn/J), obtained from the Jackson Laboratory. We characterized *Smc3* conditional deficient mice at 6-8 weeks old and both genders were used. Whenever possible, littermate controls were used for all experiments. CBCs were measured using Hemavet 950 (Drew Scientific Group).

All mice were on the C57BL/6 background and were cared for in the Experimental Animal Center of Washington University School of Medicine. The Washington University Animal Studies Committee approved all animal experiments.

Intracellular* Smc3 *staining

Intracellular Smc3 was detected with the Pharmingen™ Transcription Factor Buffer Set (562574 BD Biosciences) according to the manufacturer's instructions. BM cells were isolated from femurs and tibias and lysed with ACK lysis buffer (150mM NH_4Cl, 10mM $KHCO_3$, 0.1mM Na_2EDTA [Na_2-ethylenediaminetetraacetic acid], PH7.2-7.4). Cells were stained with cell-surface markers to identify cell type by flow cytometry and then fixed for 40 minutes at 4°C. Cells were washed with perm wash buffer and incubated with primary antibody against Smc3 (1:100 dilution, ab9263, Abcam) for 30 minutes at 4°C. Cells were washed in perm wash buffer and incubated in secondary antibody (1:500 dilution, chicken anti-rabbit Alexa Fluor 647, Molecular Probes) for 30 minutes at 4°C. Cells were rinsed in perm wash buffer and analyzed by flow cytometry. The mean fluorescence intensity was calculated for the AF647 signal.

Flow cytometry

After lysis of red blood cells by ACK lysis buffer, peripheral blood, BM, spleen cells, or thymocytes were treated with anti-mouse CD16/32 (eBioscience; clone 93) and stained with the indicated combinations of the following antibodies (all antibodies are from eBioscience unless

noted otherwise): CD34 FITC (clone RAM34), CD11b PE (clone M1/70), c-Kit PECy7 (clone 2B8) or BV421 (BioLegend, clone 2B8), Sca1 PE-Dazzle™ 594 (BioLegend, clone D7) or APC (clone D7), Gr-1 FITC, PECy7, APC (clone RB6-8C5), or BV421 (BioLegend, clone RB6-8C5), B220 PE, PECy7, APC (clone RA3-6B2), or APC-Cy7 (BioLegend, clone RA3-6B2), CD3 PECy7 (clone 145-2C11), CD71 PE(clone R17217), Ter-119 PECy7 or APC (clone TER-119), CD16/32 BV510 (clone 93), CD150 PE (BioLegend 115903, clone TC15-12F12.2), CD48 APC-Cy7 (BioLegend, clone HM48-1), Ly5.1 APC (clone A20) or AF700 (BioLegend, clone A20), Ly 5.2 PE or e450 (clone 104). The following flow phenotypes were used for stem and progenitor cell flow: Lin- (lineage negative): B220-, CD3e-, Gr-1-, Ter-119-, CD4-, CD8-, CD19-, CD127-; KL: Lin-, cKit+, Sca1-; KLS: Lin-, cKit+, Sca-1+; KLS-SLAM: Lin-, cKit+, Sca-1+, CD150+, CD48-; GMP: Lin-, cKit+, Sca-1-, CD34+, CD16/32+; CMP: Lin-, cKit+, Sca-1-, CD34+, CD16/32-; and MEP: Lin-, cKit+, Sca-1-, CD34-, CD16/32-.

Analysis was performed using a FACScan (Beckman Coulter) or Gallios flow cytometer (Beckman Coulter). Cell sorting was performed using I-Cyt Synergy II sorter (I-Cyt Technologies). Flow cytometry data were analyzed with FlowJo Software Version 10 (TreeStar), Excel (Microsoft), and Prism 7.02 (GraphPad Software).

Competitive transplantation

Competitive transplantation was performed using 0.5 x 10^6 whole BM cells from indicated donor mice (CD45.2) mixed with 0.5 x 10^6 competitor whole BM cells wild-type CD45.1 (Ly5.1) x CD45.2 mice. Mixture cells were injected intravenously into 6-8 weeks old CD45.1 recipient mice that received 1,100 cGy total body irradiation (Mark 1 Cesum irradiator, J.L. Shepard) 24 hours prior to transplantation. For *Smc3*$^{fl/fl}$*/ERT2-Cre*$^{+/-}$ or *Smc3*$^{fl/+}$*/ERT2-Cre*$^{+/-}$ transplantation, recipient mice were treated with tamoxifen (dissolved in sterile corn oil, Sigma-Aldrich) 6 weeks

post-transplant via oral gavage for 9 doses (3 mg/day/mouse, 3 days/week). Peripheral blood was examined for donor cell chimerism at indicated time points after transplantation. Recipient mice BM were analyzed at the end of experiment.

Colony replating assay

BM cells were harvested and plated in duplicate (10,000 BM cells/plate) in complete mouse methylcellulose medium with stem cell factor, IL-3, IL-6, and Epo (R&D Systems). Colonies were counted on day 7, and cells were collected from methylcellulose in warm Dulbecco modified Eagle medium containing 2% fetal bovine serum, washed, and replated as before. An aliquot of cells was taken for analysis of myeloid (Gr1, CD11b) and mast cell markers (cKit, FcER1) by flow cytometry. This process was repeated for 4 weeks or until colony formation failed.

RNA sequencing of multipotent progenitors and analysis

Multipotent progenitors (KLs; Lin-, cKit+, Sca1-) were sorted from three wild-type or *Smc3*$^{fl/+}$*/Vav1-Cre*$^{+/-}$ mice into DMEM media. Flow cytometry of samples after sorting validated >93% sort accuracy. RNA was extracted from cell pellets using a miRNeasy kit (QIAGEN) and genomic DNA was removed by RNase-Free DNase Set (QIAGEN). RNA was analyzed for degradation using the RNA Nano Chip (Agilent #5067-1521). An input of 300ng was taken forward for each sample using the TruSeq Stranded Total RNA with Ribo-Zero Globin Kit (Ilumina #20020612). Final Libraries were analyzed using the High Sensitivity DNA Chip (Agilent# 5067-4626). All Libraries were pooled and run across 3 lanes of HiSeq4000. RNAseq data were aligned to the human reference with Tophat v2.0.8 (denovo mode, params: --

library-type fr-firststrand --bowtie-version=2.1.0). Expression levels were calculated with Cufflinks v2.1.1 (params: --max-bundle-length 10000000 --max-bundle-frags 10000000).[52]

ATAC-Sequencing of multipotent progenitor and analysis

Chromatin accessibility assays using the bacterial Tn5 transposase were performed using multipotent progenitors (KLs; Lin-, cKit+, Sca1-) sorted from *Smc3*$^{fl/+}$ or *Smc3*$^{fl/+}$*/Vav1-Cre*$^{+/-}$ mice in triplicate. DNA was prepared from 75,000 sorted cells and >93% sorting accuracy verified with post-sort analysis. ATAC libraries were generated exactly as described[53] and pooled and sequenced on a HiSeqX instrument (Illumina) to obtain between 133 and 152 million 2x150 bp paired-end reads. Raw sequencing reads were adapter trimmed with trim galore using cutadapt version 1.8.1 (Martin EMBnet 2011) and then aligned to the mouse reference genome (mm10) using bwa mem (Li H. arXiv:1303.3997v1 (2013)). Peaks in each sample were identified with macs2[54] using the -f BAMPE parameter and then filtered to retain peaks with a q-value <0.01. Peak summits from all samples were merged together with BEDtools merge[55] using parameters to combine summits within 50 bp of each other. Read counts at the merged peak summits were obtained for all samples using the deepTools multiBamSummary command[56] with the minimum mapping quality set to 1, and then processed using DESeq2[57] with default parameters to obtain normalized counts for each peak summit and to perform differential analysis across all peaks between wild-type and mutant mice.

Statistics

Statistical analysis was performed using Prism 7.02 (GraphPad Software) and Excel (Microsoft). Unpaired two-tailed *t*-test, one-way, and two-way ANOVA with Turkey's multiple comparisons tests were performed, as appropriate. *P* values < 0.05 were considered statistically significant.

Error bars represent standard deviation. Data points without error bars have standard deviation below Prism 7.02's limit to display.

Results

Generation of Smc3 conditional knockout mice

To investigate the effects of *Smc3* loss on hematopoiesis, we generated *Smc3* conditionally deficient and haploinsufficient mice using Smc3<tm1a(EUCOMM)Wtsi> mice obtained from EUCOMM (*Smc3*^{trap}). The *Smc3*^{trap} allele has a lacZ-neomycin-gene-trap cassette inserted in intron 4 with two Frt sites on each side of the cassette, and two loxP sites flanking exon 4. The gene trap is predicted to lead to an early transcription stop after splicing into lacZ-neomycin. The conditional knockout *Smc3*^{fl} allele was created by excising the gene-trap cassette with Flp recombinase and was used for further characterizations because homozygous deletion could be achieved using the *Smc3*^{fl} allele (Figure 2.1A). We validated the integration of the loxP sites surrounding exon 4 in the *Smc3*^{fl} allele using whole genome sequencing (Figure 2.1B).

We examined the transcriptional consequences of the *Smc3*^{fl} allele using RNA-Seq and intracellular flow cytometry. In BM cells from three *Smc3*$^{fl/+}$*/Vav1-Cre*$^{+/-}$ mice, nearly 50% (48.4%) of transcripts spliced from exon 3 to exon 5, consistent with deletion of exon 4 while all the wild-type transcripts spliced from exon 3 to exon 4 and exon 4 to exon 5 (Figure 2.1C and Figure 2.7A-F). Analysis of reads spanning exons 3-5 suggests that this results in a frameshift mutation and a stop codon after 59 amino acids, although this truncated protein could not be detected using N-terminal antibodies. Using C-terminal antibodies, intracellular Smc3 protein level was reduced to approximately half of littermate control, as would be expected with a heterozygous allele and confirming *Smc3* haploinsufficiency (Figure 2.1D). In addition, we noted

that Smc3 protein level was regulated during normal hematopoiesis, with higher expression in KLS (Lin-cKit+Sca1+) stem/progenitor cells vs. SLAM (Lin-cKit+Sca1+CD150+CD48-) stem cells (Figure 2.1E). Representative primary intracellular flow data shown in Figure 2.8.

Homozygous Smc3 deletion

To understand whether *SMC3* mutations might have dominant-negative effects or phenocopy loss-of-function effects, we compared the consequences of *Smc3* deficient and haploinsufficient mouse models. We found that hematopoietic homozygous deletion of *Smc3* led to embryonic lethality. In heterozygous $Smc3^{fl/+}/Vav1\text{-}Cre^{+/-}$ intercrosses, we observed 0 out of 75 pups with homozygous *Smc3* alleles (Figure 2.2A). To determine whether the cause of death in $Smc3^{fl/fl}/Vav1\text{-}Cre^{+/-}$ embryos was from hematopoietic failure, we examined E13.5 embryos. Grossly, the $Smc3^{fl/fl}/Vav1\text{-}Cre^{+/-}$ embryos were indistinguishable in size and appearance from other genotypes, except the lack of obvious fetal livers (Figure 2.2B-C). A severe decrease in fetal liver hematopoietic cells was verified by cell count and flow cytometry with near-complete absence of CD45+ Gr1+ CD11b+ cells demonstrating myeloid-biased hematopoietic failure (Figure 2.2D-F).

We investigated somatic homozygous *Smc3* deletion in adult mice using the $Smc3^{fl/fl}/ERT2\text{-}Cre^{+/-}$ mice. *Smc3* deletion was achieved by treating mice with oral tamoxifen (TAM) at 6 weeks of age and reduction in Smc3 protein confirmed with western blot (Figure 2.9A). After 4 doses of TAM, mice were moribund and therefore sacrificed for analysis. Complete blood counts (CBC) data showed the $Smc3^{fl/fl}/ERT2\text{-}Cre^{+/-}$ mice had lower white blood cell counts, percentages of lymphocytes and monocytes, and fewer platelets than TAM-treated littermates (Figure 2.3A). The $Smc3^{fl/fl}/ERT2\text{-}Cre^{+/-}$ mice had decreased spleen weights (Figure 2.3B) and their spleens were smaller in size (Figure 2.9B). Total number of cells in BM, spleen, and

thymus of the *Smc3*$^{fl/fl}$*/ERT2-Cre*$^{+/-}$ mice were significantly reduced in comparison to *Smc3*$^{fl/fl}$ mice after TAM treatment (Figure 2.3C). The reduction of cells occurred across all lineages in the BM (Figure 2.3D), spleen, and thymus (Figure 2.9C-D) of the *Smc3*$^{fl/fl}$*/ERT2-Cre*$^{+/-}$ mice, suggesting complete hematopoietic collapse.

Because activation of ERT2-Cre leads to *Smc3* deletion in a wide range of cells and tissues, we repeated these studies, isolating hematopoietic cells via a competitive transplantation. Equivalent engraftment of transgenic CD45.2+ and competitor CD45.1+ CD45.2+ cells was verified 6 weeks after transplantation. Following tamoxifen-induced *Smc3* deletion, the *Smc3*$^{fl/fl}$*/ERT2-Cre*$^{+/-}$ donor cells were quickly outcompeted, indicating complete loss of hematopoietic stem and progenitor cell (HSPC) functions in the *Smc3*$^{fl/fl}$*/ERT2-Cre*$^{+/-}$ BM. Once again, the effect was most pronounced within the myeloid compartment (Figure 2.3E-F), suggesting that myeloid hematopoiesis is sensitive to *Smc3* deletion and therefore, the AML-associated *SMC3* mutations are unlikely to have simple dominant-negative effects.

Steady-state heterozygous Smc3 deletion

In the ExAC database (exac.broadinstitute.org), no *SMC3* loss-of-function mutations are observed in available human data (0 observed vs. 58.5 expected mutations), suggesting potential embryonic lethality or reduced fitness associated with *Smc3* haploinsufficiency. We, therefore, determined whether *Smc3* haploinsufficiency might be tolerated in mice. Because *CMV-Cre* is X-linked and expressed during early embryogenesis, we examined the ratio of male: female pups and compared difference between genders to determine whether embryonic *Smc3* haploinsufficiency altered hematopoiesis. We found that *Smc3* haploinsufficiency led to a normal number of female pups in *CMV-Cre* intercrosses (Figure 2.10A), and the female pups had no obvious defects in complete blood counts, total number of BM cells, and percentages of HSPCs

and cells in different lineages (Figure 2.10B-E). Hence, embryonic *Smc3* haploinsufficiency could be tolerated and did not grossly perturb steady-state hematopoiesis in mice.

We next assessed the effects of somatic *Smc3* haploinsufficiency on hematopoiesis using the inducible *Smc3*$^{fl/+}$*/ERT2-Cre*$^{+/-}$ mice. *Smc3* haploinsufficiency did not alter the proportions of immunophenotypic HSPCs and cells of different lineages (Figure 2.4A-B).

