

**DNA POLYMERASE THETA COOPERATES WITH ADH5 AND ALDH2
DEHYDROGENASES TO PROTECT HEMATOPOIETIC CELLS
FROM FORMALDEHYDE-MEDIATED DNA DAMAGE**

A Dissertation
Submitted to
the Temple University Graduate Board

In Partial Fulfillment
of the Requirements for the Degree
DOCTOR OF PHILOSOPHY

by
Jessica Atkins
August 2025

Examining Committee Members:

Tomasz Skorski, Advisory Chair, Cancer and Cellular Biology
Raza Zaidi, Cancer and Cellular Biology
Nathaniel Snyder, Cardiovascular Sciences
Neil Johnson, Cancer and Cellular Biology
Beata Kosmider, External Reader, Microbiology, Immunology and Inflammation

©
Copyright
2025

by

Jessica Atkins
All Rights Reserved

ABSTRACT

Hematopoietic stem and progenitor cells (HSPCs) are exposed to physiological levels of formaldehyde but can occasionally be challenged by high levels of formaldehyde generated by various endogenous and exogenous sources. In addition, leukemia cells stressed by oncogenic mutations continuously produce excessive amounts of formaldehyde. Here, we show that DNA polymerase theta (Pol θ) cooperates with alcohol dehydrogenase 5 (ADH5) and aldehyde dehydrogenase 2 (ALDH2) to protect healthy and malignant HSPCs challenged by formaldehyde. ADH5 and ALDH2 metabolize formaldehyde while Pol θ -mediated DNA repair by microhomology-dependent end-joining (TMEJ) protects cells from the lethal effect of DNA double strand breaks (DSBs) resulting from formaldehyde-mediated DNA-protein crosslinks (DPCs). Genetic or pharmacological targeting of ADH5 or ALDH2 enhanced the effect of Pol θ inhibitors in leukemia cells. Thus, ADH5 and ALDH2 cooperate with Pol θ to protect normal and malignant HSPCs challenged by high levels of formaldehyde and inhibition of Pol θ and ADH5 or ALDH2 may exert an anti-leukemic effect.

To my beloved grandparents,

Carmen and Geraldine Purcaro

You are two of the greatest people I've ever known. Your love, wisdom and unwavering support have shaped my life in ways words cannot fully express.

I couldn't have done this without you.

Love always,

Jessica

ACKNOWLEDGEMENTS

First, I would like to thank my thesis advisor, Tomasz Skorski, for giving me the opportunity to join the lab and for seeing my potential. It has been a privilege to learn from someone so dedicated to their work. I want to thank Anna-Mariya, who took me under her wing from the moment I started in the lab. From troubleshooting experiments to talking through new ideas, she has been an invaluable source of both intellectual and personal support. Her curiosity, integrity, and passion for helping others is what inspires me every day. I'd also like to thank Monika for believing in me and lifting my spirits, especially on my lowest days. Our daily lunches and frequent trips to our favorite place, Levain Bakery are some of the highlights of my PhD. Both Anna-Mariya and Monika have been wonderful coworkers, mentors and friends who made the long hours and many challenges more manageable and rewarding. I also want to thank the rest of the Skorski lab, both past and present members, for their kindness and encouragement throughout my PhD journey. I'd also like to thank my colleagues and collaborators for their time and effort and whose scientific work contributed to my first-author publication.

I would like to thank my family and friends for supporting me every step of the way. I want to give a very special thanks to my three sisters, Ashley, Nicole and Jaclyn, for never letting me give up. You believed in me even when I didn't believe in myself and that made all the difference. So much of who I am is because of you, so this accomplishment is just as much yours as it is mine. I'd like to thank my mom, whose resilience inspires me daily and is something I carry with me in everything I do. I would like to thank my two bonus brothers, Cole and Jessup, for always making me laugh and giving the best advice. I'd like to thank my friends for being such a great support system and giving me the confidence to accomplish my goals.

Finally, I would like to thank my grandparents, both of whom have had such a profound impact on my life. My heroes, my inspiration, my everything. Over the years, my grandpa has listened to me excitedly talk about my research, shared in my failures and frustrations, and offered words of encouragement. My grandma is the person who inspired me to become a scientist. When she passed away from a highly aggressive brain cancer, it completely changed the trajectory of my life. My grief and frustration, coupled with my innate curiosity led me to science and where I am today. I'd like to think she's been beside me throughout this entire journey and would be proud of how far I have come.

TABLE OF CONTENTS

	Page
ABSTRACT	iii
DEDICATION	iv
ACKNOWLEDGMENTS	v
LIST OF FIGURES	xi
CHAPTER	
1. INTRODUCTION	1
The ubiquitous nature of formaldehyde	1
Exogenous sources of formaldehyde.....	1
Endogenous sources of formaldehyde	4
Intracellular formaldehyde metabolism.....	7
Formaldehyde causes genome instability	9
ICL-mediated repair pathway.....	10
Bone marrow failure syndromes	11
DPC-mediated repair pathways.....	15
DSB repair pathways.....	16
<i>POLQ</i> as a therapeutic target for cancer.....	21
Interplay between metabolic enzymes and DNA repair pathways	24
2. METHODS.....	28
Mouse husbandry	28
Mouse breeding	28
Mouse genotyping.....	29
Polq-specific PCR reaction.....	30
Adh5-specific PCR reaction.....	30

Aldh2-specific PCR reaction	31
Mouse organ extraction for histopathology.....	32
Mouse bone marrow cell (mBMC) collection	32
Hematological analysis of mouse peripheral blood.....	33
AML and MPN patient cells	34
Cell lines	34
Mycoplasma detection	35
Mycoplasma-specific PCR reaction	35
Flow cytometry.....	36
Immunophenotyping: mouse peripheral blood	36
Immunophenotyping: mBMCs	37
γ H2AX and 8-Oxo-G analysis.....	37
Apoptosis	38
Human CD19 detection	39
Western blots.....	39
γ H2AX immunofluorescence (IF)	40
Detection of DPCs	41
Buffer preparation.....	41
Part I: isolation of DPC-associated DNA and free DNA	42
Part II: quantification of DNA by pico green dsDNA assay	44
Detection of DSBs.....	44
Reagent preparation.....	44
Preparation of samples and slides.....	45
Assay protocol.....	45
Clonogenic assay.....	46
Formaldehyde detection.....	47

<i>In vitro</i> formaldehyde challenge	47
<i>In vivo</i> methanol challenge.....	48
Drug treatments	48
mBMC drug treatment	49
Nalm6 parental and <i>RAD54</i> ^{-/-} treatment.....	49
Patient cell treatment.....	50
ADH activity assay	50
ALDH activity assay	51
Xenograft model	51
Virus construction and cell infection	52
Sensitivity to hydrogen peroxide (H ₂ O ₂)	52
H2DCFDA assay	53
TMEJ activity assay	53
GSH sensitivity	54
<i>In vitro</i> cigarette smoke exposure	54
<i>In vivo</i> cigarette smoke exposure.....	54
3. RESULTS	56
Polθ does not cooperate with ADH5 or ALDH2 to protect hematopoietic cells from physiological levels of formaldehyde.....	56
Polθ cooperates with ADH5 or ALDH2 to protect hematopoietic cells challenged by high levels of formaldehyde.....	62
Pharmacological inhibition of ADH5 or ALDH2 increases formaldehyde concentrations and reduced the growth of <i>Polq</i> ^{-/-} HSPCs	67
Polθ and ADH5 or ALDH2 protect AML and MPN cells from the genotoxic effects of formaldehyde induced by OTKs.....	71
Simultaneous inhibition of Polθ and ADH5 or ALDH2 exerts an anti-leukemia effect.....	76
Genetic knockout mice/cells challenged by cigarette smoke	79

4. DISCUSSION	85
Dual targeting of Polθ and formaldehyde detoxifiers (ADH5 or ALDH2)	85
Formaldehyde and lung cancer	94
Acetaldehyde and cancer	96
Identifying formaldehyde response genes	100
Conclusions	102
REFERENCES	105
APPENDIX	
A. SUPPLEMENTARY DATA	120

LIST OF FIGURES

Figure	Page
1. Exogenous sources of formaldehyde.....	3
2. Endogenous sources of formaldehyde.....	6
3. Oxidation of formaldehyde and acetaldehyde.....	8
4. “2-Tier” protective mechanism in hematopoietic cells.....	14
5. Repair of formaldehyde-mediated DNA damage.....	20
6. Synthetic lethality in cancer cells.....	22
7. Simultaneous inactivation of BRCA1/2 and Polθ is synthetically lethal.....	23
8. Breeding protocol to obtain double knockout mice.....	56
9. Analysis of organs from aged genetic knockout mice.....	57
10. Analysis of peripheral blood from young (3-4-months-old) mice.....	58
11. Analysis of peripheral blood from aged (18-24 months-old) mice.....	59
12. Analysis of bone marrow from young mice.....	61
13. Analysis of bone marrow from aged mice.....	61
14. Methanol increases formaldehyde levels in murine blood.....	63
15. Polθ and ADH5 or ALDH2 cooperate to protect murine HSPCs from methanol.....	63
16. Analysis of peripheral blood from mice after methanol challenge.....	64
17. Polθ and ADH5 or ALDH2 cooperate to protect HSPCs from formaldehyde.....	66
18. L-BSO enhances synthetic lethality in <i>Aldh2^{-/-};Polq^{-/-}</i> HSPCs.....	66
19. Simultaneous inhibition of Polθ and ALDH2 is toxic for murine HSPCs.....	67
20. Simultaneous inhibition of Polθ and ADH5 is toxic for murine HSPCs.....	68
21. Simultaneous inhibition of Polθ and ADH5 or ALDH2 is toxic to murine HSPCs.....	69
22. <i>Polq^{-/-}</i> HSPCs are not more sensitive to oxidative DNA damage.....	70
23. Polθ and ADH5 or ALDH2 protect murine HSPCs from formaldehyde challenge triggered by FLT3(ITD) or JAK2(V617F) OTKs.....	72

24. Polθ and ADH5 or ALDH2 protect AML patient cells from formaldehyde challenge triggered by FLT3(ITD).....	73
25. Polθ and ADH5 or ALDH2 protect MPN patient cells from formaldehyde challenge triggered by JAK2(V617F)	74
26. Global analysis of ADH5, ALDH2 and POLQ expression in cancer	75
27. Combination of Polθi + ALDHi and Polθi + ADH5i increased DNA damage in HR-deficient leukemia cells	76
28. Combination of Polθi + ALDHi and Polθi + ADH5i decreased proliferation of HR-deficient leukemia cells	77
29. Experimental timeline of mice injected with HR-deficient leukemia cells	78
30. Combination of Polθi + ALDHi and Polθi + ADH5i decreased leukemia burden and increased survival time in mice injected with HR-deficient leukemia cells	79
31. Analysis of peripheral blood from mice exposed to CSE for 3 months	81
32. Analysis of peripheral blood from aged mice initially exposed to CSE.....	82
33. Analysis of bone marrow from aged mice initially exposed to CSE	83
34. Genetic knockout mice are more sensitive to cigarette smoke <i>in vitro</i>	84
35. Correlation between <i>ADH5/POLQ</i> and <i>ALDH2/POLQ</i> expression in LUAD	84
36. Newly discovered “3-Tier” protective mechanism in hematopoietic cells	86
37. Validation of genetic knockout mice	120
38. Combination of Polθi + ALDHi and Polθi + ADH5i induced apoptosis in HR-deficient Nalm6-RAD54 ^{-/-} cells	121

CHAPTER 1

INTRODUCTION

The ubiquitous nature of formaldehyde

Aldehydes are a class of organic compounds in which a carbon atom shares a double bond with an oxygen, a single bond with a hydrogen atom, and a single bond with another atom or group of atoms. The double bond between carbon and oxygen is a characteristic of all aldehydes and is known as the carbonyl group. Formaldehyde is the simplest aldehyde, with the carbonyl group bonded to two hydrogen atoms, making its structural formula CH_2O . Formaldehyde is an electrophile and can react with nucleophilic molecules, making it a highly reactive molecule (1). Formaldehyde is a colorless, flammable gas that has become one of the most widely used chemicals in the United States. Exposure to formaldehyde most commonly occurs by inhalation but can also happen by absorption through the skin or by ingestion. Formaldehyde plays an important role in everyday life with most people unaware of just how prevalent it is within our world.

Exogenous sources of formaldehyde

There are many industrial applications for formaldehyde which include its use as a resin and binding agent. Resins are widely used in plastics, adhesives, coating materials and laminates. Binding agents are important in the production of wood products including plywood, particleboard and paneling which are all used to make furniture and other household items (1-3). The release of formaldehyde from these products is the main source of indoor-related exposure. Over the years, indoor formaldehyde emissions have decreased with stricter regulations and improvements in manufacturing processes. Formaldehyde is also commonly used as a preservative, disinfectant and sterilizer in mortuaries and medical facilities (2, 4). In research laboratories, it is used to fix cells and

tissues while serving as a tool for cross-linking DNA, RNA and proteins. In healthcare settings, formaldehyde is used to inactivate viruses and bacteria for certain vaccines with trace amounts of formaldehyde proven to be safe for humans (5, 6). In the agriculture industry, formaldehyde is used as fungicide and germicide, protecting crops from different fungi and insects (2, 4). Household items including cosmetics and cleaning products also contain formaldehyde (1, 3).

Formaldehyde can also originate from the breakdown of many dietary components such as methanol and aspartame (7). Methanol naturally occurs in fruits, vegetables and fish with certain alcoholic beverages such as beer and wine also containing trace amounts (8). When consumed, methanol becomes oxidized by alcohol dehydrogenase 1b (ADH1b) in the liver, producing formaldehyde which can be further metabolized to formic acid or formate (8). Accumulation of formate via methanol metabolism may lead to the development of acute methanol toxicity which is a systemic condition primarily impacting the central nervous system and is marked by visual impairment and metabolic acidosis (9). Generation of formaldehyde might also occur through aspartame consumption. Aspartame is an artificial sweetener commonly used in the beverage industry as a low-calorie alternative to sucrose. When it's consumed, it is further broken down into methanol and then formaldehyde. The United States Food and Drug Administration (FDA) has approved aspartame as a safe artificial sweetener to consume, but there have been many ongoing debates about its long-term safety (8). Methanol is both an exogenous and endogenous source of formaldehyde with the latter discussed later in this chapter.

Combustion significantly contributes to the production of exogenous formaldehyde. Cigarette smoke, automobile exhaust, factories, and fossil fuels all produce this toxic metabolite (1, 3, 10, 11). When sugars present in cigarettes are burnt, formaldehyde is produced as a byproduct and is present in both the mainstream and side

stream smoke, so primary smokers are exposed as well as non-smokers through secondary exposure (10-12). There are dozens of carcinogens and toxins in cigarette smoke, but recent studies have suggested aldehydes, including formaldehyde, are the predominant forces inducing DNA damage and inhibiting DNA repair in tobacco smoke carcinogenesis (10). Cars, buses, and airplanes all generate copious amounts of exhaust from engine combustion and contribute to the production of exogenous formaldehyde (12). The combustion of wood is a natural contributor to formaldehyde emissions which can occur through forest fires as well as household use of wood stoves or fireplaces (1, 12). Not only are there seemingly unlimited exogenous sources of formaldehyde, but this toxic metabolite can also be naturally generated by various metabolic reactions within the cellular niche of living organisms.

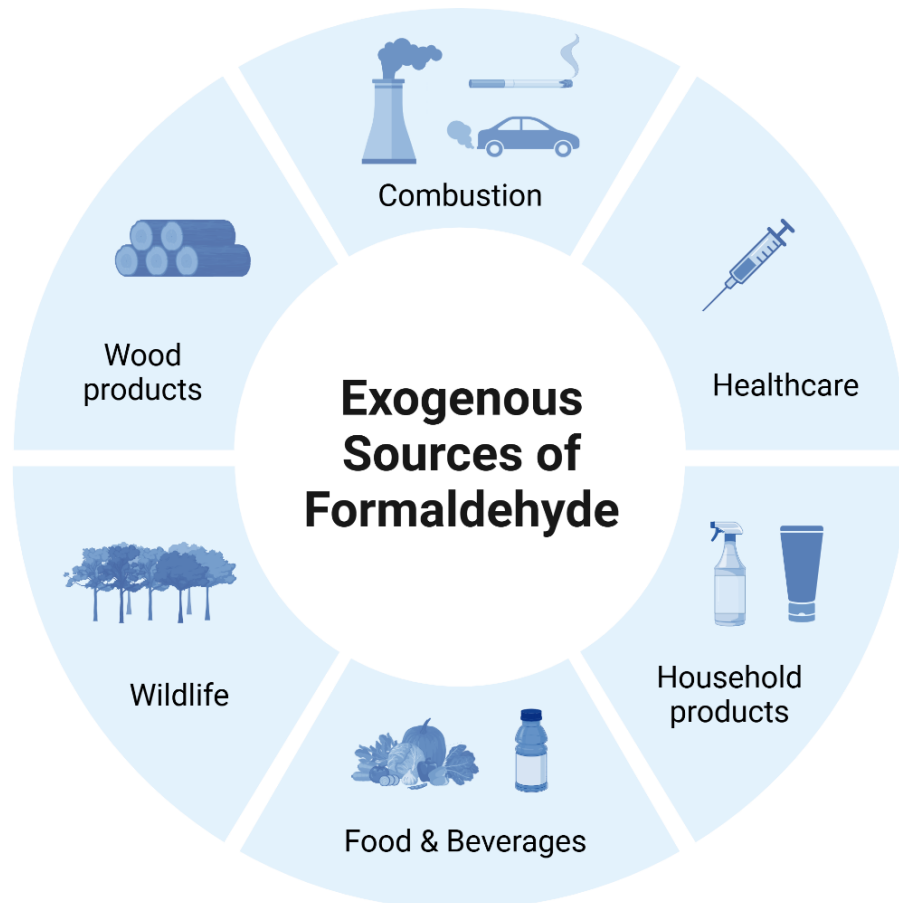


Figure 1. Exogenous sources of formaldehyde.

Endogenous sources of formaldehyde

Formaldehyde is the most abundant aldehyde within the human body with physiological concentrations ranging anywhere from 50-100 μM (13). It has been shown that endogenous formaldehyde is produced at high enough levels to threaten genome stability and can originate from various metabolic processes including methanol metabolism, demethylation and the one-carbon cycle.

Endogenously produced methanol can be directly produced by some bacteria through fermentation or the degradation of pectin, a soluble fiber found in many fruits and vegetables while indirectly generated by metabolic reactions involving S-adenosylmethionine (SAM) (8, 14). SAM is a universal methyl donor for several biological processes, including the methylation of various proteins, RNA and DNA (8). Carboxymethylation is a specific type of post-translational protein modification that involves the addition of a methyl group to a carboxyl group on certain amino acids. This process is catalyzed by enzymes called methyltransferases. Proteins that have undergone carboxymethylation where SAM served as a methyl donor can be hydrolyzed and the subsequent hydrolysis of methyl ester bonds produces methanol which is further broken down into formaldehyde and then formate (8, 15). Thus, methanol is both an exogenous and endogenous source of formaldehyde.

The second source of endogenous formaldehyde is demethylation. Oxidative demethylation is a chemical process that involves the removal of a methyl group from proteins, RNA or DNA, producing formaldehyde as a byproduct of this reaction (16, 17). Histone demethylases such as Jumonji C (JmjC)-containing oxygenases and lysine specific demethylase 1 (LSD1) have been shown to generate formaldehyde as a byproduct of their activity (14). When hematopoietic stem/progenitor cells (HSPCs) differentiate, they undergo rapid transcriptional reprogramming. This is a requirement for

HSPCs so that they can become mature blood cells. This process generates formaldehyde which can be attributed to the activity of demethylating enzymes, including JmjC and LSD1, since epigenetic modifications, including demethylation are essential for HSC differentiation (18, 19). There are also mitochondrial specific demethylases such as dimethylglycine dehydrogenase (DMGDH) and sarcosine dehydrogenase (SARDH) that catalyze the oxidative demethylation of dimethylglycine (DMG) to sarcosine and sarcosine transitions to glycine, producing formaldehyde as a byproduct (20, 21). Formaldehyde can also be generated through lipid peroxidation and the oxidation of glycerol via cytochrome P450 (CYP450) (21).

Another source of endogenous formaldehyde is the one-carbon (1C) cycle. The 1C cycle is a series of reactions that result in the synthesis of new nucleotides and amino acids that are critical for cell proliferation and genome stability. Formaldehyde is a key byproduct of many reactions within the 1C cycle. During this process, formaldehyde is generated from the enzymatic cleavage of serine (22). This process is carried out by serine hydroxymethyltransferase 1 and 2 (SHMT1 and SHMT2) which are the two key enzymes responsible for the cleavage of serine and are localized in the cytoplasm and mitochondria, respectively (14, 22). The oxidative decomposition of tetrahydrofolate (THF), an essential cofactor of the 1C cycle, also generates formaldehyde (22). As THF moves through the 1C cycle, it undergoes many chemical transitions, generating other folate derivatives and this turnover results in formaldehyde production. While formaldehyde is a product of the 1C cycle, it is also an important source of 1C units and supports nucleotide synthesis. 1C metabolism and formaldehyde detoxification systems overlap with each other to ensure homeostasis of physiological concentrations of formaldehyde within cells (22, 23).

Formaldehyde has become ubiquitous within the environment and our cells. It is the backbone of so many different industries, present in our own homes and pollutes the air we breathe every day. Our own cells continuously generate formaldehyde throughout our lifetime. Lifestyle choices like cigarette smoking and alcohol consumption or certain occupations such as industrial workers and embalmers also increase exposure to this molecule (24). Formaldehyde has become an invisible threat to humans. It is highly reactive and since it is present in our everyday lives, it is essential our cells can properly maintain a homeostatic balance of physiological levels of formaldehyde to avoid any cellular toxicity that could result in adverse health outcomes.

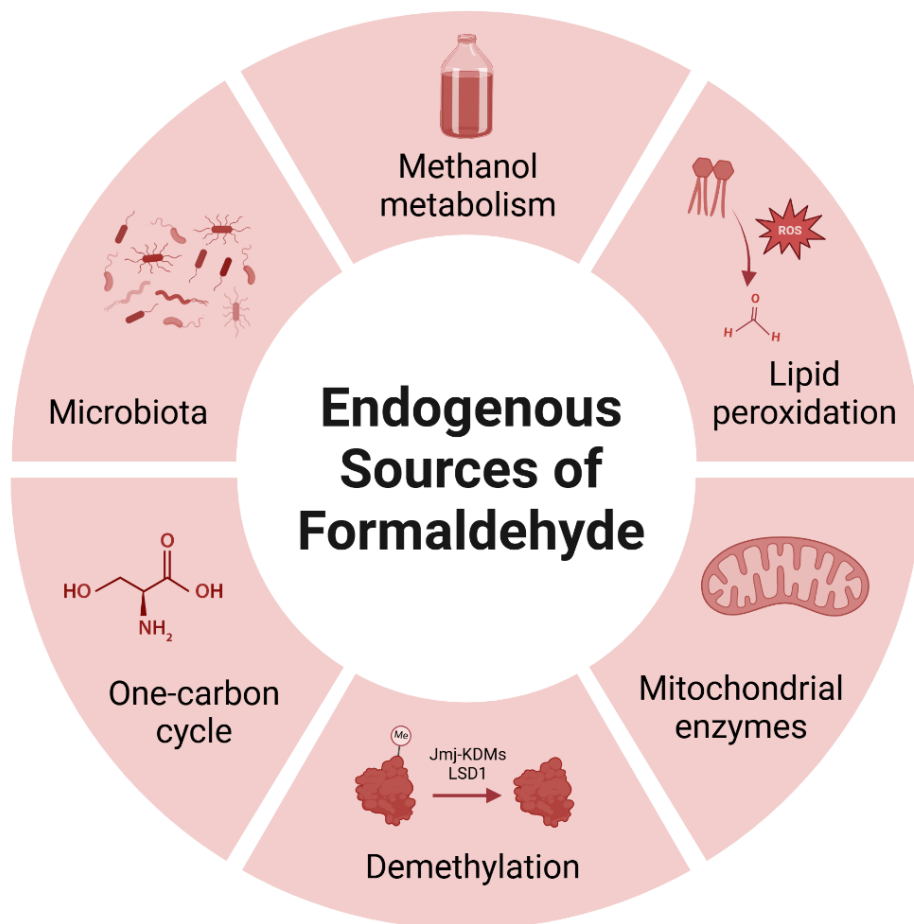


Figure 2. Endogenous sources of formaldehyde.