Furthermore, *Smc3* haploinsufficiency did not increase the number of colonies formed in methylcellulose or the average number of cells per colony, and the *Smc3* haploinsufficient BM cells did not replate beyond two weeks (Figure 2.4C-E). At the end of each week, the colonies on each plate were collected, washed, and characterized by immunophenotype. At the end of week 1, the cells were predominantly Gr1+ CD11b+ for both *Smc3*$^{fl/+}$*/ERT2-Cre*$^{+/-}$ and *Smc3*$^{fl/+}$ genotypes. However, starting week 2, the colonies shifted to cKit+ FcER1+ mast cells. In week 3 and 4, the few colonies left were exclusively mast cells (Figure 2.11A-B). Similar results were observed using BM cells from *Smc3*$^{fl/+}$*/Vav1-Cre*$^{+/-}$ mice.

We performed RNA-Sequencing to measure global gene expression in *Smc3* haploinsufficient hematopoietic progenitors (Lin-cKit+Sca1-) using the constitutive *Smc3*$^{fl/+}$*/Vav1-Cre*$^{+/-}$ model. This model was chosen because it required minimal manipulation of the mice, provided hematopoietic-restricted deletion, and would evaluate steady-state hematopoietic conditions. Multipotent progenitors (KLs) were sorted from age-matched individual wild-type and *Smc3*$^{fl/+}$*/Vav1-Cre*$^{+/}$ mice for RNA-Seq. KLs were selected because *Smc3* haploinsufficiency resulted in severe multi-lineage competitive disadvantage *in vivo*, suggesting potential defect in the functions of *Smc3* haploinsufficient KLs. However, minimal global transcriptional changes were detected. Using t-tests and Significance Analysis of Microarrays (SAM)[26] 149 genes were identified with differential expression in *Smc3* haploinsufficient KLs in comparison to wild-type

controls (most with < 2 fold changes) (Figure 2.4F). KEGG pathway analysis showed significance ($p<0.002$ and $p<0.005$) for progesterone mediate oocyte maturation and toxoplasmosis respectively, but these are not related to hematopoiesis.[27] *Smc3* expression was not observed to be different when analyzed using total reads across the entire gene. However, we observed a two-fold reduction in expression of exon 4 consistent with deletion of this exon (Figure 2.1C).

To determine whether *Smc3* haploinsufficiency might lead to alterations in global chromatin structure that may be biologically relevant, but which did not lead to measurable altered gene transcription, we performed transposase-accessible chromatin sequencing (ATAC-Seq). Chromatin accessibility peaks of the *Smc3*$^{fl/+}$*/Vav1-Cre*$^{+/}$ KLs and littermate *Smc3*$^{fl/+}$controls revealed by ATAC-Seq were not significantly different except for peaks in proximity of three genes: *Vav1*, *Tnpo3*, *Tgfb* (Figure 2.4G). The two-fold difference in *Vav1* was expected for the heterozygous allele and therefore indicated fidelity of data generated by the assay.

Phenotypes of Smc3 haploinsufficiency following competitive transplantations

AML emerges following clonal expansion. Thus, we conducted competitive transplantation using the inducible *ERT2-Cre* model instead of the constituent hematopoietic *Vav1-Cre*, so that complete engraftment could be verified 6 weeks after transplant prior to deletion of the *Smc3* allele. In competitive transplantations, we observed a significant competitive disadvantage in the *Smc3*$^{fl/+}$*/ERT2-Cre*$^{+/-}$ BM cells (Figure 2.5A). Endpoint analysis of BM cells also showed competitive disadvantage in the *Smc3*$^{fl/+}$*/ERT2-Cre*$^{+/-}$ HSPCs and across B and T-cell lineages, implying impaired HSPC functions due to *Smc3* haploinsufficiency in the BM, and not a defect in hematopoietic peripheralization or maturation (Figure 2.5B-F). The competitive disadvantage was observed first in the Gr1 myeloid compartment, perhaps due to higher turn-over of these

cells (Figure 2.5D). To verify the competitive disadvantage observed was not due to toxicity of ERT2-Cre, we repeated the competitive transplantation with the *ERT2-Cre*$^{+/-}$ control mice. The chimerisms of overall CD45.2+ cells and of CD45.2+ cells in all lineages were well-preserved, eliminating the possibility of ERT2-Cre toxicity (Figure 2.12A-B).

The absence of pre-leukemic delayed maturation or augmented self-renewal in *Smc3* haploinsufficient mice was unexpected. Hence, we determined whether *Smc3* haploinsufficiency might increase self-renewal if it occurred in combination with specific conditions of hematopoietic stress. We again observed a competitive disadvantage in the *Smc3*$^{fl/+}$*/ERT2-Cre*$^{+/-}$ BM cells following tamoxifen induction. Intriguingly, the significant myeloid competitive disadvantage was ameliorated at 18 weeks post-transplantation in the pIpC-treated group, whereas it was accelerated in the 5-fluorouracil (5-FU) treated group, although this effect was transient and by week 26 the donor cell population were equivalently reduced (Figure 2.5G-J).

Dnmt3a haploinsufficiency partially abrogated myeloid competitive disadvantage in Smc3 haploinsufficient BM cells

In AML patients, *DNMT3A* mutations co-occurred in approximately one-third of the cases with *SMC3* mutations that assess additional mutations.[4, 5, 15, 16] We therefore asked whether *Smc3* haploinsufficiency might lead to a competitive advantage if it occurred in the background of *Dnmt3a* haploinsufficiency[20].

We observed that with the addition of *Dnmt3a* haploinsufficiency, the severe competitive disadvantage was partially abrogated in the *Smc3*$^{fl/+}$*/ERT2-Cre*$^{+/-}$ myeloid cells, but the significant competitive disadvantage in other lineages remained intact (Figure 2.6A-D). The same phenotype was observed in the *Smc3*$^{fl/+}$*/ERT2-Cre*$^{+/-}$*/Dnmt3a*$^{+/-}$ BM upon endpoint analysis

(Figure 2.6E). Accordingly, even with constitutive *Dnmt3a* haploinsufficiency, *Smc3* haploinsufficiency did not result in competitive growth advantage in hematopoietic cells.

Discussion

AML is a genetically heterogenous disease characterized by clonal expansion of immature myeloblasts. associated with recurrent mutations including the cohesin complex.[25, 4, 5, 15, 16, 28] Mutations in the subunits of the cohesin complex, *SMC1A*, *SMC3*, *RAD21*, and *STAG2*, have been found as early subclonal events in AML, although they are not observed in people with CHIP.[5, 15, 16, 18, 19] In contrast, *DNMT3A* mutations are among the most common initiating mutations in normal karyotype AML patients and the most frequently mutated genes in subjects with CHIP.[25, 29] Cohesin mutations are mutually exclusive of one another and fall into two general categories: mutations in *RAD21* and *STAG2* are mainly truncations and frameshifts, whereas the majority of mutations in *SMC1A* and *SMC3* are missense. In AML, cohesin mutations are not associated with genomic instability, complex karyotypes, or monosomy karyotypes, suggesting alternative pathologic mechanisms.[4, 5, 15]

To understand whether leukemia-associated *SMC3* missense mutations might have dominant-negative activities or phenocopy loss-of-function effects, we compared the consequences of *Smc3* deficiency and *Smc3* haploinsufficiency on murine hematopoiesis using conditionally deleted strategies. We began by validating the *Smc3* allele using whole genome sequencing, RNA-Seq, and intracellular flow cytometry, which demonstrated correct integration, splicing of approximately 50% of alleles around exon 4 leading to a frameshift mutation and an early nonsense mutation, and reduced protein levels. Our findings suggest that leukemia-associated *SMC3* mutations are unlikely to have novel dominant negative activities because homozygous *Smc3* deletion was incompatible with embryonic (Figure 2.2) or adult hematopoiesis (Figure

2.3). In these experiments, we observed the effects first in the myeloid compartment. However, because myeloid cells have a shorter half-life than other hematopoietic cell types, the augmented temporal phenotypes observed in these cell fractions may be influenced by greater turn-over. Collectively, these studies demonstrate that *Smc3* is indispensable for embryonic and adult hematopoiesis and normal HSPC functions. Similar severe consequences for *Smc3* deficiency[30] and *Rad21* deficiency[31] have been observed, and thus cohesin genes appear to be essential in hematopoietic cells.

Leukemia-associated *SMC3* mutations are observed across all domains of the protein, and nearly one third are nonsense or splice-site variants, suggesting that many of these mutations are likely to be associated with loss of function. Therefore, we investigated the effects of *Smc3* haploinsufficiency on murine hematopoiesis. Because these mutations are associated with leukemia, we predicted that *Smc3* haploinsufficiency would augment colony forming capacity and provide hematopoietic cells a competitive advantage. However, we observed neither phenotype. Following *Smc3* haploinsufficiency induced with three different Cre models (*CMV-Cre*, *Vav1-Cre*, and *ERT2-Cre*) we observed normal CBCs, normal bone marrow hematopoietic population distributions, and normal colony forming (Figure 2.4A-E). We further examined expression signatures and ATAC-Seq under these steady-state conditions in *Vav1-Cre* mice where hematopoietic cells have consistently undergone heterozygous deletion and external perturbations are minimized; we observed little global dysregulation of gene expression or chromatin structure (Figure 2.4F-G). In both studies, internal markers (*Smc3* expression and peaks within the *Vav1* locus) served as controls and markers of the expected dynamic range.

In contrast, under conditions of chimeric competition, *Smc3* haploinsufficiency actually led to competitive disadvantage *in vivo*, with progressive population loss over time (Figure 2.5A-F). In

these studies, *Smc3* deletion was induced using *ERT2-Cre* following a period of 6 weeks post-transplant to facilitate engraftment and stem cell homeostasis prior to deletion. Under these conditions, activation of ERT2-Cre alone does not lead to stem cell toxicity and competitive disadvantage (Figure 2.12A-B), whereas activation of ERT2-Cre just prior to transplantation does.[32] Analysis of BM populations at the end of the study suggested reduction of populations with *Smc3* haploinsufficiency across progenitor and mature cell types, eliminating the possibility that *Smc3* haploinsufficiency led to a profound maturation block that prevented leukocyte peripheralization. The competitive disadvantage induced by somatic *Smc3* acquisition was unexpected. Therefore, we determined whether specific forms of hematopoietic stress might enable a competitive advantage that could facilitate stem cell expansion and ultimately enable leukemogenesis. We again observed a competitive disadvantage that persisted following a stem cell stressor (5-FU exposure) and an inflammatory stressor (pIpC exposure) (Figure 2.5G-J). Finally, because *SMC3* mutations may not be the first acquired mutation during leukemogenic chronicity, we investigated whether *Dnmt3a* haploinsufficiency might facilitate *Smc3* phenotypes. Germline *Dnmt3a* haploinsufficiency partially abrogated the myeloid competitive disadvantage of somatically acquired *Smc3* haploinsufficiency (Figure 2.6), suggesting that *SMC3* mutations may require pre-existing cooperating mutations to facilitate their action. Additionally, these studies do not eliminate the possibility that the frequently observed *SMC3* missense mutations may possess novel gain-of-function activity not accessed in these *Smc3* haploinsufficient studies.

Thus, under conditions of homeostasis, where all hematopoietic cells have *Smc3* haploinsufficiency, murine *Smc3* haploinsufficiency does not appear to grossly dysregulate hematopoietic feedback mechanisms or alter normal hematopoietic maturation or self- renewal

ex vivo. However, under conditions of competitive transplantation, we observed a disadvantage in hematopoietic cells across both myeloid and lymphoid lineages suggesting reduced cell production at a multipotent progenitor level.

These results contrast with previously published work using either knock-down strategies in CD34+ cord blood cells or using Mx1-Cre activation with pIpC. Specifically, knocking down of *Smc3* using shRNA or *RAD21* and *SMC1A* mutants have been shown to increase self-renewal in human cord blood $CD34^+$ HSPCs *ex vivo*.[33, 34] *Smc3* haploinsufficiency induced by Mx1-Cre exhibited shifts in hematopoietic cell populations, colony forming, and competitive transplantation advantage when deleted using Mx1-Cre two weeks after transplantation[30]. These data suggest that differences in the models may interact with the biological consequences of *Smc3* reduction through yet undefined mechanisms.

In addition, It is worth noting that other MDS or AML-associated mutations such as U2AF1[35, 36], SRSF2[37, 38], SF3B1[39, 40, 41, 42, 43], ASXL1[44, 45] are associated with having competitive disadvantage, which may seem counterintuitive for recurring leukemia mutations observed in patients, but appears to be recurrent biology.

The observed defects in hematopoietic cells with *Smc3* deficiency and haploinsufficiency may reflect population data from the ExAC database, where germline cohesin mutations are observed at lower than expected frequencies, suggesting a significant disadvantage in population fitness. No loss-of-function variants are detected in *SMC3, SMC1A, STAG2*, or *RAD21* (based on statistical models of case numbers and gene size, the expected numbers of loss-of-function variants were 58.5, 32, 42.7, and 21.8, respectively). Missense variants were also significantly underrepresented in *SMC3, SMC1A*, and *STAG2*, but not in *RAD21* ($z = 6.25, 6.59, 5.11$, and 2.76; more positive scores indicate fewer variants observed than expected). Of the published

AML-associated missense mutations, only 1 is reported in ExAC (K795E occurring in 3/121,384 alleles), although synonymous changes (R155R, Q367Q, R391R), and alternative amino acid changes (N604S and I1001L) are noted.[46]

In recent decades, mutations in cohesin complex genes have been associated with genetic syndromes, referred to as cohesinopathies. Several important features differ between cohesinopathies and AML-associated cohesin mutations. Mutations associated with cohesinopathy tend to be in cohesin adapter proteins, such as *NILS*, *HDAC8,* and *ESCO2,* with fewer mutations observed in *SMC3*, *SMC1A, RAD21,* or *STAG2*.[47] Cohesinopathies are associated with facial dysmorphism, cognitive impairment, pre- and post-natal growth delay, and multi-organ involvement and the clinical manifestations appear milder in cases with *SMC3* and *SMC1A* mutations, compared with *NIPBL* mutations.[48] Hematopoietic alterations have not been reported with cohesinopathy, nor has the development of AML. Likewise, the accumulation of aneuploidies and other chromosomal aberrations has been a recurrent feature of cohesinopathy, whereas this phenotype is largely absent in cohesin mutated AML cases, which typically present with normal karyotypes. Intriguingly, copy number gains of *STAG2* or *SMC1A* also have been associated with cohesinopathy phenotypes,[49, 50, 51] suggesting that there may be a critical window of adequate cohesin activity and that alterations in either direction may be detrimental. In contrast to these human data, in our mouse model, germline heterozygous *Smc3* deletion was tolerated using X-linked *CMV-Cre*, which is expressed during early embryogenesis. The heterozygous *Smc3*$^{+/-}$*/CMV-Cre*$^{+/-}$ female progenies had no obvious developmental defects and had normal hematopoietic homeostasis in the bone marrow (Figure 2.10). The normal hematopoietic cell numbers and differentials in the mice reflect the maintained hematopoiesis of cohesinopathies, whereas the normal number of *Smc3* haploinsufficient pups contrasts with the

near absence of cohesin mutations in the human population data. This discrepancy may be due to differences between mouse and human biology; alternatively, cohesinopathy mutations may be associated with gain of function activity not recapitulated with this allele, or activity not related directly to the SMC1A/SMC3 complex.