Intracellular formaldehyde metabolism

Formaldehyde is a highly reactive metabolite which is why cells are equipped with highly specific proteins that are responsible for the breakdown of formaldehyde, maintaining the delicate metabolic balance of physiological concentrations of formaldehyde. Alcohol dehydrogenase 5 (ADH5/GSNOR) and aldehyde dehydrogenase 2 (ALDH2) are two of the most important enzymes in formaldehyde detoxification. ADH5 and ALDH2 are dehydrogenases and oxidoreductases, and their main function is to oxidize specific substrates by reducing electron acceptors. More specifically, they are responsible for the oxidation of aldehydes, alcohols, and ketones.

ADH5 is a cytosol specific enzyme that belongs to the ADH family and has the highest affinity to formaldehyde with a K_m reported as low as $0.12 \mu\text{M}$ (25). The carbonyl group of formaldehyde reacts with the redox-active thiol group of glutathione (GSH), resulting in the formation of the intermediate product S-hydroxymethyl-GSH (HSMGSH) (26). HSMGSH is a substrate for ADH5, and this intermediate molecule is oxidized by ADH5 to yield formate which can be utilized for other biological processes like the 1C cycle (26). ADH5 dually functions as a S-nitrosoglutathione reductase (GSNOR), an important regulator of nitric oxide (NO) signaling and homeostasis (27).

On the other hand, ALDH2 is a mitochondrial specific enzyme that belongs to the ALDH family and has been shown to oxidize both formaldehyde and acetaldehyde. While ALDH2 has a low K_m for acetaldehyde ($< 1 \mu\text{M}$), it has been shown to metabolize formaldehyde with a K_m ranging from $170\text{-}400 \mu\text{M}$, making it a major compensatory pathway for formaldehyde detoxification (25, 28). Out of the 19 ALDH family members, ALDH2 is the only one known to play a role in the breakdown of formaldehyde (29). ADH5 and ALDH2 are the two main formaldehyde detoxifiers and are highly expressed in hematopoietic progenitor cells (HPCs) (25, 28).

Methanol and ethanol are oxidized by members of the ADH family, generating the toxic metabolites formaldehyde and acetaldehyde, respectively. ADH5 and ALDH2 are important for the detoxification of these toxic intermediate molecules. Their enzymatic activities result in the generation of the less toxic metabolites formate and acetate which are then used in other biological processes. Acetate is a prerequisite to the generation of acetyl coA while formate or formic acid can further be converted to CO₂ and H₂O.

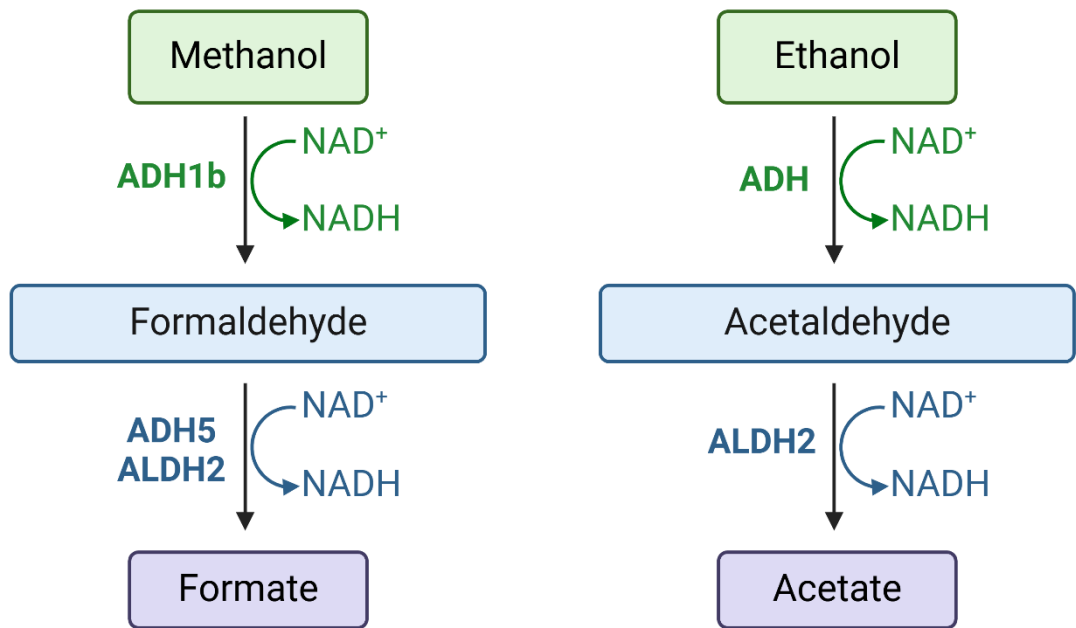


Figure 3. Oxidation of formaldehyde and acetaldehyde. ADH5 and ALDH2 are the primary enzymes responsible for formaldehyde detoxification. ALDH2, but not ADH5, is also important in acetaldehyde detoxification.

Cytochrome P450-dependent monooxygenase (CYP450) has also been shown to have the capacity to metabolize formaldehyde but plays a considerably minor role compared to ADH5 and ALDH2. CYP450 resides in the smooth endoplasmic reticulum and is known to be involved in the metabolism of various molecules including aldehydes and alcohols. These monooxygenases have also been shown to generate formaldehyde through the oxidation of glycerol in an NADPH-dependent process (14, 30).

Formaldehyde causes genome instability

Formaldehyde is a highly reactive aldehyde due to the lack of electron-donating groups attached to its carbonyl carbon, making it a strong electrophile (1). This allows formaldehyde to react with nucleophilic groups of nucleic acid bases and/or nucleophilic amino acids of proteins, so it can easily react with DNA, creating DNA damage (29). It is also a small molecule, allowing it to easily diffuse through all cellular compartments. These characteristics make formaldehyde a threat to genome stability. Enzymes like ADH5 and ALDH2 function to maintain a balance of physiological levels of formaldehyde while DNA repair pathways minimize formaldehyde-mediated DNA damage. Dysfunction of these metabolic enzymes and/or DNA repair proteins have proven to be deleterious for cells, resulting in genome instability, a key hallmark of cancer (31). This can lead to diseases such as bone marrow failure (BMF) and cancer. Thus, formaldehyde can be further characterized as genotoxic and carcinogenic.

Fanconi anemia (FA) is a rare inherited BMF syndrome where patients have mutations in FA DNA repair genes. The FA DNA repair pathway consists of about 22 DNA repair genes (*FANC*) that are crucial for maintaining genomic stability by repairing interstrand crosslinks (ICLs) in cells (32). FA is characterized by the inability to repair ICLs, hematopoietic stem cell (HSC) failure due to high levels of genome instability and developmental defects (33). Patients with FA also have an increased predisposition to developing aplastic anemia, myelodysplastic syndrome (MDS) and leukemia (32, 33). Cells with mutations in FA genes show extreme sensitivity to formaldehyde. Reactive metabolites like formaldehyde and acetaldehyde are known to crosslink with DNA to create aldehyde-mediated ICLs in HSCs. Furthermore, mutations in FA DNA repair genes result in the accumulation of formaldehyde-mediated ICLs, increasing genome instability and the likelihood of developing FA-associated cancers.

ICL-mediated repair pathway

Replication fork progression can be impeded by the presence of bulky DNA lesions such as ICLs. The FA DNA repair pathway is responsible for the removal of these ICLs, so DNA replication and transcription can continue without any issues. Stalled replication forks at ICLs are detected by E3 ubiquitin-protein ligase (UHRF1) or by a complex comprised of FA proteins, FA-associated proteins (FAAP24) and histone fold-containing proteins MHF1 and MHF2 (FANCM-FAAP24-MHF1-MHF2) during the S phase of the cell cycle (33, 34). The ataxia telangiectasia (ATR) and checkpoint kinase 1 (CHK1) cell-cycle checkpoint is activated, slowing DNA replication and allowing time for DNA repair to occur (33). The FANCM-FAAP24-MHF1-MHF2 complex is responsible for the recruitment of the FA core complex (composed of 10 FA proteins) and this core complex monoubiquitinates the FANCD2-FANCI heterodimer along with FANCT (33, 34). Ubiquitylated FANCD2-FANCI recruits FANCP/SLX4, a scaffolding protein for the DNA endonucleases MUS81, SLX1 and FANCG so they can unhook the ICL, creating a DNA adduct and a double strand break (DSB) (33, 34). The strand of DNA containing the DNA adduct needs to be bypassed for DNA replication to resume and this is executed by REV1, REV3 and FANCV/REV7 (REV3/REV7 are both subunits of Pol ζ) (33, 34). The ICL-derived DSB can then be repaired by multiple DSB repair pathways.

The FA dependent homologous recombination (HR) repair pathway is a high-fidelity DNA repair mechanism, ensuring accurate repair of the ICL-derived DSB. The ICL-derived DSB is removed by the MRN complex, CtBP (C-terminal binding protein) interacting protein (CtIP), exonuclease 1 (EXO1) and DNA replication helicase/nuclease 2 (DNA2), generating a 3' single-stranded DNA (ssDNA) overhang that is subsequently coated by replication protein A (RPA) (34). The MRN complex consists of meiotic recombination 11 (MRE11), RAD50 and Nijmegen breakage syndrome protein 1 (NBS1).

This complex retains endo and exonuclease activity through the Mre11 subunit which is essential for initiating end resection of DNA ends (35). This protein complex acts as a highly sensitive molecular sensor, important in detecting and processing DSBs. RAD51 recognizes and displaces RPA from the ssDNA overhang, leading to the formation of a RAD51 nucleofilament with the help of breast cancer type 1 (BRCA1), breast cancer type 2 (BRCA2) and partner and localizer of BRCA2 (PALB2) (34). Strand invasion and synthesis occurs followed by extension, ligation and completion of the ICL-derived DSB (33). In addition to the repair of ICLs, the FA DNA repair pathway has been known to be involved in maintaining replication fork stability, telomere length maintenance, cytokinesis and even autophagy (34). Biallelic germline mutations in many of the FA complementation group (*FANC*) genes including *FANCA*, *FANCB*, *FANCC*, *FANCD2*, *FANCE*, and *FANCI* cause FA which is characterized by hypersensitivity to ICL-inducing agents and the predisposition to myeloid malignancies, solid tumors and congenital abnormalities (34). Carriers of monoallelic mutations in *BRCA1/2*, and *PALB2* genes have an increased risk of developing multiple cancers including breast and ovarian cancers (33, 34). Acquired mutations in FA genes occur in multiple cancers and a portion of aggressive forms of ovarian, breast and prostate cancer have alterations in the FA DNA repair pathway (34).

Bone marrow failure syndromes

In mouse models of FA, a single mutation in *FANCD2*, an important gene in the FA DNA repair pathway, only resulted in mild effects. It wasn't until the additional deletion of *ADH5* or *ALDH2* that mice began to develop severe developmental defects such as HSC failure and leukemia (23, 36, 37). *Adh5^{-/-};Fancd2^{-/-}* and *Aldh2^{-/-};Fancd2^{-/-}* mice shared similar clinical features seen in FA patients, implicating formaldehyde as a genotoxic agent responsible for HSC failure in FA (23, 36, 37). Together, these findings demonstrate that formaldehyde-mediated ICLs play a significant role in the pathogenesis of bone marrow

failure in FA patients. More importantly, it was discovered that the metabolic enzymes ADH5 and ALDH2 work together with the FA DNA repair pathway to maintain normal HSC and blood cell functions. Mice lacking both *ADH5* and *ALDH2* exhibited partial synthetic lethality, growth failure and the development of lymphoid malignancies such as leukemia (25). Hematopoietic functions were severely disrupted with a significant reduction in HSCs in the *Adh5^{-/-};Aldh2^{-/-}* double knockout mice (25, 28). To confirm the severe phenotype in *Adh5^{-/-};Aldh2^{-/-}* mice is due to the accumulation of formaldehyde, blood samples were taken from these mice to measure formaldehyde concentrations. Formaldehyde levels in *Adh5^{-/-};Aldh2^{-/-}* double knockout mice were significantly higher compared to wild-type and single knockout counterparts which also correlated with an increase in DNA damage in HSCs (28). Using induced pluripotent stem cells (iPSCs), it was confirmed that the loss of *ADH5* and *ALDH2* led to cytotoxicity brought on by accumulation of DNA damage, even with intact DNA repair systems (38). This emphasized the dual requirement of *ADH5* and *ALDH2* in formaldehyde detoxification. *ADH5* is the primary defense against formaldehyde damage while *ALDH2* provides backup (25, 28, 38). Strikingly, *ADH5/ALDH2*-deficient iPSCs stopped growing and showed increased DNA damage when they were forced into hematopoietic differentiation *in vitro*, consistent with epigenetic changes that result in the generation of abundant intracellular levels of formaldehyde which may play a role during development (38).

Based on *in vivo* and *ex vivo* studies, it was thought that there may be a human disease that could arise from the simultaneous loss of *ADH5* and *ALDH2*. In humans, there are over 500 million people worldwide that have a dominant negative *ALDH2*2* polymorphism (39). This small nucleotide polymorphism (SNP) known as *ALDH2*2* or rs671 encodes an E504K missense mutation, where the amino acid glutamate is substituted for lysine and is abundant within the East Asian population, reducing the

enzymatic activity of the protein by about 90% in a dominant negative manner (28, 40). Presence of the mutation results in a red flushing reaction upon alcohol consumption due to the buildup of acetaldehyde in the body (41). Patients who are carriers of rs671 in addition to mutations in FA genes showed accelerated progression of FA (38, 42). Through whole-exome and targeted exome sequencing, it was found that a set of Japanese patients with BMF of unknown origin possessed an *ADH5* mutation and the *ALDH2**2 polymorphism (28, 38). Patients did not possess any other pathogenic alterations in genes known to be associated with other BMF syndromes like FA, but did share many clinical features seen in FA such as aplastic anemia that later developed into MDS or leukemia. Additional overlapping conditions include skin pigmentation, short stature and other developmental defects. The FA DNA repair pathway was found to be intact, confirming a unique BMF syndrome independent of FA. This rare inherited disease has been characterized as Aldehyde Degradation Deficiency Syndrome (ADDS) or AMeD (aplastic anemia, mental retardation, dwarfism) syndrome. Patients have no mutations in FA DNA repair genes, but rather mutations in both *ADH5* and *ALDH2* genes (32, 43, 44). Individuals with ADDS all share clinical features of aplastic anemia, mental retardation and dwarfism. With the overlapping clinical features and genetic interplay between *ADH5*, *ALDH2* and FA DNA repair pathway, it confirms that formaldehyde plays a role in FA pathogenesis as well as ADDS (23, 36, 38, 45, 46).

ADH5 and *ALDH2* are highly expressed in HSPCs, indicating the blood as a critical site of production of physiological formaldehyde (28). Therefore, maintaining genome stability is essential for normal HSC functions. To protect HSCs from genome instability, coordination between metabolic enzymes and DNA repair systems is essential. There is a “2-Tier” mechanism that exists within hematopoietic cells that protects them from formaldehyde-mediated DNA damage and genotoxicity (21, 23, 36, 37). The first tier of

protection involves ADH5 and ALDH2, providing formaldehyde detoxification while the FA DNA repair pathway eliminates formaldehyde-mediated ICLs, providing the second tier of protection (Figure 4). Dysfunction of either or both tiers results in the improper clearance of formaldehyde and its accumulation overwhelms DNA repair, leading to genomic instability and the introduction of mutations. This can result in lethal consequences including HSC failure and an increased likelihood of developing BMF syndromes, MDS, leukemia and/or age-related disorders. Thus, maintaining physiological levels of formaldehyde is essential for normal hematopoiesis and HSC functions.

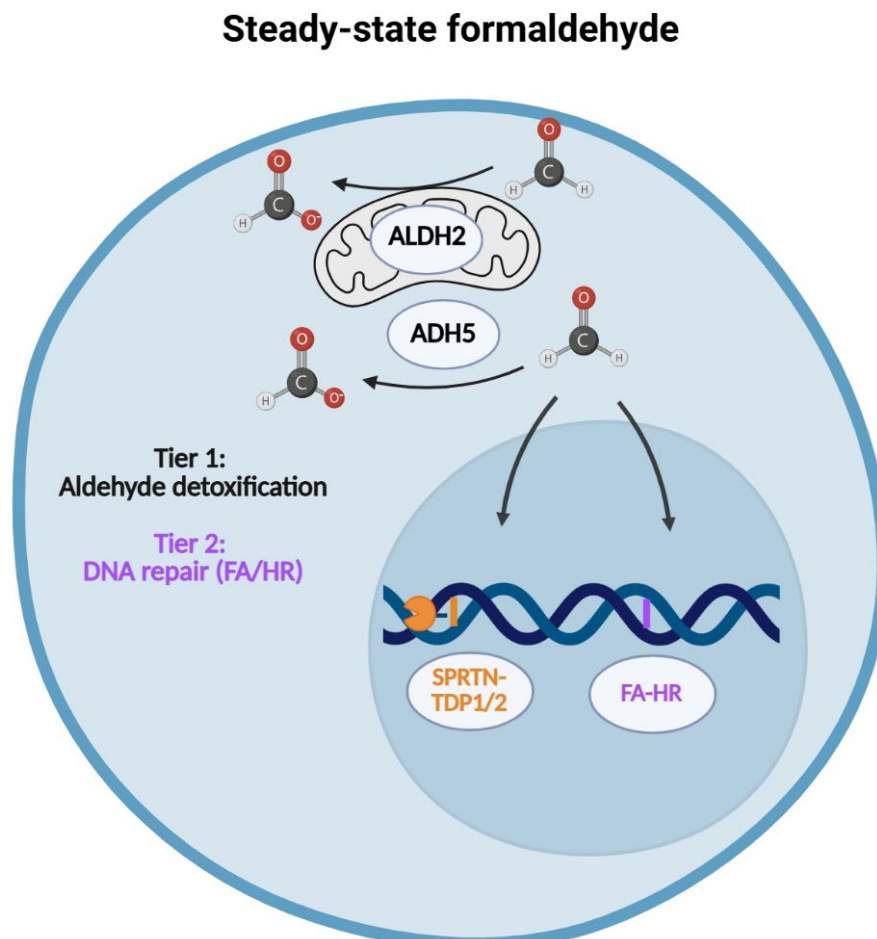


Figure 4. “2-Tier” protective mechanism in hematopoietic cells. Tier-1 protection: ADH5 and ALDH2 break down formaldehyde. Tier-2 protection: FA pathway is responsible for the repair of formaldehyde-mediated ICLs. Infrequent DPCs are digested by SPRTN and TDP1/2. This 2-Tier system protects HSPCs from formaldehyde-mediated genotoxicity.

Aldehydes are metabolic drivers of BMF syndromes like FA and ADDS, but their influence on carcinogenesis has remained elusive. Formaldehyde has been a known human carcinogen for over four decades, but *how* formaldehyde contributes to carcinogenesis is largely unknown. The dominant negative *ALDH2*2* polymorphism has been associated with an increase in cardiovascular diseases and certain types of cancer while *ADH5* mutations have been mostly associated with cardiovascular diseases (47-52). The interplay between *ADH5*, *ALDH2* and DNA repair proteins may be the key to understanding how formaldehyde contributes to cancer, specifically hematological malignancies such as leukemia.

DPC-mediated repair pathways

In addition to ICLs, aldehydes can generate DNA-protein crosslinks (DPCs), highly toxic DNA lesions that occur when DNA and proteins become covalently crosslinked (53-56). Non-enzymatic or non-specific DPC formation can occur from the use of many chemotherapeutic agents (platinum-based, PARP inhibitors) or highly reactive metabolites such as formaldehyde and acetaldehyde (35). Enzymatic DPC formation is caused by specific enzymes such as topoisomerases (*TOP1/2*) that form covalent intermediates with DNA. Etoposide, doxorubicin and camptothecin are all chemotherapeutic agents that selectively trap topoisomerase within the DNA, creating DPCs with the bulky lesions often forming at the site of double strand breaks (DSBs) (35, 57). The bulky nature of DPCs can prevent replication fork progression, resulting in the blockage of DNA replication and transcription (14, 35, 54, 55). There is great variability when it comes to the size, origin and type of protein or DNA structure involved in DPCs which is why they cannot be repaired directly by canonical DNA repair pathways.

There are three distinct components of a DPC that can each be targeted for repair: DNA, protein and the covalent crosslink (Figure 5) (35). Nuclease-dependent repair

requires nucleases, specifically the MRN complex, to make specific cuts on the DNA to excise the DPC, producing a 3' ssDNA overhang that is further repaired by canonical DNA repair mechanisms (58, 59). Protease-dependent repair of DPCs occurs by proteasomal degradation of DPC-containing proteins through a metalloprotease called SPRTN. In DPCs, DNA and proteins are covalently crosslinked and tyrosyl-DNA phosphodiesterase 1 or 2 (TDP1, TDP2) can hydrolyze the covalent link to release the protein from DNA (35). Canonical repair pathways such as homologous recombination (HR) and the FA DNA pathway have been implicated in the downstream repair of DPCs, but there is currently no clear or defined role for each DNA repair pathway or how much they contribute to the direct repair of DPCs (35, 60-62).

DSB repair pathways

DSBs are the most lethal kind of DNA damage, posing an immediate threat to genome stability. If DSBs are not quickly addressed, they may cause genome instability, promote tumorigenesis and/or cell death in mammals. Exogenous sources of DSBs can arise from exposure to ionizing radiation (IR), reactive oxygen species (ROS) and chemotherapeutic agents while endogenous sources of DSBs occur when DNA replication forks encounter unrepaired DNA damage such as ICLs or DPCs which cause replication fork collapse (63). HR and canonical non-homologous end joining (c-NHEJ) are the two canonical DSB repair pathways while DNA polymerase theta mediated end-joining (TMEJ) has emerged as an alternate end-joining DSB repair pathway (Figure 5). c-NHEJ is an error prone pathway that joins two DNA ends with minimal reference to DNA sequence while HR is an error-free repair pathway that requires extensive sequence homology and shares the same initial end resection step with TMEJ (64). DSB repair pathways are highly regulated to ensure accurate repair, and DSB repair pathway choice depends on multiple factors including cell cycle phase and the length or size of the DSBs (63, 65, 66).

c-NHEJ is a DSB repair pathway that is active throughout the cell cycle, but most prominently in the G0/G1 phases where there is a high abundance of c-NHEJ specific proteins such as p53-binding protein 1 (53BP1) and Ku autoantigen proteins p70 and p80 (Ku70-Ku80) (65). This DSB repair pathway directly relegates DNA ends without requiring substantial end processing by introducing small insertions or deletions at break sites, making this an error-prone DSB repair pathway. The Ku70-Ku80 heterodimer detects DSBs, binds to the DNA ends and helps recruit other factors to promote the ligation of DNA ends. 53BP1 and Ku70-80 are responsible for DNA end protection while DNA-dependent protein kinase catalytic subunit (DNA-PKcs) is responsible for recruiting DNA ligase 4 (LIG4) and x-ray repair cross complementing 4 (XRCC4) for the ligation of processed ends (Figure 5) (67).