In summary, we did not observe evidence of impaired differentiation or augmented self-renewal *ex vivo* or *in vivo* when *Smc3* haploinsufficiency was generated using *CMV-Cre, ERT2-Cre*, and *Vav1-Cre.* Instead, *Smc3* haploinsufficiency was associated with competitive disadvantage, with an early bias towards phenotypes in the myeloid compartment. In AML patients, *SMC3* mutations are typically early, but not initiating, genetic events. These data also suggest that pre-existing mutations may be required to enable leukemogenic consequences of *SMC3* mutagenesis and to permit productive clonal expansion. Future studies are needed to determine the combination of cooperating mutations that predispose HSPCs to *SMC3*-induced leukemic transformation and clonal dominance.

Figure Legends

Figure 2.1. Generation of *Smc3* conditional deficient mice and allele validation.

(A) *Smc3* haploinsufficient mouse model ($Smc3^{trap/+}$) was obtained from the European Mouse Mutagenesis Program (EUCOMM). *Smc3* conditionally deficient mice were generated by removing the gene-trap cassette, which retains the loxP sites flanking exon 4 ($Smc3^{fl/+}$) and crossing these mice with either $Vav1\text{-}Cre^{+/-}$ or $ERT2\text{-}Cre^{+/-}$ to delete the allele ($Smc3^{\Delta/+}$). All mice are on the C57BL/6J background. (B) Whole genome sequencing validation of $Smc3^{fl}$ integration sites. (C) RNA-Seq data of the $Smc3^{fl/+}/Vav1\text{-}Cre^{+/-}$ mice showed 227 transcripts spliced from exon 3 to 4 and then 313 transcripts from exon 4 to 5 while 279 transcripts from the other allele spliced from exon 3 to 5 (average data from 3 mice). (D) *Smc3* haploinsufficiency was confirmed by reduced Smc3 level in the bone marrow (BM) cells of the $Smc3^{fl/+}/Vav1\text{-}Cre^{+/-}$ mice measured using intracellular flow cytometry (n=5). *Denotes statistical significance by t-test. ***$p<0.001$. (E) Smc3 level is significantly higher in KLS (Lin-cKit+Sca1+) cells and progenitor populations than Lin- and SLAM (Lin-cKit+Sca1+CD48-CD150+). *Denotes statistical significance by one-way ANOVA with Turkey's multiple comparisons test. ****$p<0.0001$.

Figure 2.2. Embryonic hematopoietic *Smc3* deletion.

(A) No $Smc3^{fl/fl}/Vav1\text{-}Cre^{+/-}$ pups were observed following $Smc3^{fl/+}$ and $Smc3^{fl/+}/Vav1\text{-}Cre^{+/-}$ intercrosses (n=11 litters). *Denotes statistical significance by Chi-square test. ****$p<0.0001$. The E13.5 $Smc3^{fl/fl}/Vav1\text{-}Cre^{+/-}$ embryos (B) lacked gross fetal livers but retain (C) normal body weight compared with littermates. (D) The E13.5 $Smc3^{fl/fl}/Vav1\text{-}Cre^{+/-}$ embryos had decreased total fetal liver cells and (E)fetal liver hematopoietic cells (CD45.2+). (F) Myeloid

(Gr1+CD11b+) cells were reduced and increased proportions of B220+ and CD3e+ lymphocytes were observed in E13.5 *Smc3*$^{fl/fl}$*/Vav1-Cre*$^{+/-}$ fetal livers compared to littermate controls. (C-F) n=7 embryos per group, *Denotes statistical significance by one-way (D-E) and two-way (F) ANOVA with Turkey's multiple comparisons test. ***p<0.01*, ****p<0.0001*.

Figure 2.3. Homozygous somatic *Smc3* deletion.

(A) *Smc3*$^{fl/fl}$*/ERT2-Cre*$^{+/-}$ and *Smc3*$^{fl/fl}$ littermate control mice were treated with 4 doses of tamoxifen (3 mg orally on days 1st, 3rd, 5th, 8th and analyzed on day 8, n = 4 mice in each group) (A) Peripheral blood analysis. (B) Body weight and spleen weight. (C) Total number of cells in the bone marrow, spleen, and thymus. (D) Analysis of lineage percentages within total bone marrow cells. (A-D) n=4 mice per group, *Denotes statistical significance by t-test. **p<0.05*, ***p<0.01*, ****p<0.001*. (E) Experimental schema of the *Smc3*$^{fl/fl}$*/ERT2-Cre*$^{+/-}$ competitive transplantation. (F) Recipient mice were treated with tamoxifen after 6-week engraftment. After tamoxifen-mediated deletion, *Smc3* deficient cells were rapidly outcompeted, with earliest cell loss in the Gr1+ myeloid compartment, showing as complete competitive disadvantage. *Denotes statistical significance by 2-way ANOVA with Turkey's multiple comparisons test. *****p<0.0001*. (E-F) n=10 mice per group.

Figure 2.4. Hematopoietic *Smc3* haploinsufficiency.

(A and B) Distribution of bone marrow stem, progenitor, and lineage populations in *Smc3*$^{fl/+}$*/ERT2-Cre*$^{+/-}$ and littermate *Smc3*$^{fl/+}$ mice following 9 doses of tamoxifen (n=6 mice per group). (C) Experimental schema of serial replating assay. (D-E) Colony numbers and average cells per colony on indicated week of plating in methylcellulose (n=4 mice per group). *Denotes statistical significance by t-test. **p<0.05*, ***p<0.01*. (F) Expression analysis by RNA-Seq data

of KL (Lin-cKit+Sca1-) bone marrow cells from *Smc3*$^{fl/+}$*/Vav1-Cre*$^{+/-}$ compared to wild-type cells (n=3 mice per group). (G) Comparison of relative peak intensity identified by ATAC-Seq of KL bone marrow cells from relative peak *Smc3*$^{fl/+}$*/Vav1-Cre*$^{+/-}$ compared to wild-type cells (n=3 mice per group)

Figure 2.5. Competitive transplantation of *Smc3* haploinsufficient bone marrow cells.

(A-F) Competitive repopulation assay using *Smc3*$^{fl/+}$*/ERT2-Cre*$^{+/-}$ BM cells and littermate *Smc3*$^{fl/+}$ BM cells with competitor CD45.1 x CD45.2 bone marrow cells (3 donor mice per group and 10 recipient mice per group). Following 6 weeks of engraftment, equal peripheral chimerism was validated and recipient mice were treated with 9 doses of tamoxifen. (B - C) Following 42 weeks, bone marrow chimerism was analyzed (n=3 mice per group). *Denotes statistical significance by t-test, *$p<0.05$, **$p<0.01$, ****$p<0.0001$. (D-F) At interval time-points during follow-up peripheral blood chimerism was evaluated within the Gr1, B220, and CD3e compartments. *Denotes statistical significance by 2-way ANOVA with Turkey's multiple comparisons test, *$p<0.05$, **$p<0.01$, ****$p<0.0001$. (G-J) Competitive repopulation assay of *Smc3*$^{fl/+}$*/ERT2-Cre*$^{+/-}$ BM cells under hematopoietic stresses (n=10). As before, recipient mice were treated with 9 doses of tamoxifen after 6-week engraftment. PIpC and 5-FU were given 16 weeks post-transplant, respectively. *Denotes statistical significance by 2-way ANOVA with Turkey's multiple comparisons test, **$p<0.01$, ***$p<0.001$, ****$p<0.0001$.

Figure 2.6. Effect of *Dnmt3a* haploinsufficiency on competitive disadvantage in *Smc3* haploinsufficient BM cells.

(A-D) Competitive repopulation assay of *Smc3*$^{fl/+}$*/ERT2-Cre*$^{+/-}$*/Dnmt3a*$^{+/-}$ BM cells and indicated littermate controls (n=10 mice per group). As in Figure 2.5, total bone marrow cells were

allowed to engraft for 6 weeks and equivalent chimerism was validated before treatment of all cohorts with 9 doses of tamoxifen (3 mg/day). Peripheral blood chimerism was evaluated by flow cytometry at indicated time points. (E) Bone marrow chimerism assessed by flow cytometry 26 weeks after engraftment. KL: Lin-cKit+Sca1-. (n=4 mice in each group). *Denotes statistical significance by 2-way ANOVA with Turkey's multiple comparisons test. *$p<0.05$, **$p<0.01$, ***$p<0.001$, ****$p<0.0001$.

Figure 2.7. Splicing analysis of exon 3 to exon 5 in wild-type and *Smc3*$^{fl/+}$*/Vav1-Cre*$^{+/-}$ KL cells.

Lin-cKit+Sca1- bone marrow cells were subjected to RNA-Seq (Figure 2.1C and 2.4F). Schema indicates total number of reads spanning each splice junction from indicated mice.

Figure 2.8. Representative plot of intracellular flow cytometry data (Figure 2.1D-E).

(A) Percentages of Smc3+ cells (left) and mean fluorescence intensity (MFI) (right) of WBM from mice used in Figure 2.1D. (B). Overlay of the two MFI plots. (C) Percentages of Smc3+ cells (left) and mean fluorescence intensity (MFI) (right) of lin-, KLS, SLAM, and KL cells from mouse used in Figure 2.1E. The height of the peak is proportional to the number of events collected.

Figure 2.9. Analysis of homozygous somatic *Smc3* deletion.

(A) Western blot of Smc3 in total bone marrow cells following 4 doses of tamoxifen in indicated mice. (B) Image of spleens from *Smc3*$^{fl/+}$*/ERT2-Cre*$^{+/-}$ and *Smc3*$^{fl/+}$ littermate controls following 4 doses of tamoxifen. (C-D) Proportion of bone marrow cells and thymocytes with indicated immunophenotypes following 4 doses of tamoxifen in *Smc3*$^{fl/+}$*/ERT2-Cre*$^{+/-}$mice and littermate

controls (n=4 mice per group), *Denotes statistical significance by t test, *$p<0.05$, **$p<0.01$, ***$p<0.001$, ****$p<0.0001$.

Figure 2.10. Analysis of germline heterozygous *Smc3* deletion.

(A) Numbers of male vs. female pups generated from *Smc3*$^{fl/fl}$ and *CMV-Cre*$^{+/-}$ intercrosses (of note, *CMV-Cre* is X-linked). (B-E) Complete blood counts, total number of bone marrow cells, percentages of HSPCs, and cells of myeloid (Gr1+ CD11b+), B cells (B220+ CD4-), and erythroid cells (Ter119+ CD71+) in *Smc3*$^{+/-}$*/CMV-Cre*$^{+/-}$ females and *Smc3*$^{fl/+}$ littermate males, (n=4 mice per group).

Figure 2.11. Immunophenotypic analysis of colonies in serial replating assay *ex vivo* (Figure 2.4C-E).

(A) Percentages of Gr1+ CD11b+ and cKit+ FcER1+ cells in the *Smc3*$^{fl/+}$ and *Smc3*$^{fl/+}$*/ERT2-Cre*$^{+/-}$ colonies week 1-4 respectively (n=4 mice per group). (B) Representative plot of the *Smc3*$^{fl/+}$ and *Smc3*$^{fl/+}$*/ERT2-Cre*$^{+/-}$ colonies week 1-4. *Denotes statistical significance by t test, *$p<0.05$.

Figure 2.12. Competitive transplantation of *ERT2-Cre*$^{+/-}$ bone marrow cells.

(A) Competitive repopulation assay using *ERT2-Cre*$^{+/-}$ BM cells with competitor CD45.1 x CD45.2 bone marrow cells (3 donor mice and 5 recipient mice). Following 6 weeks of engraftment, equal peripheral chimerism was validated and recipient mice were treated with 9 doses of tamoxifen. (B) At interval time-points during follow-up peripheral blood chimerism was evaluated within the Gr1, B220, and CD3e compartments.

Figure 2.1. Generation of *Smc3* conditional deficient mice and allele validation.

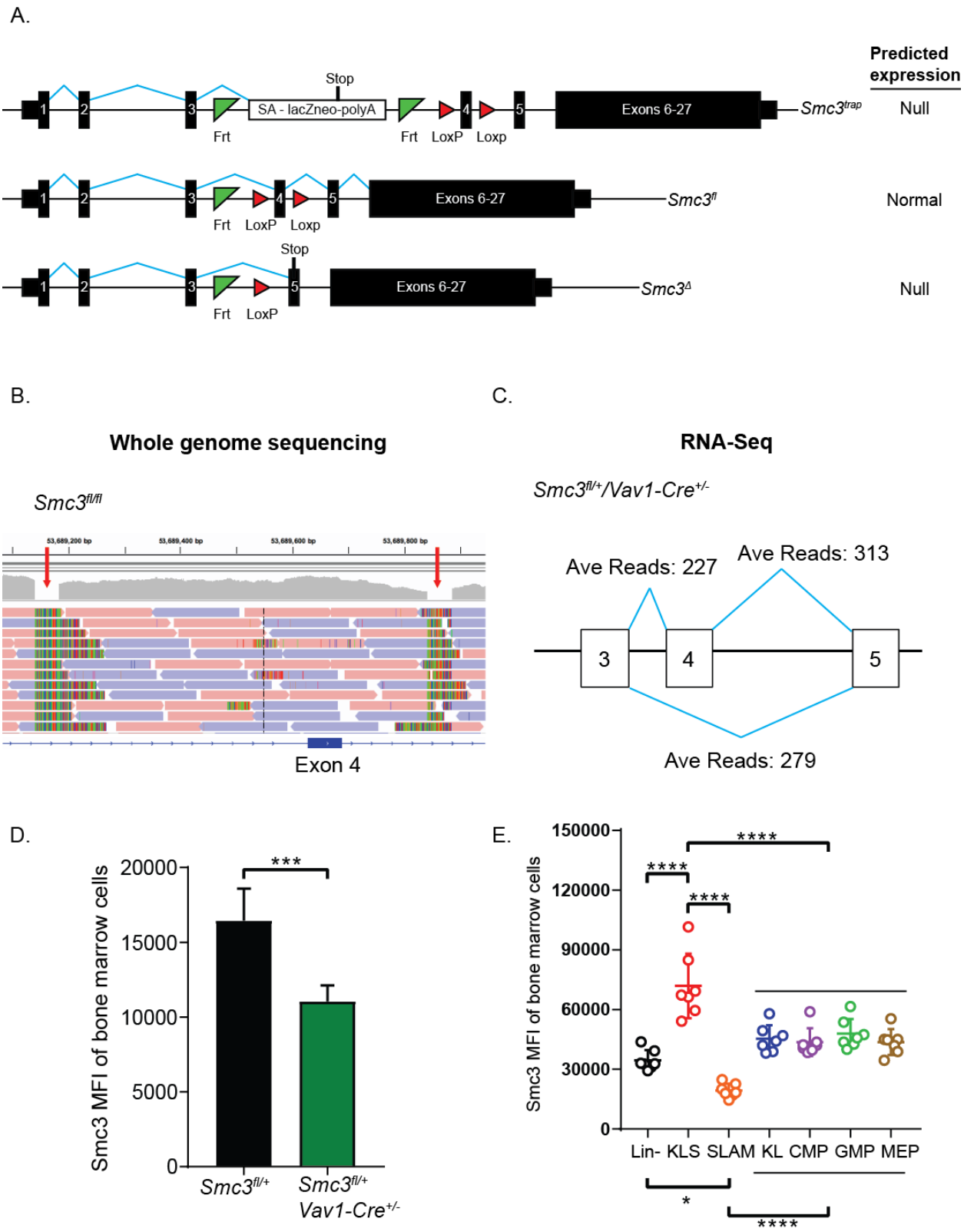


Figure 2.2. Embryonic hematopoietic *Smc3* deletion.