Unlike c-NHEJ, HR repair occurs selectively in the S and G2 phases of the cell cycle because of its reliance on a sister chromatid to use as a complementary template for DNA repair (65). HR is initiated when the resection of the DSB takes place, producing a 3' ssDNA overhang. Short-range end resection is carried out by the MRN complex along with CtIP, with Mre11 exhibiting 3'-5' exonuclease and endonuclease activity while long-range end resection requires EXO1 along with DNA2 and other helicases (65). The MRN complex also recruits ataxia-telangiectasia mutated kinase (ATM) which phosphorylates histone H2AX, forming γ H2AX foci (common marker of DSBs) (68, 69). The formation of γ H2AX foci results in chromatin relaxation and the recruitment of BRCA1 and BRCA1-Associated RING Domain 1 (BARD1) (70). BARD1 is the main binding partner of BRCA1, forming the stable BRCA1-BARD1 complex that has many functions in HR repair including mediating efficient and prolonged end resection to generate 3' ssDNA (70, 71). The 3' ssDNA overhang is coated with replication protein A (RPA) to protect the ssDNA strand, and RPA is recognized and replaced by RAD51 recombinase with the help of BRCA1,

BRCA2, and PALB2 (65, 72). BRCA1 facilitates the recruitment of BRCA2 to the DSBs by binding to PALB2, a scaffold protein which in turn interacts with BRCA2 and RAD51 (73, 74). The RAD51 nucleofilament is used for homology search and promotes strand exchange using a sister chromatid as a template for accurate restoring of the original DNA sequence (Figure 5). RAD51 recombinase promotes strand exchange during which ssDNA invades homologous duplex DNA forming a D-loop and the 3' end in the D-loop is extended by DNA synthesis, and then the newly synthesized DNA strand dissociates to anneal to the other DNA end to complete the reaction (66). RAD54 is a protein that interacts with RAD51 and stimulates the RAD51 DNA strand exchange activity (75).

In more recent years, an alternate end-joining (alt-EJ) repair pathway was discovered, referred to as microhomology-mediated end-joining (MMEJ) or DNA polymerase theta mediated end-joining (TMEJ). The change in terminology reflects the discovery that DNA polymerase theta (Pol θ) is required for this type of repair (72, 76, 77). TMEJ is a DSB repair pathway that is independent of both c-NHEJ and HR and involves the alignment of microhomology sequences for end joining (76, 78). TMEJ uses microhomology sequence tracts for DSB repair which often leads to template insertions or deletions, making it an error-prone DSB repair pathway (72, 76). TMEJ is initiated by Poly (ADP-ribose) polymerase 1 (PARP1), an important DNA damage sensor that is recruited to resected DNA ends (72). Once PARP1 binds to DNA ends, it effectively blocks Ku proteins from binding to the ends, relegating the cell to TMEJ (65). The MRN complex and CtIP are responsible for end resection, generating 3' ssDNA overhangs. Pol θ then binds to the ssDNA overhangs and anneals sequences with 2-6 base pairs of microhomology to use them as primers for DNA synthesis (72). The stabilized DNA ends are then ligated by DNA ligase 3 (LIG3) and x-ray repair cross complementing 1 (XRCC1) (Figure 5) (72).

Pol θ is encoded by the *POLQ* gene and is classified as an error-prone translesion polymerase that has an N-terminal helicase domain and a C-terminal polymerase domain, separated by a central domain (76, 79). It is the only eukaryotic polymerase currently known to have a helicase domain, adding to its versatility. The polymerase domain is responsible for strand annealing and template extension while the helicase domain contains both ATPase and helicase activity and can counteract RPA and RAD51 by displacing them from ssDNA, preventing DNA repair by HR (72, 79-82). The helicase domain of Pol θ can also bind adjacent to the polymerase domain to prevent the snap back of ssDNA, allowing TMEJ to continue (72).

Both HR and TMEJ share the same end resection process via MRN complex, but TMEJ only requires limited (short-range) end resection while HR depends on extensive (long-range) resection using additional repair factors (77, 83). TMEJ is an error-prone DSB repair pathway which is why it has been thought to merely be a backup repair pathway to HR (72, 83). But certain types of replication blocking DNA lesions are converted to DSBs and since these cells cannot undergo replication and a sister chromatid is unavailable, TMEJ is the only remaining repair pathway (72). TMEJ is used to repair DSBs that are generated at collapsed replication forks, demonstrating its importance in helping cells with genomic stress maintenance (83). High expression of Pol θ has been shown to protect cells from replication stress even in the presence of HR (76, 84). Cells with *POLQ* depletion are also sensitive to replication fork stalling agents such as topoisomerase inhibitors. *Polq*^{-/-} genetic knockout mice are viable and grow normally but are sensitive to ionizing radiation (76). A cell's decision to undergo HR or TMEJ is context dependent and still under investigation.

POLQ as a therapeutic target for cancer

For decades, chemotherapy has been the standard of care for cancer patients. Chemotherapeutic agents are cytotoxic drugs that not only kill cancer cells, but they also affect normal cells which is why adverse side effects are a common occurrence. To combat this issue, there has been a shift towards personalized medicine in cancer research that is focused on analyzing genetic mutations and vulnerabilities unique to each patient's cancer. This approach is helpful in identifying potential therapeutic targets, specifically tailoring treatment for each patient. The transition from a "one-drug-fits all" to a more personalized approach allows for more effective and highly selective targeting of cancer cells while minimizing side effects (72).

The DNA damage response (DDR) is a complex cellular network that is responsible for sensing and responding to DNA damage, maintaining genome integrity and stability (85). Inherited or acquired mutations in DDR genes are associated with the accumulation of genome instability, contributing to cancer development (85, 86). One approach to personalized medicine is to target tumor-specific mutations in DNA repair genes. Cancer cells often acquire mutations in DNA repair genes, and they respond by rewiring their DNA repair network, depending on other compensatory pathways for survival. The targeting of these compensatory pathways leads to synthetic lethality (Figure 6). Therefore, therapeutic targeting of compensatory pathways offers a lot of promise with the selective killing of tumor cells while leaving normal cells unharmed. Mutations in DDR genes create vulnerabilities specific to cancer cells that can be exploited therapeutically. PARP inhibitors are a great example of synthetic lethality where the combination of a genetic defect (inactivation of *BRCA1/2*) and pharmacological inhibition of PARP proteins leads to cell death. The PARP family is a group of proteins that detect DNA damage (SSBs, DSBs), recruit DNA repair machinery and stabilize the replication fork during repair (85).

PARP inhibitors inhibit the catalytic activity of PARP with cells accumulating high levels of SSBs which left unrepaired, result in stalled replication forks, trapping of PARP within the strand of DNA and the conversion of SSBs to DSBs, leading to cell death (85). There are multiple PARP inhibitors that are FDA approved and are very effective in treating patients with HR-deficient cancers such as breast and ovarian cancers, but unfortunately PARP resistance has been a major hurdle in this therapeutic avenue (87). Finding new therapeutic targets to induce synthetic lethality and overcome drug resistance in cancer cells is crucial. Targeting other key components of the DDR such as *ATM*, *ATR* and *CHK1/2* are currently undergoing pre-clinical testing (88). In recent years, *POLQ* has also emerged as an exciting therapeutic target.

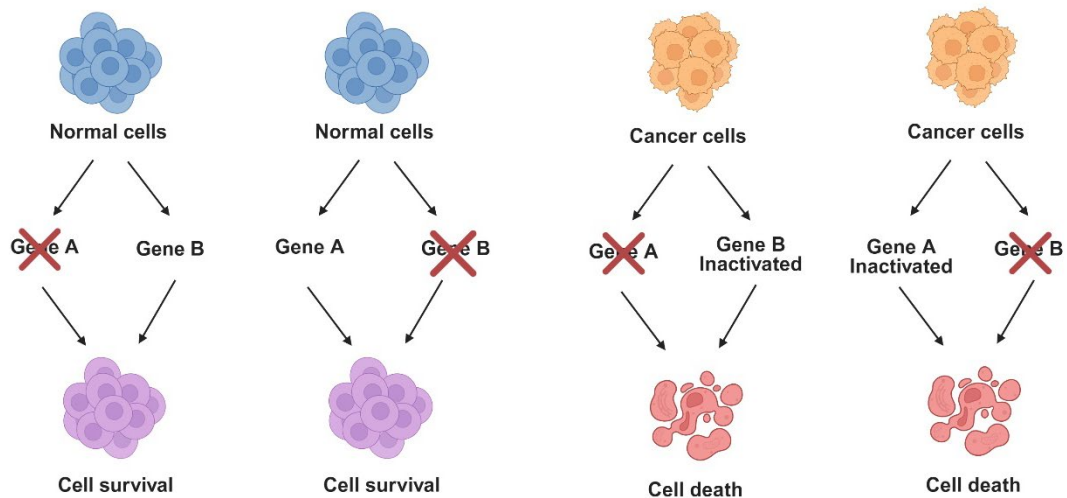


Figure 6. Synthetic lethality in cancer cells. Synthetic lethality occurs when two redundant pathways promote cell survival in normal cells. In cancer cells, one of these pathways is inactivated and they rely on the alternative pathway for cell survival. Targeting the alternative pathway will selectively result in cancer cell death while normal cells will be protected.

POLQ is overexpressed in multiple cancers and is also associated with poor prognosis (72, 81, 89, 90). Many HR-deficient cancer types have increased expression of Polθ to counteract the increase in replication stress and accumulation of DSBs. Synthetically lethal interactions have been identified between *POLQ* and other DNA repair

genes frequently mutated in cancer, including multiple genes important in the FA/HR repair pathways such as *BRCA1*, *BRCA2* and *FANCD2* (Figure 7) (72, 81, 91). The depletion of Pol θ overwhelms HR-deficient cells with DNA damage, triggering cell death (81). Both the helicase and polymerase domain of Pol θ contain druggable sites which has led to the development of multiple Pol θ inhibitors (Pol θ i) (72). Recently, ART558 was discovered as a novel drug that inhibits the polymerase domain of Pol θ , preventing cells from undergoing TMEJ and was shown to induce synthetic lethality in BRCA-deficient cancers while enhancing the effect of PARP inhibitors and even overcoming PARP resistance (92). Novobiocin (NVB) is another Pol θ i, but it directly binds to the ATPase domain, inhibiting its ATPase activity and HR-deficient cancer cells that acquired resistance to PARPi were found to be sensitive to NVB (93). Thus, Pol θ activity is essential in HR-deficient cells, making it a promising target for cancer treatment.