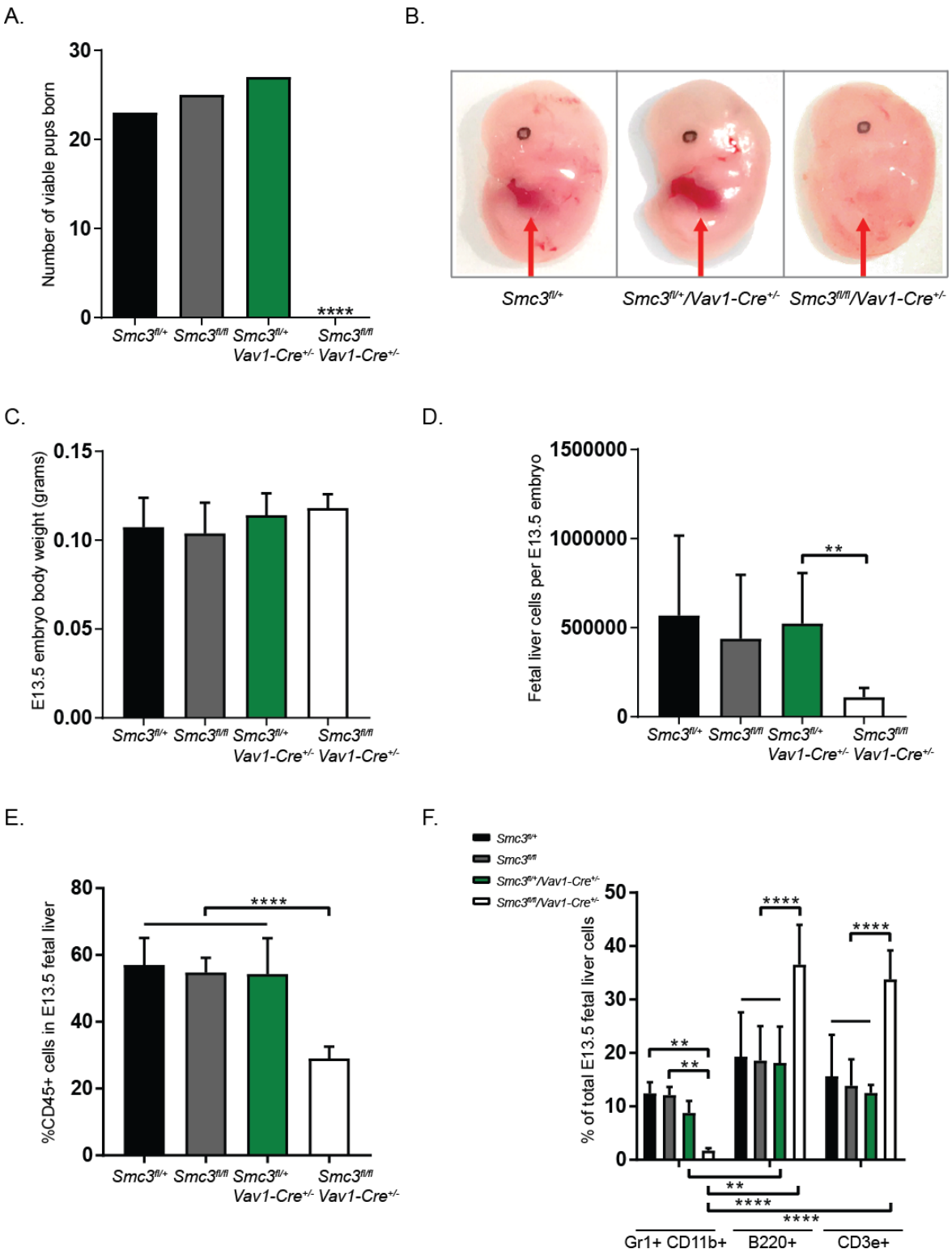


Figure 2.3. Homozygous somatic *Smc3* deletion.

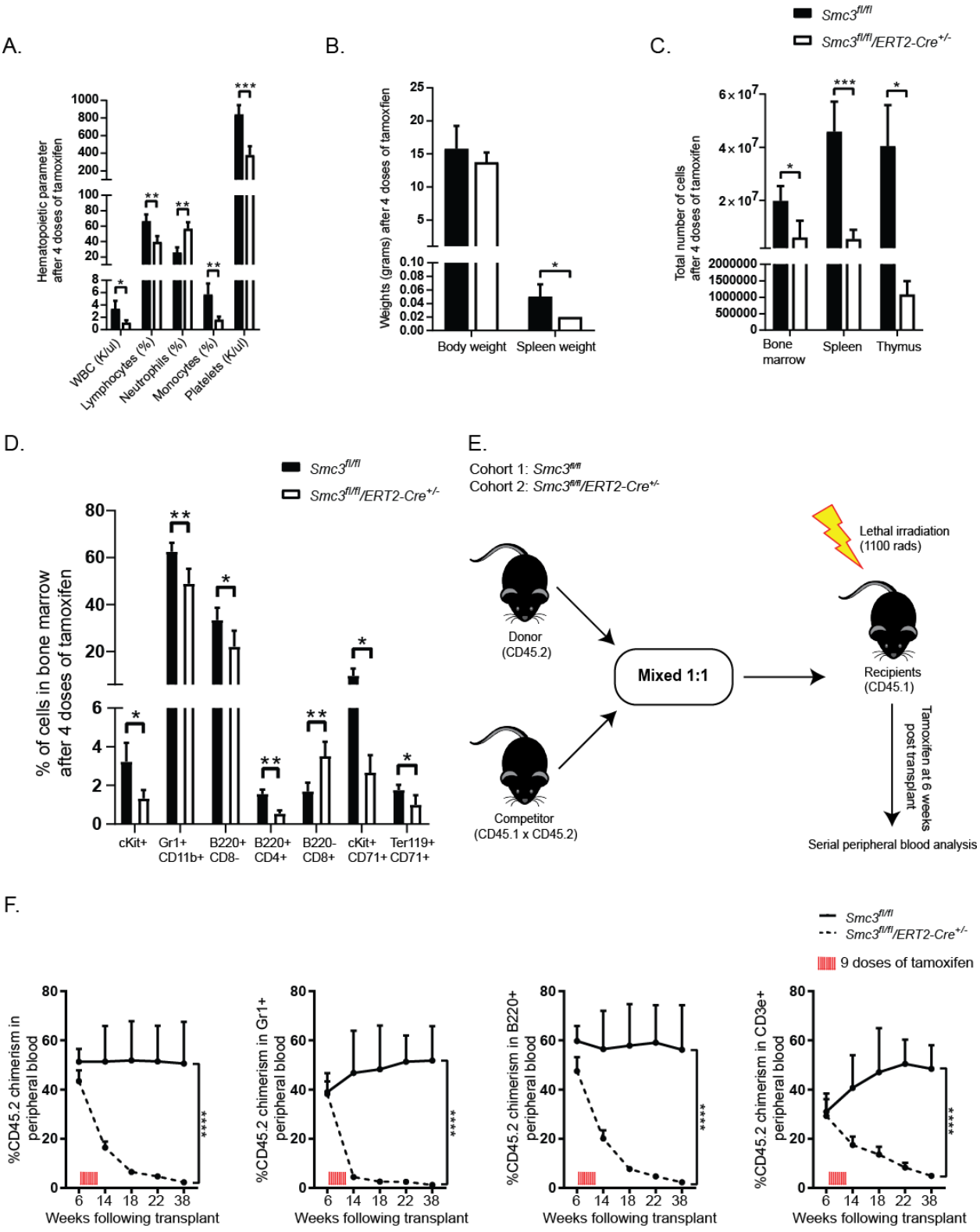


Figure 2.4. Hematopoietic *Smc3* haploinsufficiency.

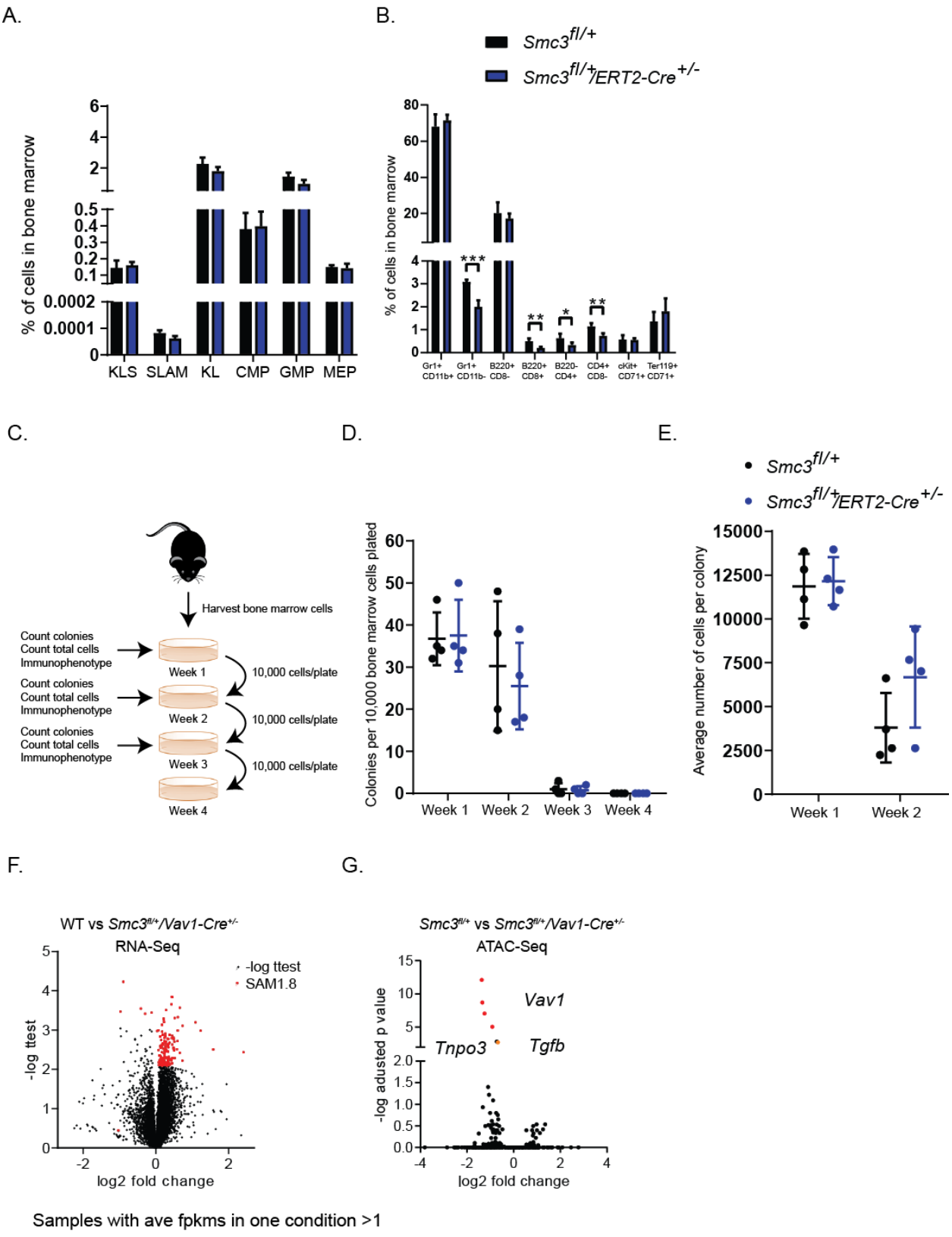


Figure 2.5. Competitive transplantation of *Smc3* haploinsufficient bone marrow cells.

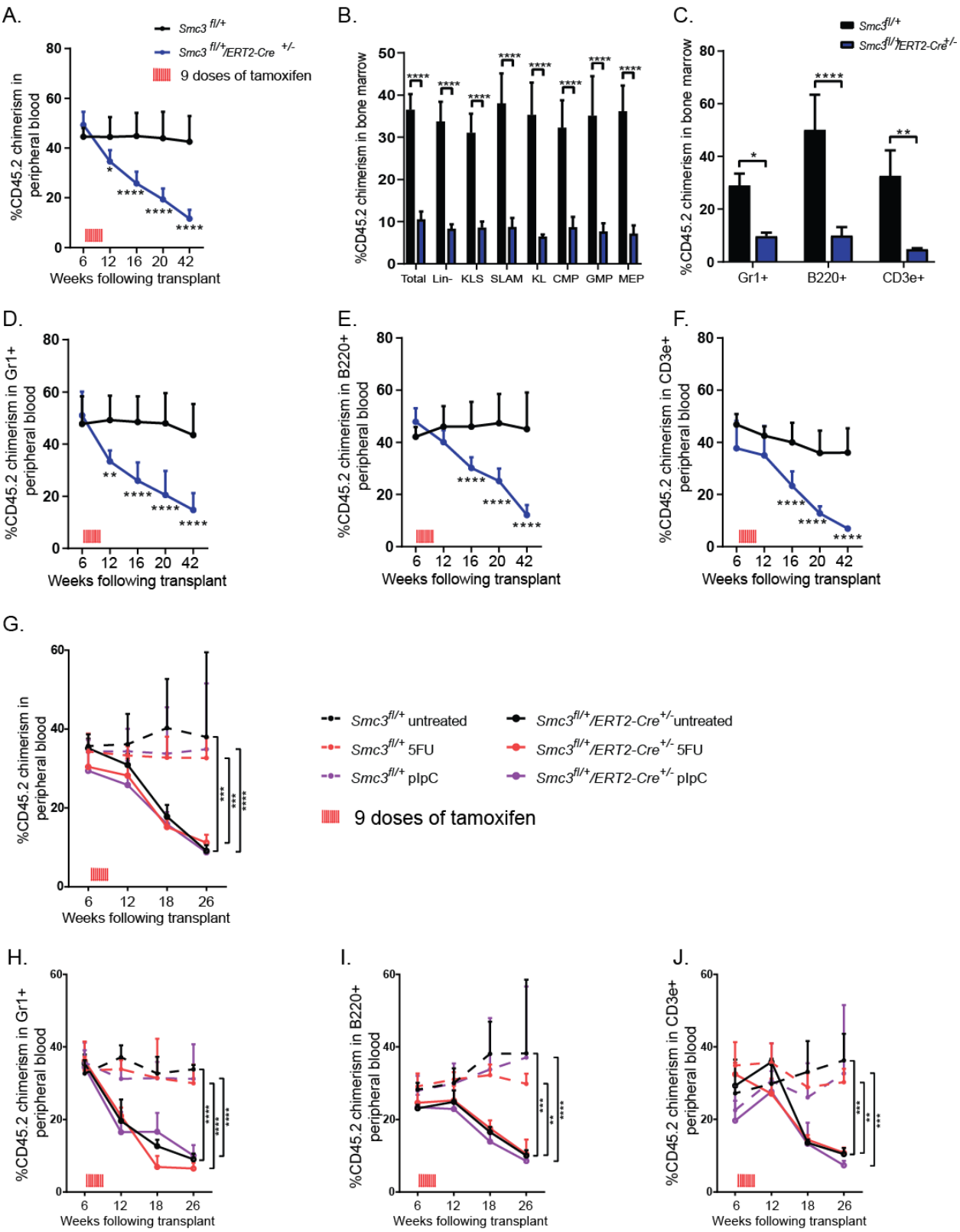


Figure 2.6. Effect of *Dnmt3a* haploinsufficiency on competitive disadvantage in *Smc3* haploinsufficient BM cells.

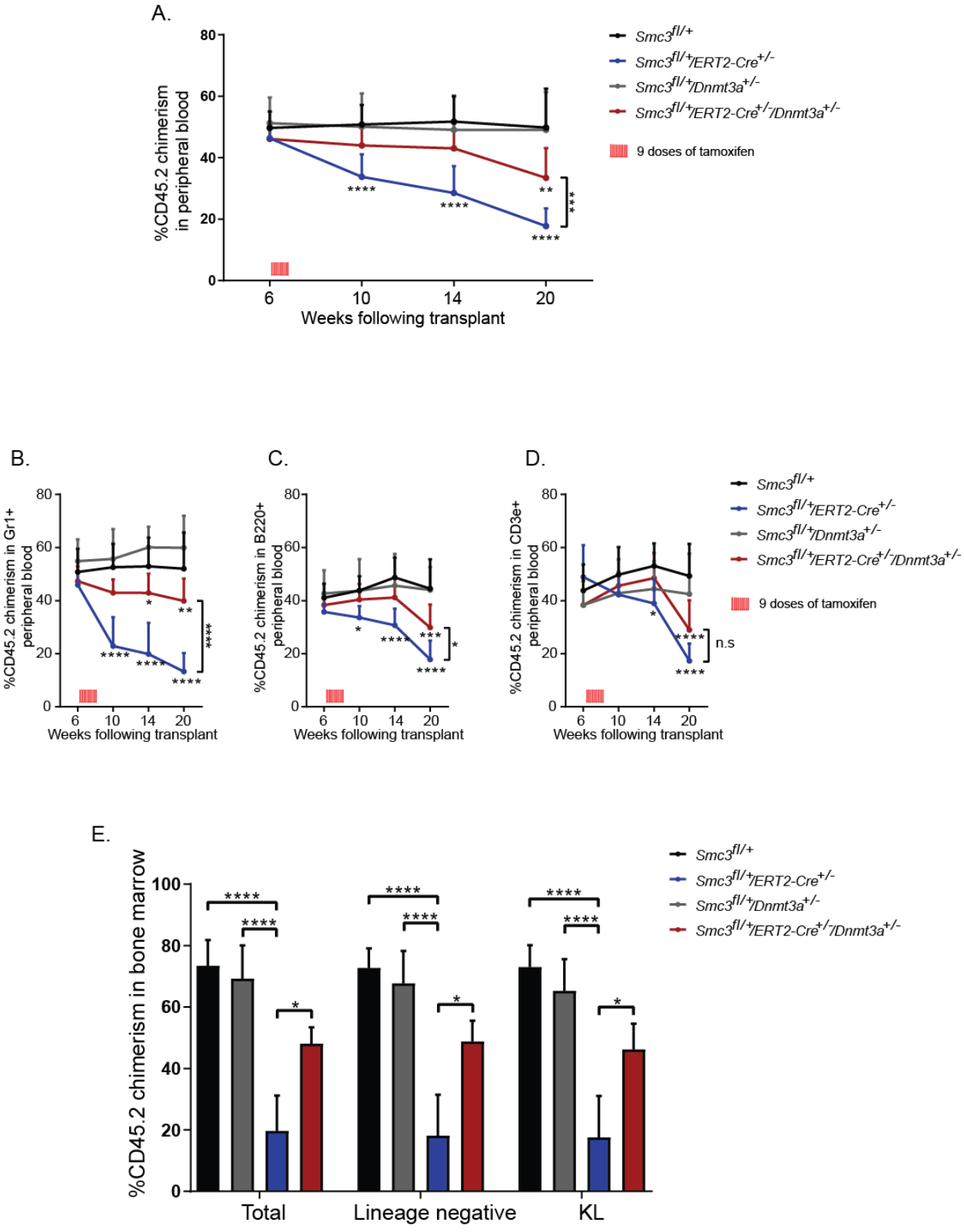
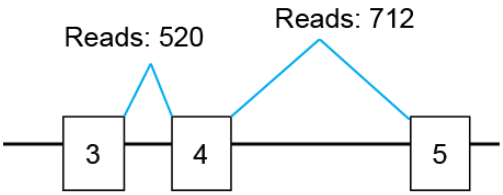


Figure 2.7. Splicing analysis of exon 3 to exon 5 in wild-type and *Smc3*$^{fl/+}$*/Vav1-Cre*$^{+/-}$ KL cells.

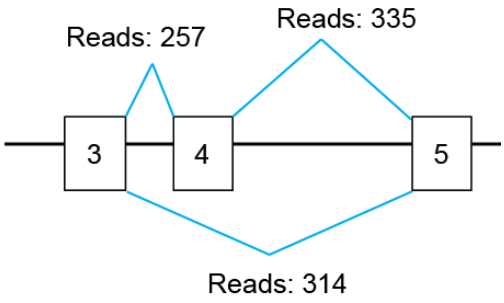
A.

Mouse 10433, Wild type



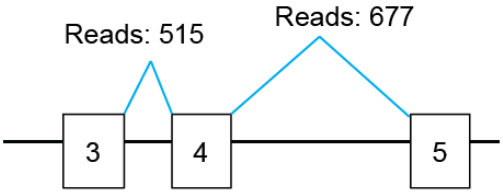
B.

Mouse 10410, *Smc3*$^{fl/+}$*/Vav1-Cre*$^{+/-}$



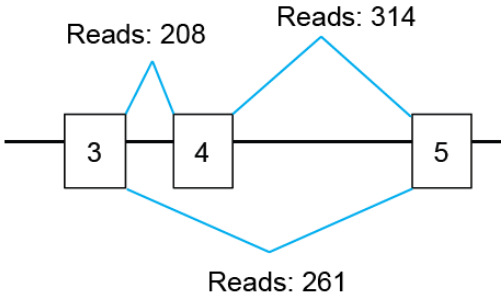
C.

Mouse 10442, Wild type



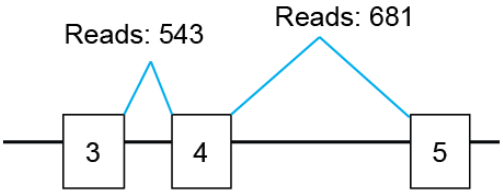
D.

Mouse 10411, *Smc3*$^{fl/+}$*/Vav1-Cre*$^{+/-}$



E.

Mouse 10841, Wild type



F.

Mouse 10412, *Smc3*$^{fl/+}$*/Vav1-Cre*$^{+/-}$

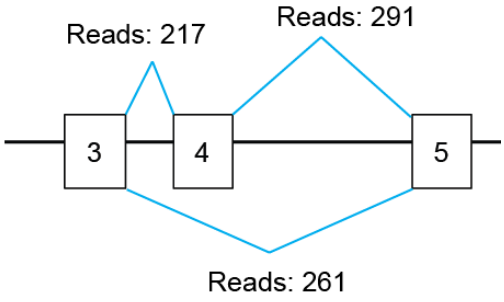
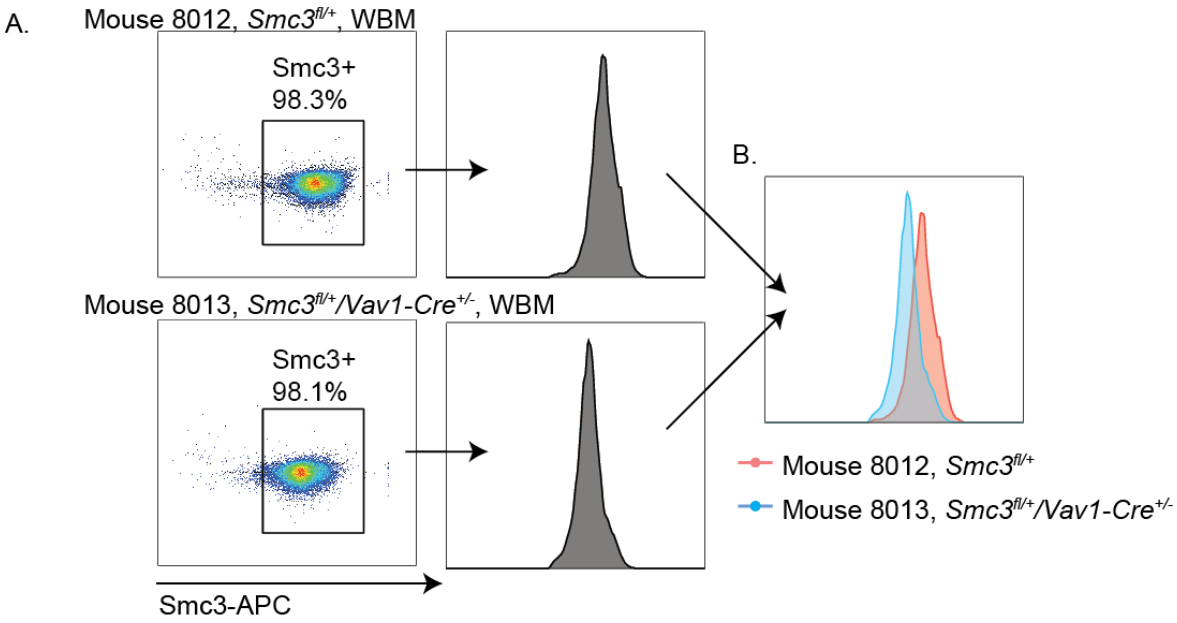


Figure 2.8. Representative plot of intracellular flow cytometry data (Figure 2.1.D-E).



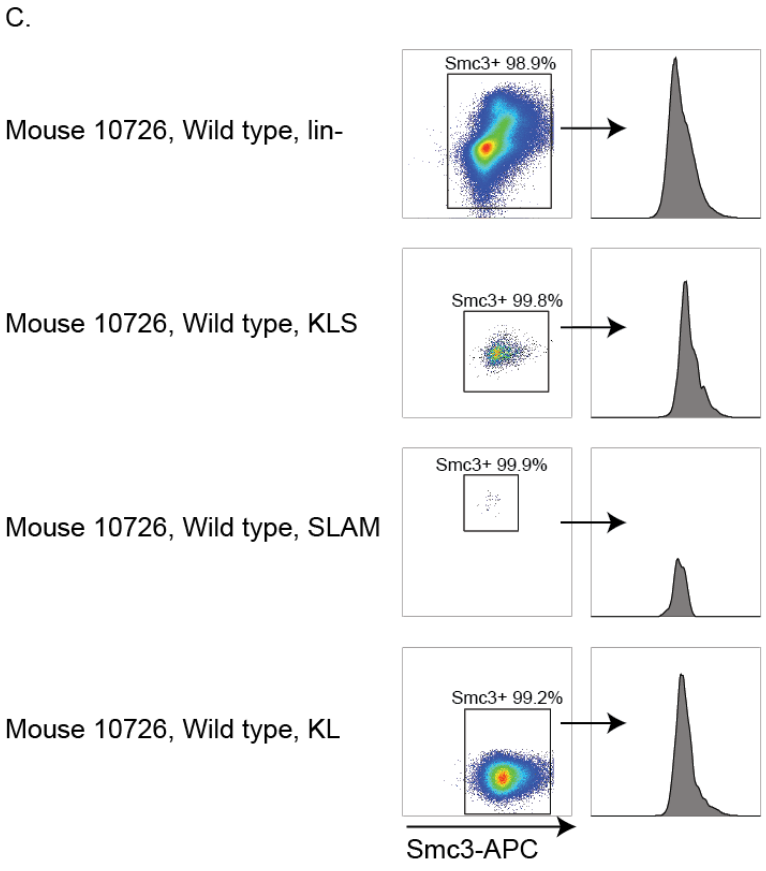


Figure 2.9. Analysis of homozygous somatic *Smc3* deletion.