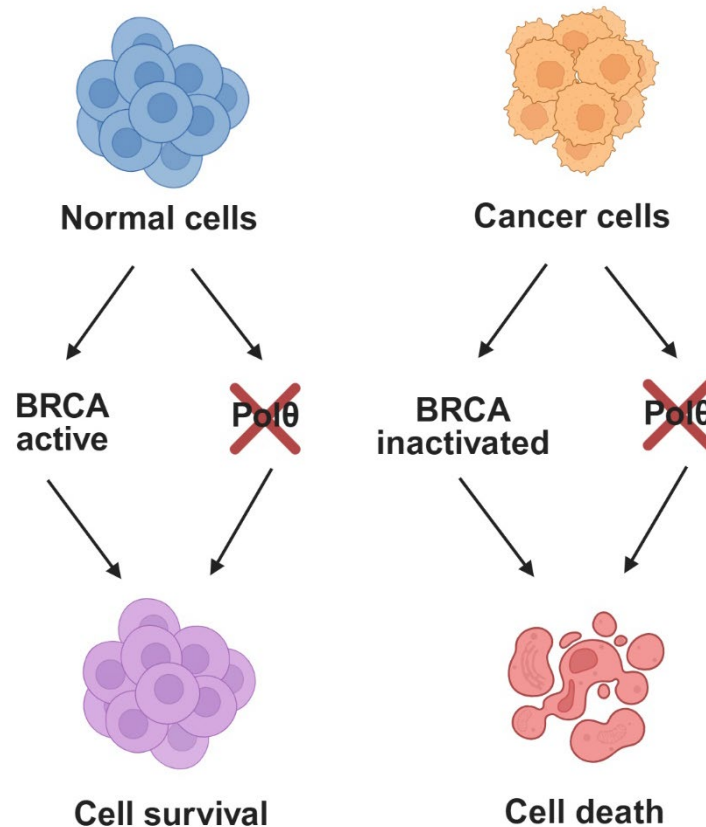


Figure 7. Simultaneous inactivation of BRCA1/2 and Pol θ is synthetically lethal.

Interplay between metabolic enzymes and DNA repair pathways

Formaldehyde causes more cancer than any other toxic air pollutant (94). It has become a pervasive threat to humans due to its broad industrial use and its release as a byproduct of combustion from factories, automobiles and cigarette smoke. It is also an occupational hazard for many industrial workers and medical professionals who are consistently exposed to formaldehyde. Studies have shown that increased formaldehyde exposure in the workplace induces DNA damage and increases the risk of certain cancers (95-99). *In vivo* exposure to formaldehyde led to the development of squamous cell carcinomas in rodents (97, 100-105). The correlation between formaldehyde exposure and blood cancer has been more widely debated with some epidemiological studies suggesting medical professionals such as embalmers and pathologists exposed to formaldehyde have increased risk of leukemia, while other studies have not found an increased risk of developing any hematological malignancy (95, 96, 99, 106, 107).

The biological mechanism for how formaldehyde contributes to leukemia has remained elusive. There is evidence that occupational exposure to formaldehyde leads to chromosomal damage in blood cells but whether significant levels of formaldehyde can get to the bone marrow, the site of leukemia induction is still uncertain (108-112). In a study of 94 factory workers in China, leukemia-specific chromosome changes, including monosomy 7 and trisomy 8 were observed and peripheral blood counts were significantly decreased (96). Additionally, studies have found increased DPCs in the bone marrow of mice exposed to formaldehyde (113-116). DPCs are not easily repairable and if they are not resolved quickly, it may result in permanent DNA alterations and genome instability (35, 117). These findings were consistent with the notion that formaldehyde exposure can have adverse effects on the hematopoietic system and leukemia induction by formaldehyde is biologically plausible (96).

A new analysis found that formaldehyde exposes everyone to an elevated risk of developing cancer no matter where they live (94). It's not only a threat to workers who use it, but everyone is at risk as it contaminates the air we breathe and is released from products commonly found in our homes.

There is a critical need to investigate the biological plausibility of a causal relationship between formaldehyde and blood cancers such as leukemia. Formaldehyde-mediated DPCs may be a potential biological mechanism for how formaldehyde can target the bone marrow, contributing to leukemogenesis. Formaldehyde-mediated DNA lesions are not equally genotoxic or mutagenic, so it is crucial to understand which lesions are responsible for diseases such as leukemia and the DNA repair pathways that are responsible for repairing them. This can lead to the identification of more targeted therapies and better clinical outcomes.

DPCs and the DNA repair pathways responsible for their repair are poorly understood, especially when it comes to their role in carcinogenesis (118). Understanding which DNA repair pathways are involved in downstream DPC repair will allow us to identify new therapeutic targets and overcome cellular resistance to chemotherapeutics that induce DPCs for diseases with high genomic instability like cancer. It was previously unknown whether Pol θ and TMEJ contribute to the repair of formaldehyde-induced DPC-associated DSBs. Pol θ provides resistance to chemotherapeutic agents like etoposide and camptothecin as well as ionizing radiation, suggesting TMEJ may play a role in the repair of DPCs that accumulate at DSBs (57, 76, 119, 120). Recently, it was discovered that Pol θ promotes the repair of 5' DPC-associated DSBs by TMEJ and Pol θ also confers resistance to formaldehyde (57). Pol θ promotes repair through TMEJ, as well as restarting stalled replication forks and filling ssDNA gaps during replication fork progression (72, 76, 121-123).

Knowing that Pol θ promotes the repair of DPCs in normal cells, the next step was to determine whether this phenomenon occurs in cancer cells. *POLQ* is highly upregulated in acute myeloid leukemia (AML), making it an attractive therapeutic target for AML (124). We found that formaldehyde generated by the 1C cycle contributed to the accumulation of DPCs in leukemia cells, especially in cells harboring oncogene tyrosine kinases (OTKs) such as FLT3(ITD) and JAK2(V617F) (124). Proto-oncogenes are genes that help regulate cell growth and division. Mutations in proto-oncogenes can result in their constitutive activation, creating an oncogene. More specifically, mutations in OTKs lead to their constitutive activation which can lead to uncontrollable cell growth and the induction of many tumors (125). Myeloid leukemias often express FLT3(ITD) whereas myeloproliferative neoplasms (MPNs) express JAK2(V617F). These OTKs enhanced the expression of Pol θ via ERK1/2 serine/threonine kinase dependent inhibition of c-CBL E3 ligase-mediated ubiquitination of Pol θ , preventing its proteasomal degradation (124). The overexpression of Pol θ ensures efficient repair of DPC-associated DSBs in leukemia cells.

Now, it is known that TMEJ can repair formaldehyde-mediated DPC-associated DSBs in both normal and malignant HSPCs. Therefore, we next wanted to determine whether Pol θ and TMEJ cooperate with formaldehyde detoxifiers ADH5 and ALDH2 to protect HSPCs from formaldehyde-mediated DNA damage, specifically DPC-associated DSBs. There is a "2-tier" protection mechanism that exists when it comes to the repair of formaldehyde-mediated ICLs in HSPCs, preventing the onset of BMF syndromes like Fanconi anemia and ADDS. However, it has remained unknown whether this "2-tiered" mechanism provides enough protection for HSPCs, especially in the presence of elevated levels of formaldehyde as seen in leukemia.

Here, a comprehensive genetic and pharmacological approach is described to determine the relationship between ADH5, ALDH2 and Pol θ and whether they work

together to protect normal and malignant HSPCs from formaldehyde-induced DPC-associated DSBs. This was done by generating *Adh5^{-/-};Polq^{-/-}* and *Aldh2^{-/-};Polq^{-/-}* double knockout mice for *in vitro* experiments to study ADH5, ALDH2 and Polθ under steady state formaldehyde concentrations. *In vitro* and *in vivo* formaldehyde challenges were also conducted to further elucidate the cooperativity of ADH5, ALDH2 and Polθ in HSPCs exposed to increased levels of formaldehyde. Leukemic cell lines and patient cells were treated with ADH5, ALDH and/or Polθ inhibitors to determine whether ADH5, ALDH2 and Polθ are viable therapeutic targets. Overall, this led to the discovery that ADH5, ALDH2 and Polθ do not seem to cooperate under physiological circumstances. However, when HSPCs are challenged by formaldehyde levels, mimicking environmental pressures, Polθ does cooperate with ADH5 and ALDH2 to provide a “3rd tier” or extra layer of protection against formaldehyde in HSPCs.

CHAPTER 2

METHODS

Mouse husbandry

All mouse experiments were approved by the Institutional Animal Care and Use Committees (IACUC) review board at Temple University. Mice were housed in a temperature and light-controlled animal facility under a 12h light/dark cycle and were provided standard food and water *ad libitum*. Wild-type and *Polq*^{-/-} mice were obtained from Jackson Laboratories (Cat#: JAX000664, JAX006194). *Adh5*^{-/-} mice were provided by Jonathan Stamler (Institute for Transformative Molecular Medicine, Case Western Reserve University School of Medicine, Cleveland, OH) and *Aldh2*^{-/-} were provided by Lopa Mishra (The Institute for Bioelectronic Medicine, Feinstein Institutes for Medical Research & Cold Spring Harbor Laboratory, Department of Medicine, Division of Gastroenterology and Hepatology, Northwell Health, Manhasset, New York).

Mouse breeding

To obtain wild-type, *Polq*^{-/-}, *Adh5*^{-/-}, *Aldh2*^{-/-}, *Adh5*^{-/-};*Polq*^{-/-} and *Aldh2*^{-/-};*Polq*^{-/-}, mice were set up as breeding pairs with one male and one female. The pairs were then used to maintain breeding colonies. There were at least three active colonies for each genotype, so there would be sufficient mice to use for future breeding pairs and experiments. The pups were weaned 21 days after birth. If any pups looked smaller than usual, a weaning extension was requested. Breeding pairs were allowed to deliver a maximum of six litters before they were separated. To generate *Adh5*^{-/-};*Polq*^{-/-} and *Aldh2*^{-/-};*Polq*^{-/-} double knockout mice, *Adh5*^{-/-} and *Aldh2*^{-/-} mice were first cross-bred with *Polq*^{-/-} mice, generating heterozygous mice (*Adh5*^{+/-};*Polq*^{+/-} and *Aldh2*^{+/-};*Polq*^{+/-}). Heterozygous breeding pairs (*Adh5*^{+/-};*Polq*^{+/-} x *Adh5*^{+/-};*Polq*^{+/-}, *Aldh2*^{+/-};*Polq*^{+/-} x *Aldh2*^{+/-};*Polq*^{+/-}) were then set up to obtain

double knockout mice. All mice were housed in the Medical Research and Education Building (MERB) 541. Mouse experiments that involved injecting any biological or chemical hazards were required to be housed under BSL-2 conditions in MERB 545. Therefore, all immunodeficient mice injected with primary human leukemia cells (Nalm6-*RAD54*^{-/-}) or any other biological hazard (methanol) were housed in MERB 545, the designated BSL-2 housing room.

Mouse genotyping

The genotypes of all mice were confirmed by polymerase chain reaction (PCR). To obtain a DNA sample for PCR, mice were anesthetized using an isoflurane machine located in MERB 542 and 547. While mice were anesthetized, they received an ear tag for identification purposes. An already sterilized ear tag was applied to the ear using an ear tag applicator. Once the ear tag was in place, the tail was wiped with 70% ethanol in preparation for the tail amputation. The procedure was performed using a sterile, single use, disposable scalpel. A maximum of 0.3 cm was amputated from the tail tip. Each tail snip was placed in a 1.5 mL Eppendorf tube (Thermo Fisher Scientific Cat#: 05-402-25), labeled with the corresponding ear tag number and brought back to the lab for DNA isolation and purification.

DNA isolation and purification was performed using the following protocol. 240 μ L of 50 mM NaOH was added to each tissue sample and heated at 95°C for 10-20 min with vortexing every 5 min. Then, 60 μ L of 1 M Tris HCl pH 8 was added to each sample, mixed, and centrifuged for 5 min at maximum speed (~14,000 rpm). The ratio of NaOH and Tris HCl was adjusted based on the size of the sample. For smaller tail snips, 60 μ L of NaOH and 15 μ L of 1 M Tris HCl pH 8 were used. Genotyping was performed using knockout specific primers (Eurofins) and DreamTaq PCR Master Mix (2X) (Thermo Fisher Scientific Cat#: K1081).