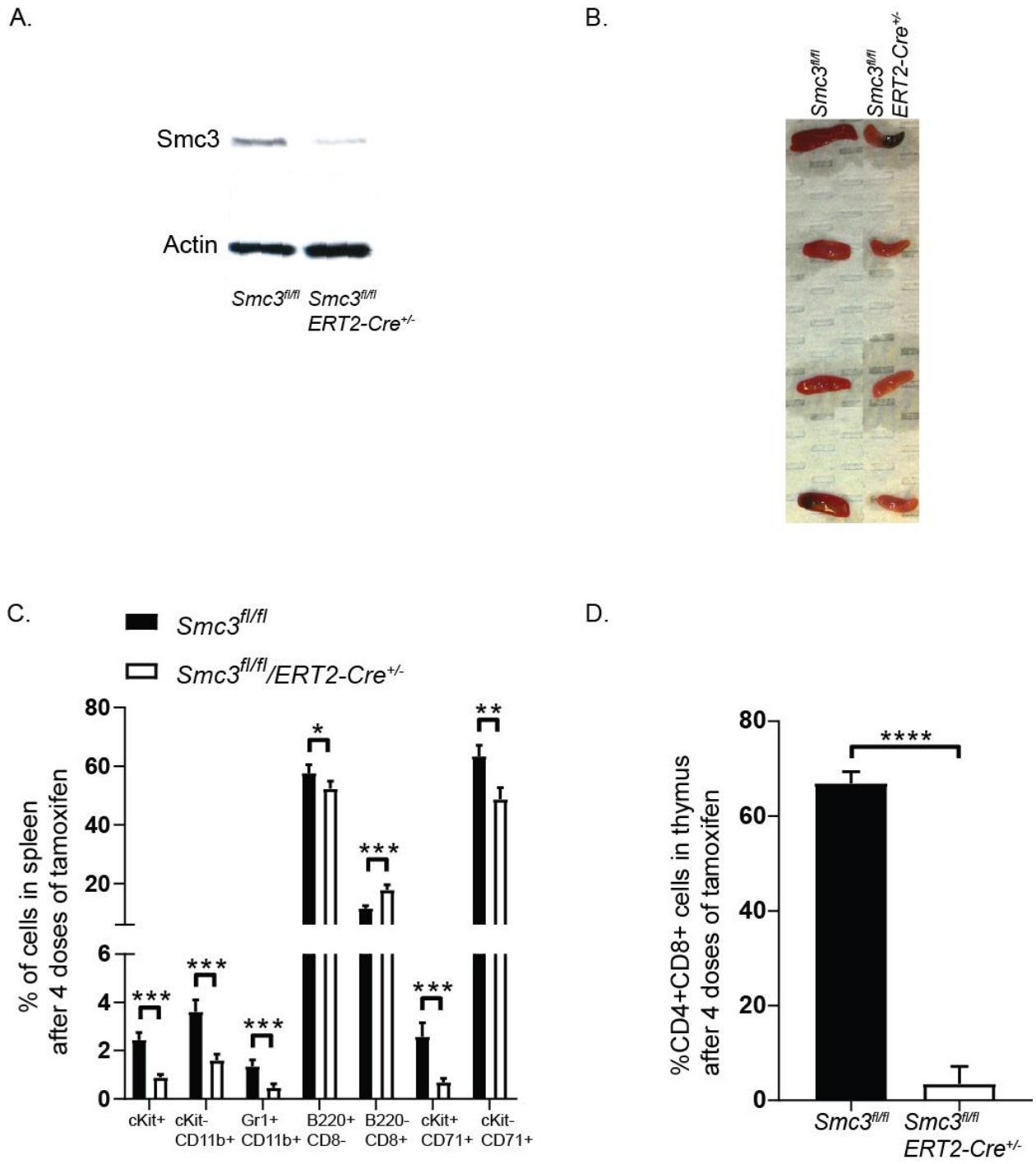


Figure 2.10. Analysis of germline heterozygous *Smc3* deletion.

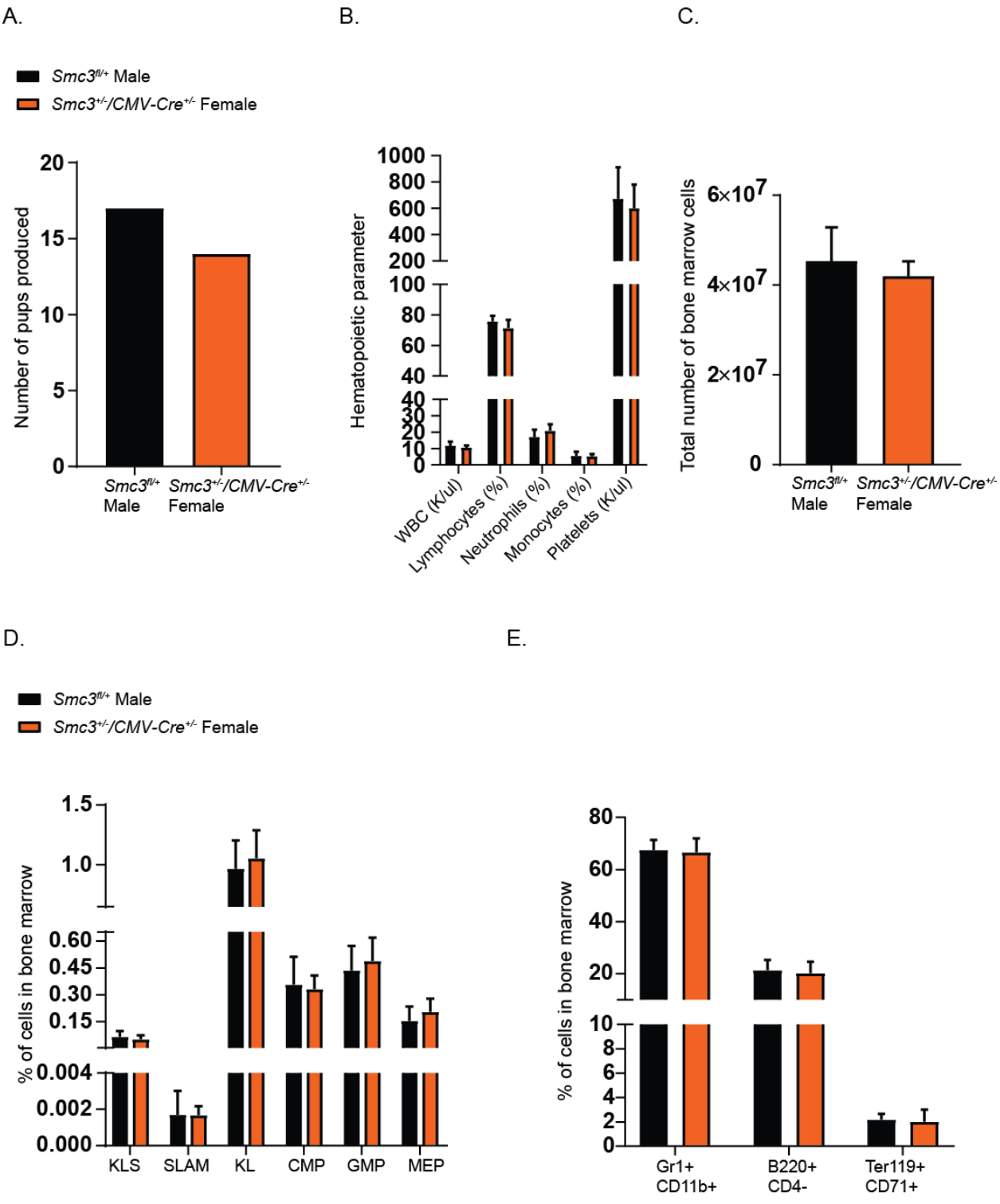


Figure 2.11. Immunophenotypic analysis of colonies in serial replating assay *ex vivo* (Figure 2.4.C-E).

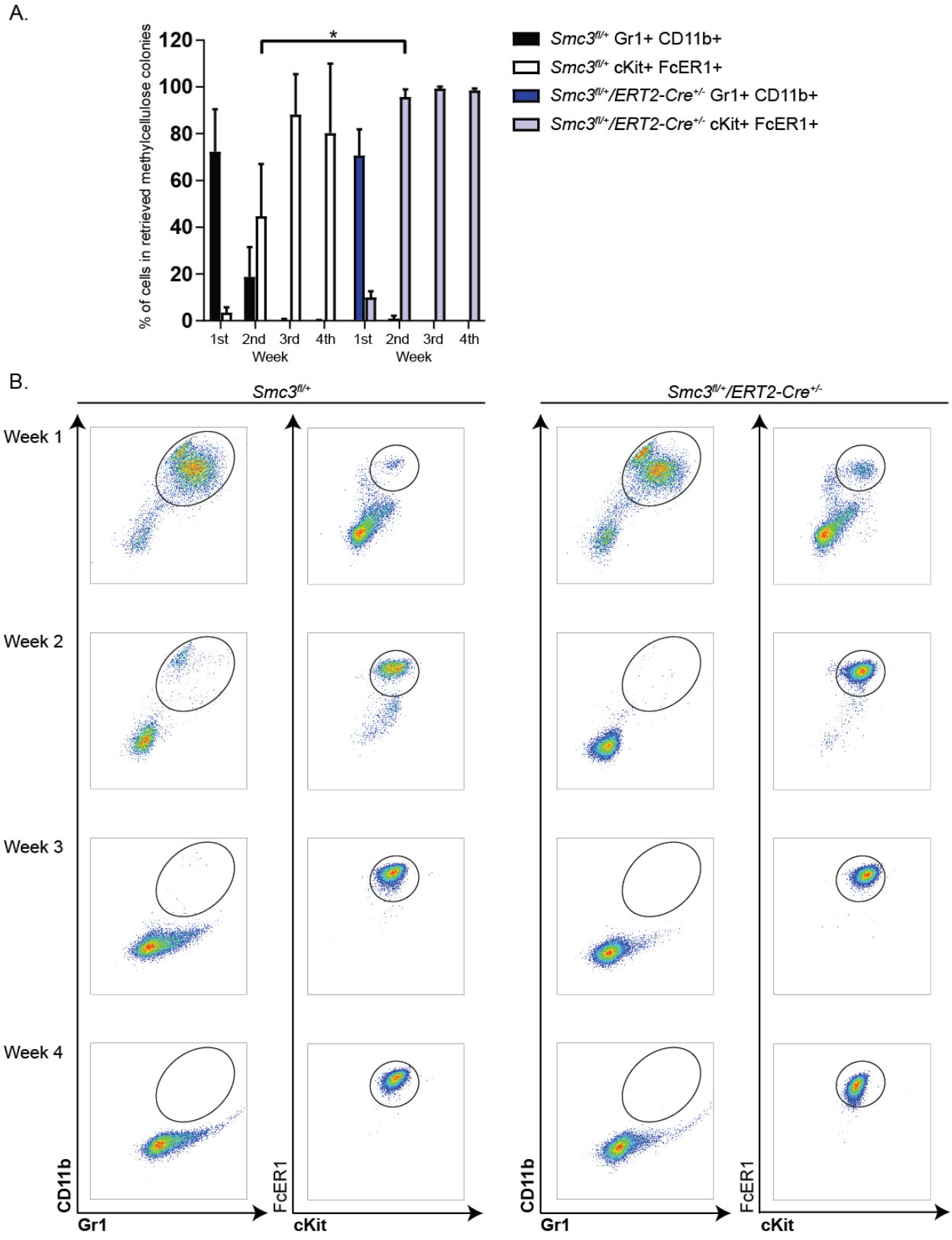
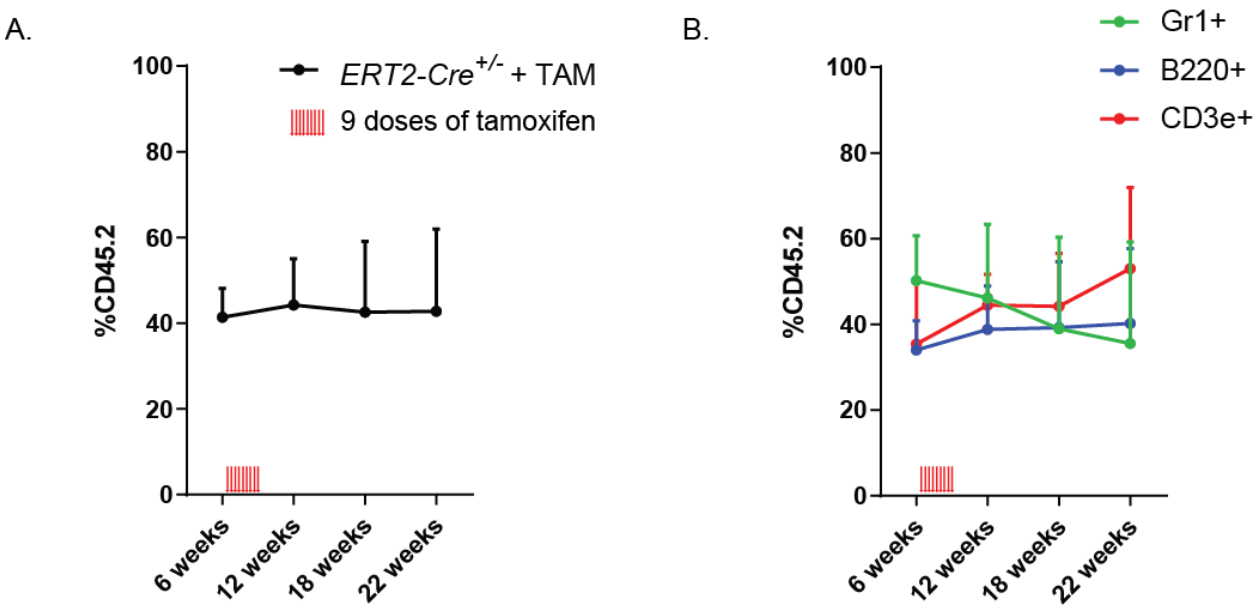


Figure 2.12. Competitive transplantation of *ERT2-Cre*$^{+/-}$ bone marrow cells.



References

1. Saultz JN, Garzon R. Acute Myeloid Leukemia: A Concise Review. *J Clin Med*. 2016;5(3). doi:10.3390/jcm5030033

2. Lagunas-Rangel FA, Chávez-Valencia V, Gómez-Guijosa MÁ, Cortes-Penagos C. Acute Myeloid Leukemia—Genetic Alterations and Their Clinical Prognosis. *Int J Hematol-Oncol Stem Cell Res*. 2017;11(4):328-339.