Polq-specific PCR reaction

Wild-type forward (1): 5' - TGCAGTGTACAGATGTTACTTTT - 3'

Wild-type reverse (2): 5' - TGGAGGTAGCATTCTTCTC - 3'

Mutant forward (3): 5' - TCACTAGGTTGGGGTTCTC - 3'

Mutant reverse (4): 5' - CATCAGAAGCTGACTCTAGAG - 3'

Two separate PCR reactions were set up. One reaction with WT primers 1 and 2 together and another reaction with mutant primers 3 and 4 together. The master mix recipe was as follows: Dream Taq (2X) 12.5 μ L, Primer 1 1.25 μ L, Primer 2 1.25 μ L and sterile H₂O 8.5 μ L. 23.5 μ L of master mix was added to each 0.2 mL PCR tube (Genesee Scientific Cat# 27-125) and 1.5 μ L of DNA was added to each sample to achieve a final reaction volume of 25 μ L. Tubes were briefly spun and placed in the T100 Thermo Cycler (Bio-Rad). The amplification conditions for *Polq* specific primers consisted of one cycle at 94°C for 3 min, followed by 34 cycles at 94°C for 30 sec, 60°C for 30 sec, and 68°C for 45 sec. Then, 68°C for 5 min and 4°C hold. The primers amplified a 190-base pair fragment if wild-type, a 300-base pair fragment if *Polq* null, and both 190- and 300-base pair fragments if heterozygous.

Adh5-specific PCR reaction

Wild-type (exon 5) 19se: 5'-GAAATGAGCCTGCCCTTCCA-3'

Wild-type (exon 5) 620as: 5'-CTGTCATCTCAACGAGGACTTC-3'

Mutant Neo: 5'-CGCCTTCTATCGCCTTCTTGACGAGT-3'

Mutant Exon7as: 5'-TTCTTCCCCTGAAGCAGCTA-3'

One reaction was set up combining all four primers together. The master mix recipe was as follows: Dream Taq (2X) 12.5 μ L, Primer 1 1.0 μ L, Primer 2 1.0 μ L, Primer 3 1.0 μ L, Primer 4 1.0 μ L and sterile H₂O 7.0 μ L. 23.5 μ L of master mix was added to each 0.2 mL PCR tube and 1.5 μ L of DNA was added to each sample to achieve a final reaction volume of 25 μ L. Tubes were briefly spun and placed in the T100 Thermo Cycler (Bio-Rad). The amplification conditions for *Adh5* specific primers consisted of one cycle at 95°C for 5 min, followed by 34 cycles at 95°C for 30 sec, 59°C for 30 sec, and 72°C for 1 min. Then, 72°C for 7 min and finally 4°C to hold. The primers amplified a 600-base pair fragment if wild type, 900-base pair fragment if *Adh5* null, and both 600- and 900-base pair fragment if heterozygous.

***Aldh2*-specific PCR reaction**

ST316 Forward (F): 5'-AAACTTTGCACACACTGTCCC-3'

ST317 Reverse 1 (R1): 5'-CCCAGATCCAACTGTTAGGAATAC-3'

ST318 Reverse 2 (R2): 5'-GCTTCACTGAGTCTCTGGCATCTC-3'

One reaction was set up combining all three primers together in a F:R1:R2=2:1:1 ratio. The master mix recipe was as follows: Dream Taq (2X) 10 μ L, Primer F 1.0 μ L, Primer R1 0.5 μ L, Primer R2 0.5 μ L and sterile H₂O 6.5 μ L. 18.5 μ L of master mix was added to each 0.2 mL PCR tube and 1.5 μ L of DNA was added to each sample to achieve a final reaction volume of 20 μ L. The amplification conditions for *Aldh2* specific primers consisted of one cycle at 95°C for 2 min, followed by 30 cycles at 95°C for 30 sec, 62°C for 30 sec, and 72°C for 30 sec. Then, 72°C for 5 min and 4°C hold. The primers amplified a 212-base pair fragment if wild type, 414-base pair fragment if *Aldh2* null and both 212- and 414-base pair fragments if heterozygous.

All PCR products were run on a 2.0% agarose gel (Invitrogen Cat#: 16500) and prepared in 1X TAE buffer. 50X TAE buffer was made using the following recipe: 121 g Tris Base, 28.55 mL acetic acid and 50 mL of 0.5 M EDTA with the total volume adjusted to 500 mL using deionized water. 20 mL of 50X TAE was dissolved in 980 mL deionized water to prepare 1X TAE. Before the gel was poured into a cast, ethidium bromide (Sigma-Aldrich Cat#: 46067) was added for product visualization. The products were resolved using an Electrophoresis System (Thermo Fisher Scientific) and visualized using an iBright 1500 Imaging System (Invitrogen). Products of interest were identified using the GeneRuler 100 bp ladder (Thermo Fisher Scientific #SM0241) which was resolved together with samples.

Mouse organ extraction for histopathology

All mice were euthanized following IACUC recommendation. Mice were placed in a CO₂ chamber that was gradually filled with CO₂ which was followed by cervical dislocation as a mandatory secondary method of euthanasia. Organs were then harvested under a sterile fume hood and placed in 4% paraformaldehyde (PFA). Organs were then given to Mariusz Wasik for analysis at the Department of Pathology, Fox Chase Cancer Center, Philadelphia, PA.

Mouse bone marrow cell (mBMC) collection

To obtain mBMCs, mice were euthanized following the exact method previously described. Using surgical scissors and forceps, the femur and tibia bones (2 femurs, 2 tibias per mouse) were isolated, placed in a 70 mm tissue culture dish filled with PBS and placed on ice until ready to isolate cells. Each set of bones were flushed with about 5 mL of PBS 5-10 times, added to a 15 mL tube and centrifuged for 5 min at 1000 rpm. PBS was aspirated from the tube and 5 mL of ammonium chloride potassium (ACK) lysis buffer (Thermo Fisher Scientific #A1049201) was added to mBMCs for red blood cell (RBC) lysis

on ice for about 5 min. Cells were centrifuged again for 5 min at 1000 rpm and were washed with PBS. Immediately after mBMC isolation, cells were counted and lineage depletion was performed using a Direct Lineage Cell Depletion Kit for mice (Miltenyi Biotec Cat#: 130-110-470), followed by Easy Sep™ Mouse CD117 (cKit+) Positive Selection Kit (StemCell Technologies Cat#: 18757). All procedures were performed using the manufacturer's protocol with no modifications. After obtaining the stem cell enriched population of cells, they were cultured using IMEM medium (Corning Cat#: 10-016-CV) supplemented with 10% FBS, 1% antibiotics, 6 ng/mL murine IL-3, 6 ng/mL murine IL-6 and 10 ng/mL murine SCF.

Hematological analysis of mouse peripheral blood

Mouse blood samples were taken either by tail nick or cardiac puncture. Blood collection by tail nick is a non-surgical procedure. The tail nick procedure was performed using a restrainer to keep the mouse in place during blood collection. A sterile, single use scalpel was used to make a small nick on the tail. About 50-100 µL blood was collected per mouse using blood collection tubes lined with K₂EDTA (BD Biosciences Cat#: 365974). Blood samples were placed on ice until they were ready to be processed. Per session, the tail was not punctured more than 4 times (2 times from each site). Hemostasis was achieved by applying pressure to the cut site. To stop bleeding, Clotisol liquid coagulant was applied to the cut site. For cardiac puncture, mice were euthanized and a 28G insulin syringe (Thermo Fisher Scientific Cat#: 14-841-31) was used to puncture the heart. Blood was collected using the same blood collection tubes lined with K₂EDTA and placed on ice. Approximately 25-30 µL of blood was used for the Hemavet HV950 Multispecies Hematology Instrument (Drew Scientific) to perform automated blood cell counts.

AML and MPN patient cells

AML 201 primary patient sample was previously described (126). Single-cell targeted DNA sequencing (sctDNA-seq) revealed that 93.48% of cells carried FLT3(D593ins) mutation. For AML1, bulk sequencing found FLT3(Y589_E611dup) mutation with the Variant Allele Frequency (VAF) of 53%. For MPN1 sample, sctDNA-seq revealed that 91.91% of cells carried JAK2(V617F) mutation. For P293 MPN sample, bulk sequencing detected the presence of JAK2(V617F) mutation. AML1 and MPN1 patient cells were acquired from the Department of Internal Medicine I, Division of Hematology & Hemostaseology, Medical University of Vienna, Austria. P293 cells were provided by the Department of Hematology, Transplantation and Internal Medicine, University Clinical Care, Medical University of Warsaw, and the Department of Hematology, Institute of Hematology and Blood Transfusion, Warsaw, Poland. These specific patient cells were cultured using StemSpan SFEM medium (StemCell Technologies Cat#: 09655) supplemented with 100 ng/mL human SCF, 20 ng/mL human IL-3, 100 ng/mL human FLT3 ligand, 20 ng/mL human IL-6, and 100 ng/mL human TPO (all purchased from Peprotech). All further procedures were performed on Lin-CD34⁺ cells obtained using EasySep™ Human Progenitor Cell Enrichment Kit with Platelet Depletion (StemCell Technologies #19356) and EasySep™ Human CD34 Positive Selection Kit II (StemCell Technologies #17856). All procedures were performed according to the manufacturer's protocol with no modifications. These patient cells were chosen based on their clonal composition where a major clone was either FLT3(ITD) or JAK2(V617F).

Cell lines

Nalm6 parental and Nalm6-*RAD54*^{-/-} cells were described previously (127). Nalm6 parental and Nalm6-*RAD54*^{-/-} were thawed from liquid nitrogen and cultured in RPMI 1640 medium (Genesee Scientific Cat#: 25-506) supplemented with 10% FBS (Gibco Cat#

10437-028) and 1% antibiotics (Corning #30-004-CI) and kept at 37°C in a humidified incubator containing 5% CO₂. Before using cells for any experiments, they were tested for mycoplasma. Once cells were confirmed negative for any mycoplasma contamination, they were ready to use. The cells were tested at multiple time points for mycoplasma during culture.

Mycoplasma detection

To detect mycoplasma in cells, a previously published direct PCR protocol was followed (128). Each time cells were tested, negative control and positive control were prepared. The negative control sample included sterile water while the positive control sample was a Mycoplasma-contaminated cell line or a standard Mycoplasma organism.

Mycoplasma-specific PCR reaction

Antisense primer MGSO: 5'-TGCACCATCTGTCACTCTGTAAACCTC-3'

Sense primer GPO: 5'-GGGAGCAAACAGGATTAGATACCCT-3'

One reaction was set up combining both primers in a 1:1 ratio. The master mix recipe was as follows: Dream Taq (2X) 5 µL, MGSO primer 1 µL, GPO primer 1 µL, water 2 µL and DNA 1 µL with a total reaction volume of 10 µL (including DNA) per sample. 9 µL of master mix was added to each PCR tube and 1 µL of DNA was added to each sample. The amplification conditions for *Mycoplasma* specific primers consisted of one cycle at 95°C for 5 min, followed by 40 cycles at 94°C for 30 sec, 55°C for 30 sec, and 72°C for 1 min. Then, 72°C for 10 min and 4°C hold until samples were taken out of the machine. Samples were run on a 2.0% agarose gel along with negative and positive controls. If the sample was negative, no band would be present, but if a sample was positive, a 270-base pair fragment would be present.

Flow cytometry

Flow cytometry analysis was performed on the BD Symphony A5 Analyzer (5 lasers 28 parameters) or BD LSRII Analyzer (4 lasers 17 parameters) and analyzed using FlowJo v10 software. Amir Yarmahmoodi, manager of the flow cytometry core facility at the Lewis Katz School of Medicine at Temple University, assisted with flow cytometry analysis and interpretation of data.