3. Cancer Genome Atlas Research Network, Ley TJ, Miller C, et al. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N Engl J Med*. 2013;368(22):2059-2074. doi:10.1056/NEJMoa1301689

4. Kon A, Shih L-Y, Minamino M, et al. Recurrent mutations in multiple components of the cohesin complex in myeloid neoplasms. *Nat Genet*. 2013;45(10):1232-1237. doi:10.1038/ng.2731

5. Thota S, Viny AD, Makishima H, et al. Genetic alterations of the cohesin complex genes in myeloid malignancies. *Blood*. 2014;124(11):1790-1798. doi:10.1182/blood-2014-04-567057

6. Yoshida K, Toki T, Okuno Y, et al. The landscape of somatic mutations in Down syndrome–related myeloid disorders. *Nat Genet*. 2013;45(11):1293-1299. doi:10.1038/ng.2759

7. Barber TD, McManus K, Yuen KWY, et al. Chromatid cohesion defects may underlie chromosome instability in human colorectal cancers. *Proc Natl Acad Sci*. 2008;105(9):3443-3448. doi:10.1073/pnas.0712384105

8. Gorringe KL, Ramakrishna M, Williams LH, et al. Are there any more ovarian tumor suppressor genes? A new perspective using ultra high-resolution copy number and loss of heterozygosity analysis. *Genes Chromosomes Cancer*. 2009;48(10):931-942. doi:10.1002/gcc.20694

9. Bailey ML, O'Neil NJ, Pel DM van, Solomon DA, Waldman T, Hieter P. Glioblastoma Cells Containing Mutations in the Cohesin Component STAG2 Are Sensitive to PARP Inhibition. *Mol Cancer Ther*. 2014;13(3):724-732. doi:10.1158/1535-7163.MCT-13-0749

10. Solomon DA, Kim J-S, Bondaruk J, et al. Frequent truncating mutations of *STAG2* in bladder cancer. *Nat Genet*. 2013;45(12):1428-1430. doi:10.1038/ng.2800

11. Balbás-Martínez C, Sagrera A, Carrillo-de-Santa-Pau E, et al. Recurrent inactivation of *STAG2* in bladder cancer is not associated with aneuploidy. *Nat Genet*. 2013;45(12):1464-1469. doi:10.1038/ng.2799

12. Solomon DA, Kim T, Diaz-Martinez LA, et al. Mutational Inactivation of STAG2 Causes Aneuploidy in Human Cancer. *Science*. 2011;333(6045):1039-1043. doi:10.1126/science.1203619

13. Remeseiro S, Cuadrado A, Gómez-López G, Pisano DG, Losada A. A unique role of cohesin-SA1 in gene regulation and development. *EMBO J*. 2012;31(9):2090-2102. doi:10.1038/emboj.2012.60

14. Mannini L, Cucco F, Quarantotti V, Krantz ID, Musio A. Mutation spectrum and genotype-phenotype correlation in Cornelia de Lange syndrome. *Hum Mutat*. 2013;34(12):1589-1596. doi:10.1002/humu.22430

15. Welch JS, Ley TJ, Link DC, et al. The Origin and Evolution of Mutations in Acute Myeloid Leukemia. *Cell*. 2012;150(2):264-278. doi:10.1016/j.cell.2012.06.023

16. Thol F, Bollin R, Gehlhaar M, et al. Mutations in the cohesin complex in acute myeloid leukemia: clinical and prognostic implications. *Blood*. 2014;123(6):914-920. doi:10.1182/blood-2013-07-518746

17. Conese M, Liso A. Cohesin complex is a major player on the stage of leukemogenesis. *Stem Cell Investig*. 2016;3:18. doi:10.21037/sci.2016.05.04

18. Xie M, Lu C, Wang J, et al. Age-related mutations associated with clonal hematopoietic expansion and malignancies. *Nat Med*. 2014;20(12):1472-1478. doi:10.1038/nm.3733

19. Jaiswal S, Natarajan P, Silver AJ, et al. Clonal Hematopoiesis and Risk of Atherosclerotic Cardiovascular Disease. *N Engl J Med*. 2017;377(2):111-121. doi:10.1056/NEJMoa1701719

20. Cole CB, Russler-Germain DA, Ketkar S, et al. Haploinsufficiency for DNA methyltransferase 3A predisposes hematopoietic cells to myeloid malignancies. *J Clin Invest*. 2017;127(10):3657-3674. doi:10.1172/JCI93041

21. Ball AR, Chen Y-Y, Yokomori K. Mechanisms of cohesin-mediated gene regulation and lessons learned from cohesinopathies. *Biochim Biophys Acta BBA - Gene Regul Mech*. 2014;1839(3):191-202. doi:10.1016/j.bbagrm.2013.11.002

22. Wendt KS, Yoshida K, Itoh T, et al. Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature*. 2008;451(7180):796-801. doi:10.1038/nature06634

23. Merkenschlager M, Odom DT. CTCF and cohesin: linking gene regulatory elements with their targets. *Cell*. 2013;152(6):1285-1297. doi:10.1016/j.cell.2013.02.029

24. Schmidt D, Schwalie PC, Ross-Innes CS, et al. A CTCF-independent role for cohesin in tissue-specific transcription. *Genome Res*. 2010;20(5):578-588. doi:10.1101/gr.100479.109

25. Genomic and Epigenomic Landscapes of Adult De Novo Acute Myeloid Leukemia. *N Engl J Med*. 2013;368(22):2059-2074. doi:10.1056/NEJMoa1301689

26. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A*. 2001;98(9):5116-5121. doi:10.1073/pnas.091062498

27. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res*. 2016;44(D1):D457-D462. doi:10.1093/nar/gkv1070

28. Walter MJ, Shen D, Ding L, et al. Clonal Architecture of Secondary Acute Myeloid Leukemia. *N Engl J Med*. 2012;366(12):1090-1098. doi:10.1056/NEJMoa1106968

29. Genovese G, Kähler AK, Handsaker RE, et al. Clonal Hematopoiesis and Blood-Cancer Risk Inferred from Blood DNA Sequence. *N Engl J Med*. 2014;371(26):2477-2487. doi:10.1056/NEJMoa1409405

30. Viny AD, Ott CJ, Spitzer B, et al. Dose-dependent role of the cohesin complex in normal and malignant hematopoiesis. *J Exp Med*. 2015;212(11):1819-1832. doi:10.1084/jem.20151317

31. Xu H, Balakrishnan K, Malaterre J, et al. Rad21-cohesin haploinsufficiency impedes DNA repair and enhances gastrointestinal radiosensitivity in mice. *PloS One*. 2010;5(8):e12112. doi:10.1371/journal.pone.0012112

32. Welch JS, Yuan W, Ley TJ. PML-RARA can increase hematopoietic self-renewal without causing a myeloproliferative disease in mice. *J Clin Invest*. 2011;121(4):1636-1645. doi:10.1172/JCI42953

33. Mullenders J, Aranda-Orgilles B, Lhoumaud P, et al. Cohesin loss alters adult hematopoietic stem cell homeostasis, leading to myeloproliferative neoplasms. *J Exp Med*. 2015;212(11):1833-1850. doi:10.1084/jem.20151323

34. Mazumdar C, Shen Y, Xavy S, et al. Leukemia-Associated Cohesin Mutants Dominantly Enforce Stem Cell Programs and Impair Human Hematopoietic Progenitor Differentiation. *Cell Stem Cell*. 2015;17(6):675-688. doi:10.1016/j.stem.2015.09.017

35. Shirai CL, Ley JN, White BS, et al. Mutant U2AF1 Expression Alters Hematopoiesis and Pre-mRNA Splicing In Vivo. *Cancer Cell*. 2015;27(5):631-643. doi:10.1016/j.ccell.2015.04.008

36. Fei DL, Zhen T, Durham B, et al. Impaired hematopoiesis and leukemia development in mice with a conditional knock-in allele of a mutant splicing factor gene U2af1. *Proc Natl Acad Sci U S A*. 2018;115(44):E10437-E10446. doi:10.1073/pnas.1812669115

37. Kim E, Ilagan JO, Liang Y, et al. SRSF2 Mutations Contribute to Myelodysplasia by Mutant-Specific Effects on Exon Recognition. *Cancer Cell*. 2015;27(5):617-630. doi:10.1016/j.ccell.2015.04.006

38. Komeno Y, Huang Y-J, Qiu J, et al. SRSF2 Is Essential for Hematopoiesis, and Its Myelodysplastic Syndrome-Related Mutations Dysregulate Alternative Pre-mRNA Splicing. *Mol Cell Biol*. 2015;35(17):3071-3082. doi:10.1128/MCB.00202-15

39. Wang C, Sashida G, Saraya A, et al. Depletion of Sf3b1 impairs proliferative capacity of hematopoietic stem cells but is not sufficient to induce myelodysplasia. *Blood*. 2014;123(21):3336-3343. doi:10.1182/blood-2013-12-544544

40. Matsunawa M, Yamamoto R, Sanada M, et al. Haploinsufficiency of Sf3b1 leads to compromised stem cell function but not to myelodysplasia. *Leukemia*. 2014;28(9):1844-1850. doi:10.1038/leu.2014.73

41. Visconte V, Tabarroki A, Zhang L, et al. Splicing factor 3b subunit 1 (Sf3b1) haploinsufficient mice display features of low risk Myelodysplastic syndromes with ring sideroblasts. *J Hematol OncolJ Hematol Oncol*. 2014;7:89. doi:10.1186/s13045-014-0089-x

42. Obeng EA, Chappell RJ, Seiler M, et al. Physiologic Expression of Sf3b1(K700E) Causes Impaired Erythropoiesis, Aberrant Splicing, and Sensitivity to Therapeutic Spliceosome Modulation. *Cancer Cell*. 2016;30(3):404-417. doi:10.1016/j.ccell.2016.08.006

43. Mupo A, Seiler M, Sathiaseelan V, et al. Hemopoietic-specific Sf3b1-K700E knock-in mice display the splicing defect seen in human MDS but develop anemia without ring sideroblasts. *Leukemia*. 2017;31(3):720-727. doi:10.1038/leu.2016.251

44. Abdel-Wahab O, Gao J, Adli M, et al. Deletion of Asxl1 results in myelodysplasia and severe developmental defects in vivo. *J Exp Med*. 2013;210(12):2641-2659. doi:10.1084/jem.20131141

45. Nagase R, Inoue D, Pastore A, et al. Expression of mutant Asxl1 perturbs hematopoiesis and promotes susceptibility to leukemic transformation. *J Exp Med*. 2018;215(6):1729-1747. doi:10.1084/jem.20171151

46. ExAC Browser. http://exac.broadinstitute.org/gene/ENSG00000108055. Accessed May 18, 2018.

47. Cucco F, Musio A. Genome stability: What we have learned from cohesinopathies. *Am J Med Genet C Semin Med Genet*. 2016;172(2):171-178. doi:10.1002/ajmg.c.31492

48. Ramos FJ, Puisac B, Baquero-Montoya C, et al. Clinical utility gene card for: Cornelia de Lange syndrome. *Eur J Hum Genet*. 2015;23(10):1431. doi:10.1038/ejhg.2014.270

49. Kumar R, Corbett MA, Van Bon BWM, et al. Increased STAG2 dosage defines a novel cohesinopathy with intellectual disability and behavioral problems. *Hum Mol Genet*. 2015;24(25):7171-7181. doi:10.1093/hmg/ddv414

50. Leroy C, Jacquemont M-L, Doray B, et al. Xq25 duplication: the crucial role of the STAG2 gene in this novel human cohesinopathy. *Clin Genet*. 2016;89(1):68-73. doi:10.1111/cge.12567

51. Baquero-Montoya C, Gil-Rodríguez MC, Teresa-Rodrigo ME, et al. Could a patient with SMC1A duplication be classified as a human cohesinopathy? *Clin Genet*. 2014;85(5):446-451. doi:10.1111/cge.12194

52. Trapnell C, Roberts A, Goff L, et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc*. 2012;7(3):562-578. doi:10.1038/nprot.2012.016

53. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods*. 2013;10(12):1213-1218. doi:10.1038/nmeth.2688

54. Zhang Y, Liu T, Meyer CA, et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol*. 2008;9(9):R137. doi:10.1186/gb-2008-9-9-r137

55. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinforma Oxf Engl*. 2010;26(6):841-842. doi:10.1093/bioinformatics/btq033

56. Ramírez F, Dündar F, Diehl S, Grüning BA, Manke T. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res*. 2014;42(Web Server issue):W187-191. doi:10.1093/nar/gku365

57. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550. doi:10.1186/s13059-014-0550-8

Chapter 3:
Exome analysis of treatment-related AML after APL suggests secondary evolution

Treatment-related acute myeloid leukemia (tAML) and treatment-related myelodysplastic syndrome (tMDS) have been associated with many types of chemotherapy and radiation.[1] Treatment-related AML or tMDS have been observed after treatment of acute promyelocytic leukemia (APL) with combination all-trans retinoic acid (ATRA)/idarubicin/cytarabine.[2, 3, 4, 5, 6] It is unknown whether tAML will emerge following exposure to all-trans retinoic acid (ATRA) and arsenic trioxide (ATO), which should not cause DNA damage or the clonal selection of chemotherapy-resistant clones that give rise to tAML.[1, 6, 7] We describe a case of an older women who presented initially with dysplasia. Shortly thereafter she developed APL and was treated with ATRA/ATO. Five years later she developed what clinically appeared to be tAML. Exome sequencing revealed a founding clone with a *TET2* mutation and shared passenger mutations that existed at all three time points. The APL sample shared these mutations, and the subsequent "tAML" emerged as a new subclone with an *NPM1* mutation, more consistent of secondary AML rather than tAML. This case demonstrates that APL may emerge within the context of clonal hematopoiesis, and that tAML emerging after ATRA/ATO should be evaluated for features consistent with secondary AML.

An 81 year-old woman presented initially with cytopenias at an outside facility. A bone marrow biopsy was performed, revealing trilineage dysplasia, 6% promyelocytes, and a normal karyotype, 46 XX[20/20] (Table 3.1). Two months later, the cytopenias persisted and a repeat bone marrow aspirate was hypercellular with 50% promyelocytes. Cytogenetics revealed 46 XX, t(15;17)[18]/46 XX[2] and 89% of cells were positive for *PML-RARA* by FISH. She was referred to Washington University. The bone marrow biopsy was repeated and banked with appropriate consent for genomic analysis. The repeat biopsy was unfortunately hemodilute with 4.5% t(15;17) by FISH [9/200], was not evaluated by cytomorphology, and RT-PCR failed (control

GAPDH primers did not amplify). She was treated with ATRA/ATO. Bone marrow biopsy 6 weeks later revealed 1% promyelocytes and normal karyotype by cytogenetics and FISH. Due to her age, no further bone marrow biopsies were performed and subsequent peripheral blood RT-PCR tests were negative. Five years later she became increasingly cytopenic and a bone marrow aspirate revealed acute myeloid leukemia with 72% blasts, 20% promyelocytes, and a normal karyotype. She was treated with decitabine. She deteriorated during the first cycle and transitioned to hospice.

Exome sequencing was performed on three samples using techniques described elsewhere and compared against a skin sample, which was used as a germline control (Figure 3.1).[8] Cells retained on a coverslip were available from the dysplastic pre-APL sample and this was used for genomic analysis. No coverslip or other material was available from the APL sample with 89% *PML-RARA*. Cryopreserved bone marrow aspirate cells were used for APL (4.5% *PML-RARA*) and AML (72% blasts) analysis. The collected APL sample only generated one cryovial, which was used for DNA synthesis. Unfortunately, no additional samples are available for RNA-Seq analysis or sequencing of subpopulations after flow sorting.

A shared founding clone existed across all three time points that contained a *TET2* mutation and 15 additional variants (Figure 3.1, black). Three additional clusters were identified: 1) variants that were present predominantly in the dysplasia sample (*SEMA4A*, and *ZBTB7A*, orange); 2) variants that were absent in the AML sample (*TET1*, *SACM1L*, *OR7D2*, *SH3TC1*, blue); 3) variants that increased in the AML sample, but were present at low variant allele frequencies or were undetected in prior samples (*NPM1* and 13 additional variants, red).

Treatment-related AML has been associated with APL therapy that includes alkylating agents,[2, 3, 4, 9] but has not been associated with ATRA/ATO, which are non-cytotoxic and do not damage

DNA. This unusual case initially appeared consistent with tAML following ATRA/ATO, however, following exome analysis, appears to be more consistent with a pre-existing dominant clone associated with dysplasia and a *TET2* mutation, and an *NPM1*-associated secondary AML. Two models are possible (Figure 3.1F). First, the APL may have emerged as a subclone of the *TET2* founding clone. Given the high *TET2* variant allele frequency (VAF, consistent with 70% and 66% tumor burden in the dysplastic and APL samples, respectively) and the high tumor burden of *PML-RARA* in the bone marrow (85%), this model seems likely. Alternatively, because the sequenced APL sample is hemodilute, it is possible that the *PML-RARA* clone is independent of the *TET2* clone, that the APL clone did not peripheralize, or that geographical heterogeneity existed in this older patient. Given the absence of additional samples, it is impossible to determine whether the APL emerged as a subclone of the *TET2* founding clone occurred independent of the *TET2* clone.

In order to determine if these subclonal patterns are observed in other APL patients, we reviewed available cytogenetic results from published APL patients. Two cases of disease progression from MDS to APL have been described[10], and cases of treatment-related APL have been described following MDS.[11] In the former cases, like this case, the timeline of evolution and dysplastic changes in the APL morphology suggests that the APL clone likely was evolutionarily related to the MDS clone. In the latter cases, the cytogenetics and timeline suggest the APL clone likely emerged independent of the MDS clone.

Clonal hematopoiesis of indeterminate potential (CHIP) and MDS have been associated with deletions involving chromosomes 5 and 7[12, 13], which occasionally co-occur with t(15;17).[14, 15, 16] Therefore, we examined the subclonal relationship of t(15;17) with possible CHIP-related variants among APL patients enrolled at large cancer centers. We identified 44 APL cases with

t(15;17) and cytogenetic abnormalities in chromosomes 5 or 7. Five cases involved chromosome 5, two of which also had abnormalities in chromosome 7. Thirty-three cases involved a monoclonal process, suggesting that the variants in chromosomes 5 or 7 co-occurred with t(15;17). Of the eleven cases with identified polyclones, the chromosome 5 or 7 variant was in a subclone of t(15;17) in ten cases and in only one case was there evidence of del(7) in a clone that was independent of the t(15;17) clone (Patient 52009 from APL 2006, 46, XY, t(15,17)(q22, q11)[13]/46, XY, del(7)(q35)[5]/46, XY [2]) (Table 3.2). These data suggest that t(15;17) may co-occur with additional cytogenetic deletions of chromosomes 5 and 7, but t(15;17) is almost invariably the founding event, and it is unusual for a concurrent CHIP or MDS clone to co-exist with t(15;17).

Other investigators have reviewed the outcomes in APL patients with cytogenetic abnormalities that occur in addition to t(15;17).[16, 17, 18] The most recurrently observed co-occurring cytogenetic abnormalities include +8 and +21, both of which frequently can be identified as subclonal progression events when analyzing metaphase cytogenetics or FISH, and have been observed as progression events in MDS and AML.[11] In these studies, additional cytogenetic abnormalities have not correlated with initial clinical characteristics, or with outcomes in patients with t(15;17).[16, 17, 18]

Morphologic disease switching also has been described in *NPM1*-mutated AML patients who subsequently developed MDS or myelofibrosis. In each of these cases, the *NPM1*-mutation was lost, and the MDS or myelofibrosis evolved from an antecedent clone with a mutation in *TET2*, *JAK2*, *ASXL1*, *IDH2*, or a spliceosome trancript, suggesting that the two clonal diseases were related through an ancestral clone.[19]

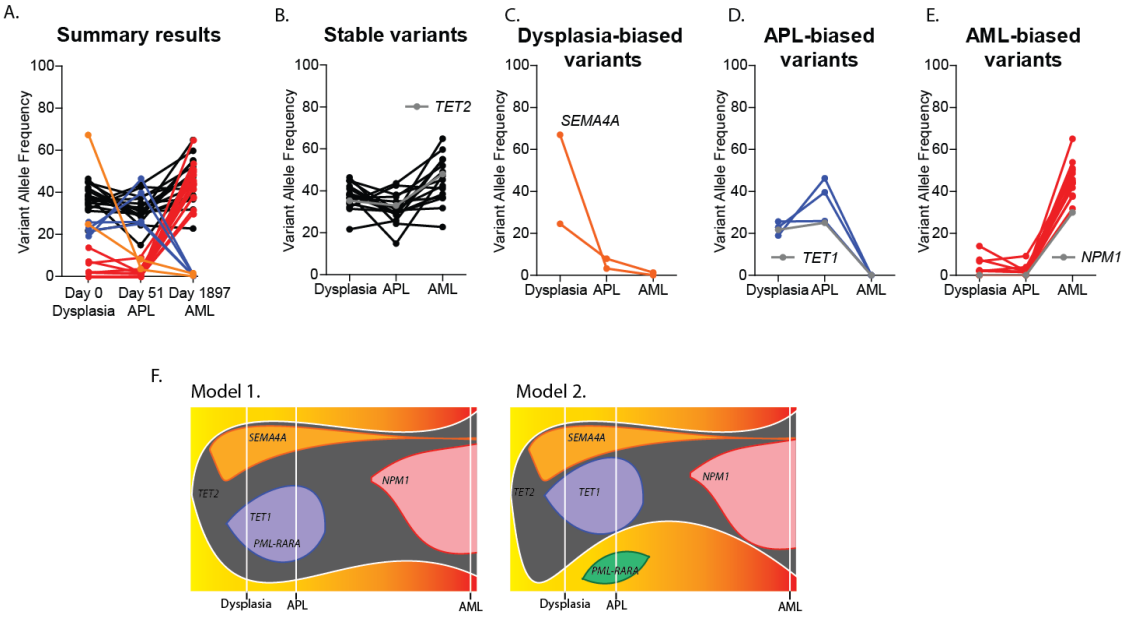
In summary, we present a case of what initially appeared to be tAML following non-cytotoxic ATRA/ATO therapy for APL. Exome analysis clearly demonstrated a *TET2*-associated, dominant clonal process that anteceded the APL diagnosis, persisted, and gave rise to AML associated with evolutionary expansion of an *NPM1*-mutated subclone. This progression would be more consistent with a secondary AML process, rather than a treatment-related process. Additional characterization of this case would be interesting. Unfortunately, samples for such analyses are unavailable. Review of additional cytogenetic abnormalities observed in APL patients did identify one rare case with del(7) independent of the APL clone, suggesting that APL can co-exist with CHIP or MDS clones, although it appears much more common for such cytogenetic abnormalities to occur as subclones of the APL clone. Collectively, these results suggest that caution must be exercised when interpreting the development of tAML following ATRA/ATO therapy, and subclonal expansion of related or well-established clones should be considered, especially in older patients.

Figure Legends

Figure 3.1. Exome analysis of patient 10DD-1029.

(A) Summary results of all somatic variants detected at any of the three time points analyzed. (B) Cluster of variants with stable VAFs across all three samples. (C) Cluster of variants present in the initial dysplastic sample, with reduced VAFs in subsequent samples. (D) Cluster of variants with VAFs in the initial two samples, but which were absent in the subsequent AML sample. (E) Cluster of variants associated with the AML progression, which were largely absent in the initial two samples. (F) Two models for subclonal expansion.

Figure 3.1. Exome analysis of patient 10DD-1029.



Tables 3.1. Clinical data of the patient.

	Dysplasia	APL	APL (banked)	post-APL	AML
	Day 0	Day 49	Day 51	Day 87	Day 1897
Sequencing	exome	ND	Exome	ND	exome
t(15;17)	0%	89%	4.5%	0%	0%
BM blasts	0%	1%	ND	1%	72%
BM promyelocytes	6%	50%	ND	1%	20%
BM myelocytes	9%	25%	ND	25%	1%
BM metamyelocytes	2%	3%	ND	3%	5%
BM bands	9%	2%	ND	2%	0%
BM dysplasia	trilineage	atypia	ND	none	none
PB WBC	NA	0.9	0.9	0.9	2.2
PB % Lymphs	NA	95%	95%	95%	88%
PB % Blasts	NA	0%	0%	0%	6%

AML, acute myeloid leukemia; APL, acute promyelocytic leukemia; BM, bone marrow; NA, not available; ND, not done; PB, peripheral blood; WBC, white blood cells.

Table 3.2. APL patients with cytogenetic abnormalities in chromosomes 5 or 7 and evaluable subclonal architecture.

Source	Patient	Karyotype	t(15;17) subclone	Clone independent of t(15;17)
APL 2006	52009	46, XY, t(15,17) (q22, q11) [13] / 46, XY, del(7)(q35)[5]/46, XY [2]		del(7)
MD Anderson	NA	46,XX,add(7)(q32),t(15;17)(q24;q21)[5] /47,idem,+8[2]/46,XX,del(7)(q22q34),t(15;17)[7]/ 46,XX,add(4)(p16),t(15;17)[5]/46, XX[1]	del(7), +8, add(4)	
APL 2006	78034	t(15 ;17), +8 [?], t(15 ;17), -7, -5 [?]	-7, -5, +8	
APL92	848	46XY,15q+,17q- [19/20] / 45,XY,-5,15q+,17q- [1/20]	-5	
APL92	108	46,XX,1p-,7q-,15q+,17q-,19q+ [14/20] / 46,XX,15q+,17q- [6/20]	del(1p), del(7q), add(19q)	
APL92	687	46,XX,t(15;17)(q22;q11 ~ 21) [19/20] / 46,idem,-7,+mar [1/20]	-7	
APL92	742	46,XX,t(15;17) [16/17] / 46,idem,add(7)(q?) [1/17]	add(7)	
MD Anderson	NA	46,XX,t(15;17)(q24;q21)[15]/ 46,idem,del(7)(q22q32)[3]; 46,XX[2]	del(7)	
MD Anderson	NA	46XX,t(15;17)(q22;q21.1)[16]/46, idem,del(7)(q22q34)[1]/ 46XX[3]	del(7)	
MD Anderson	NA	46XY,t(15;17)(q22;q21)[16]/46,idem,del(7)(q32q36)[1]/46,XY[3];	del(7)	
MD Anderson	NA	46XX, del(7)(q31q36), inv(9)(p11;q12),t(15;17)(q22;q21)[2]/47 XX,+8,inv(9), der(15), der(17)[3]/46XX, inv(9).	del(7), +8	

References

1. Wong TN, Ramsingh G, Young AL, et al. Role of TP53 mutations in the origin and evolution of therapy-related acute myeloid leukaemia. *Nature*. 2015;518(7540):552-555. doi:10.1038/nature13968

2. Park SH, Chi H-S, Cho Y-U, Jang S, Park C-J, Lee J-H. A case of therapy-related acute myeloid leukemia with a normal karyotype after sustained molecular complete remission of acute promyelocytic leukemia. *Ann Lab Med*. 2014;34(1):68-70. doi:10.3343/alm.2014.34.1.68

3. Zompi S, Viguié F. Therapy-related acute myeloid leukemia and myelodysplasia after successful treatment of acute promyelocytic leukemia. *Leuk Lymphoma*. 2002;43(2):275-280. doi:10.1080/10428190290006044

4. Ojeda-Uribe M, Schneider A, Luquet I, et al. Therapy-related acute myeloid leukemia (t-AML) with poor-risk cytogenetics in two patients with persistent molecular complete remission of acute promyelocytic leukemia. *Eur J Haematol*. 2012;89(3):267-272. doi:10.1111/j.1600-0609.2012.01805.x

5. Park TS, Choi JR, Yoon SH, et al. Acute promyelocytic leukemia relapsing as secondary acute myelogenous leukemia with translocation t(3;21)(q26;q22) and RUNX1-MDS1-EVI1 fusion transcript. *Cancer Genet Cytogenet*. 2008;187(2):61-73. doi:10.1016/j.cancergencyto.2008.06.015

6. de Thé H. Differentiation therapy revisited. *Nat Rev Cancer*. 2018;18(2):117-127. doi:10.1038/nrc.2017.103

7. Lehmann-Che J, Bally C, Letouzé E, et al. Dual origin of relapses in retinoic-acid resistant acute promyelocytic leukemia. *Nat Commun*. 2018;9(1):2047. doi:10.1038/s41467-018-04384-5

8. Welch JS, Petti AA, Miller CA, et al. TP53 and Decitabine in Acute Myeloid Leukemia and Myelodysplastic Syndromes. *N Engl J Med*. 2016;375(21):2023-2036. doi:10.1056/NEJMoa1605949

9. Zompi S, Legrand O, Bouscary D, et al. Therapy-related acute myeloid leukaemia after successful therapy for acute promyelocytic leukaemia with t(15;17): a report of two cases and a review of the literature. *Br J Haematol*. 2000;110(3):610-613.

10. Isono S, Saigo K, Nagata K, et al. A case of acute promyelocytic leukemia with morphologic multilineage dysplastic changes. *Hematol Rep*. 2013;5(1):18-20. doi:10.4081/hr.2013.e5

11. Andersen MK, Larson RA, Mauritzson N, Schnittger S, Jhanwar SC, Pedersen-Bjergaard J. Balanced chromosome abnormalities inv(16) and t(15;17) in therapy-related

myelodysplastic syndromes and acute leukemia: report from an international workshop. *Genes Chromosomes Cancer*. 2002;33(4):395-400.

12. Laurie CC, Laurie CA, Rice K, et al. Detectable clonal mosaicism from birth to old age and its relationship to cancer. *Nat Genet*. 2012;44(6):642-650. doi:10.1038/ng.2271

13. Zahid MF, Malik UA, Sohail M, Hassan IN, Ali S, Shaukat MHS. Cytogenetic Abnormalities in Myelodysplastic Syndromes: An Overview. *Int J Hematol-Oncol Stem Cell Res*. 2017;11(3):231-239.

14. Cervera J, Montesinos P, Hernández-Rivas JM, et al. Additional chromosome abnormalities in patients with acute promyelocytic leukemia treated with all-trans retinoic acid and chemotherapy. *Haematologica*. 2010;95(3):424-431. doi:10.3324/haematol.2009.013243

15. Ono T, Takeshita A, Iwanaga M, et al. Impact of additional chromosomal abnormalities in patients with acute promyelocytic leukemia: 10-year results of the Japan Adult Leukemia Study Group APL97 study. *Haematologica*. 2011;96(1):174-176. doi:10.3324/haematol.2010.030205

16. De Botton S, Chevret S, Sanz M, et al. Additional chromosomal abnormalities in patients with acute promyelocytic leukaemia (APL) do not confer poor prognosis: results of APL 93 trial. *Br J Haematol*. 2000;111(3):801-806.

17. Hernández JM, Martín G, Gutiérrez NC, et al. Additional cytogenetic changes do not influence the outcome of patients with newly diagnosed acute promyelocytic leukemia treated with an ATRA plus anthracyclin based protocol. A report of the Spanish group PETHEMA. *Haematologica*. 2001;86(8):807-813.

18. Lou Y, Suo S, Tong H, et al. Characteristics and prognosis analysis of additional chromosome abnormalities in newly diagnosed acute promyelocytic leukemia treated with arsenic trioxide as the front-line therapy. *Leuk Res*. 2013;37(11):1451-1456. doi:10.1016/j.leukres.2013.07.030

19. Morin-Zorman S, Renneville A, Adès L. Assessment of Minimal Residual Disease in Standard-Risk AML. *N Engl J Med*. 2016;375(6):e9. doi:10.1056/NEJMc1603847