Immunophenotyping: mouse peripheral blood

About 100-200 μL of murine peripheral blood was added to 5 mL of ACK lysis buffer in a 15 mL tube. After the addition of ACK, cells were placed on ice for about 5 min or until the sample was not cloudy anymore, so proper RBC lysis could occur. Cells were then spun down at approximately 1200 rpm for 5 min. All ACK and lysed RBCs were aspirated from the tube and the remaining cells were washed 2x with PBS and spun down again at 1200 rpm for 5 min. After the second wash, cells were ready to be stained and were transferred to pre-labeled 5 mL polypropylene tubes. Cells were then resuspended in 90 μL 1% BSA (Sigma-Aldrich Cat#: A9576) in PBS + 10 μL of Fc blocker (Invitrogen Cat#: 14-0161-86) for 10 min to prevent any non-specific binding. Cells were spun down and the supernatant was aspirated. A master mix was made using 1% BSA in PBS + 4 antibodies for cell surface markers Gr-1 (BD Biosciences Cat#: 557405), Mac-1 (BD Biosciences Cat#: 553311), B220 (BD Biosciences Cat#: 553092) and CD3 (BD Biosciences Cat#: 552774). To make the master mix, 500 μL of 1% BSA in PBS was needed per sample. Anti-Gr-1 and anti-Mac-1 antibodies were added to the master mix at a concentration of 10 μL per sample. Anti-B220 and anti-CD3 antibodies were added to the master mix at a concentration of 5 μL per sample. Cells were resuspended in 500 μL of the master mix. 500 μL of 1% BSA in PBS without any antibodies present was added to the negative control sample. A single stained control was also prepared for each antibody to set up

correct compensation on the flow machine. Single stained control samples received 500 μ L 1% BSA in PBS + 1 antibody. Samples were incubated for about 30 min at 4°C while protected from the light. After the incubation, samples were washed with 2-3 mL of PBS, and they were centrifuged for 5 min at 1200 rpm to remove supernatant. Samples were then fixed in 200 μ L 2% PFA and were subsequently analyzed.

Immunophenotyping: mBMCs

The total mBMC population was depleted of RBCs by using 5 mL of ACK lysis buffer. Cells were placed on ice for 5 min and spun down at 1000 rpm for 5 min. Cells were then washed in PBS and spun down again. Cells were resuspended in 90 μ L 1% BSA in PBS + 10 μ L Fc blocker for 10 min. A master mix was made containing 1% BSA in PBS + 3 antibodies for cell surface markers Lin- cocktail (BD Biosciences Cat#: 558074), cKit (BD Biosciences Cat#: 553355) and Sca-1 (BD Biosciences Cat#: 558162). To make the master mix, 500 μ L of 1% BSA in PBS was needed per sample. The Lin- antibody cocktail was added to the master mix at a concentration of 10 μ L per sample while the anti-cKit antibody was 5 μ L per sample and anti-Sca-1 was 2.5 μ L per sample. 500 μ L of 1% BSA in PBS without any antibodies present was added to the negative control sample. A single stained control was also made for each antibody to set up correct compensation on the flow machine. Single stained control samples received 500 μ L 1% BSA in PBS + 1 antibody. Cells were resuspended in 500 μ L of the master mix and incubated for 30 min at 4°C away from the light. Cells were washed using 2-3 mL of PBS, resuspended in 200 μ L 2% PFA and analyzed.

γ H2AX and 8-Oxo-G analysis

Cells were first fixed in 2% PFA for 15-30 min on ice and then washed 1x with PBS. If samples were not processed right away, they were left in PBS and placed in 4°C overnight. After washing in PBS, cells were centrifuged at 1000 rpm for 5 min and

supernatant was removed. Cells were resuspended in 100 μ L PBS + 1 mL of pre-chilled 100% methanol and incubated on ice for 15-30 min. After incubation, cells were spun down, supernatant was discarded, and cells were washed with 0.02% Triton-X in PBS. After washing, a master mix was made using 100 μ L of 1% BSA + 0.02% Triton-X in PBS per sample as well as the anti- γ H2AX (BD Biosciences Cat#: 560447) or anti-8-Oxo-G (Abcam Cat#: ab183393) antibodies which were added at a concentration of 5 μ L per sample (less was added if there was less than 1×10^6 cells/sample). For negative control, one sample had 1% BSA + 0.02% Triton-X in PBS added to it with no antibody present. For a positive control in γ H2AX analysis, mBMCs were treated with doxorubicin (SelleckChem Cat#: S1208) for approximately 2 hr and then processed simultaneously with all other samples. Master mix was added to each sample and incubated for 30 min at room temperature in the dark. After incubation, cells were washed with 2-3 mL of PBS, spun down, supernatant was removed and finally resuspended in 200 μ L 0.02% Triton-X in PBS. Cells were analyzed by flow cytometry and FlowJo software was used to determine the geometric mean of each sample.

Apoptosis

Cells were prepared and analyzed using the FITC Annexin Apoptosis Detection Kit II (BD Biosciences Cat#: 556570) to assess changes in viability and to track mechanism of cell death. Unstained cells were used as a negative control while cells stained with either FITC Annexin V (no propidium iodide (PI)) or PI (no FITC Annexin V) were used as compensation controls. Cells were washed twice with cold PBS and resuspended in 1X Binding Buffer at a concentration of 1×10^6 cells/mL. 100 μ L of the solution was transferred to a 5 mL tube and 5 μ L of FITC Annexin V and 5 μ L of PI were added. Cells were gently vortexed and incubated for 15 min at room temperature in the dark. After, 400 μ L of 1X Binding Buffer was added to each tube and analyzed via flow cytometry within 1 hr.

Human CD19 detection

To detect the development of leukemia in SCID-beige mice injected with Nalm6-*RAD54*^{-/-} cells, human CD19 positive cells were identified in the peripheral blood. About 100 μ L of blood was added to 5 mL of ACK lysis buffer in a 15 mL tube. After the addition of ACK, cells were placed on ice for about 5 min or until the sample was not cloudy anymore. Cells were then spun down at approximately 1200 rpm for 5 min. All ACK and lysed RBCs were aspirated from the tube and the remaining cells were washed 2x with PBS and spun down again at 1200 rpm for 5 min. After the second wash, cells were ready to be stained and were transferred to pre-labeled 5 mL polypropylene tubes. Cells were then resuspended in 90 μ L 1% BSA in PBS + 10 μ L of Fc blocker for 10 min to prevent any non-specific binding. Cells were spun down and the supernatant was aspirated. Nalm6-*RAD54*^{-/-} that were being cultured were used as a positive control while the negative control did not receive any antibody. A master mix was made using 1% BSA in PBS + anti-human CD19 antibody (BD Biosciences Cat#: 555413). To make the master mix, 100 μ L of 1% BSA in PBS was needed per sample. The anti-CD19 antibody was added to the master mix at a concentration of 10 μ L per sample. Cells were resuspended in 100 μ L of the master mix. Samples were incubated for about 30 min at 4°C while protected from the light. After the incubation, samples were washed with 2-3 mL of PBS, and they were centrifuged for 5 min at 1200 rpm to remove supernatant. Samples were then fixed in 200 μ L 2% PFA and were subsequently analyzed.

Western blots

Whole cell lysates were made using PierceTM IP Lysis Buffer (Thermo Fisher Scientific Cat#: 87788) and Halt Protease & Phosphatase Inhibitor Cocktail (100x) (Thermo Fisher Scientific Cat#: 78446). After centrifugation, protein concentrations were measured using Bradford assay (Bio-Rad Laboratories Cat#: 5000006). All samples were

prepared using Protein Assay Dye Reagent Concentrate (Bradford reagent) and measured using a spectrophotometer. 100 µg of protein lysates were separated using SurePAGE™ Bis-Tris 4-12% pre-cast protein gel (GenScript Cat#: M00652) and Tris-MOPS-SDS Running Buffer (GenScript Cat#: M00138). Spectra™ Multicolor Broad Range Protein Ladder (Thermo Fisher Scientific Cat#: 26634) was run on the gel alongside protein lysates. Proteins were transferred onto a nitrocellulose membrane for about 2 hr at 250 mA using transfer buffer (800 mL water, 200 mL 100% methanol, 14.4 g glycine and 3 g Tris base). Membrane was blocked using 5% BSA in PBST (100 mL 10X PBS, 900 mL water, 1 mL Tween 20) for 45 min at room temperature on a rocker. Membrane was probed overnight with 1:1000 anti-Polθ (MyBioSource Cat#: MBS9612322), anti-ADH5 (Abcam Cat#: ab177932), anti-ALDH2 (Invitrogen Cat#: MA5-17029), and anti-actin (Invitrogen Cat#: MA5-11769) antibodies. For γH2AX western blot, membrane was probed overnight with 1:2000 anti-γH2AX (phosphor S139) (Abcam Cat#: ab243906). Donkey anti-rabbit (LICORbio™ Cat#: 926-68073) and goat anti-mouse (LICORbio™ Cat#: 926-32210) secondary antibodies were diluted to 1:3000 with 5% BSA in PBST for 1.5 hr at 4°C on a rocker protected from the light. Protein bands were detected using an iBright machine.

γH2AX immunofluorescence (IF)

Cells were processed and analyzed for IF as described before (129). After 72 h of ART558 treatment, cells were fixed with 4% (v/v) PFA for 20 min at 4°C, washed two times with PBS, followed by permeabilized with 0.5% (v/v) Triton X-100 for 10 min and blocked with PBS containing 3% BSA for 1 hr. Cells were then incubated with the same buffer containing primary antibody anti-γH2AX (Abcam Cat#: ab243906) at 1:100 overnight at 4°C followed by 1 hr incubation with secondary antibody anti-mouse Alexa Fluor 594 (Life Technologies Cat#: A11062). The incubations were performed in the dark. 5x washing in

PBST for 3 min after each antibody incubation was performed. Slides were mounted in 20 μ L mounting media. Cells were visualized and imaged using a Leica SP8 Confocal microscope at a 63X objective magnification and images were analyzed using ImageJ software. For quantification, > 30 cells were counted for all conditions from three independent experiments. Intensity of individual protein was analyzed in every nuclei (N=30) by measuring their respective foci formation [as intensity (a.u)] using ImageJ.

Detection of DPCs

Detection of DPCs was performed using the ARK assay as described before (124, 130). First, all buffers were prepared prior to the start of the detection of DPCs. All buffer recipes were prepared per 50 mL and Milli Q Deionized Water was used for all buffers.

Buffer preparation

MB Buffer: recipe consisted of 5.6M guanidine thiocyanate (GTC), 10 mM Tris-HCl (pH 6.5), 20 mM EDTA, 4% Triton X-100, 1% Sarkosyl, 1% DTT and water. First, water was placed in a beaker with a stir bar and heated up using a heating plate. About 33.05 g of GTC was added and dissolved in the heated water. Then, 0.5 mL of 1 M Tris-HCl (pH 6.5), 2 mL of 500 mM EDTA, 2 mL of Triton X-100, 0.5 g of Sarkosyl and 0.5 g DTT were added to the mixture in that sequential order. Finally, water was added until the final volume of 100 mL was reached. After everything was completely dissolved in water, the buffer was placed in 4°C refrigerator for long-term storage.

Buffer 2: recipe consisted of 20 mM Tris-HCl (pH 6.5), 150 mM NaCl, 50% ethanol and water. In a 50 mL tube, 1 mL of 1 M Tris-HCl (pH 6.5), 2.5 mL of 3 M NaCl, and 25 mL of 100% ethanol were added together and then topped off with water to reach 50 mL. Buffers 2-5 and the Proteinase K buffer were all stored at room temperature.

Buffer 3: recipe consisted of 1% SDS, 20 mM Tris-HCl (pH 7.5), and water. In a 50 mL tube, 5 mL of 10% SDS and 1 mL of 1 M Tris-HCl was added and then topped off with water to reach 50 mL.

Buffer 4: recipe consisted of 200 mM KCl, 200 mM Tris-HCl (pH 7.5) and water. In a 50 mL tube, 5 mL of 2 M KCl and 1 mL of 1 M Tris-HCl (pH 7.5) were added and topped off with water to reach 50 mL.

Buffer 5: recipe consisted of 100 mM KCl, 20 mM Tris-HCl (pH 7.5) and water. In a 50 mL tube, 2.5 mL of 2 M KCl and 1 mL of 1 M Tris-HCl (pH 7.5) were added and topped off with water to reach 50 mL.

Proteinase K Buffer: recipe consisted of 100 mM KCl, 20 mM Tris-HCl (pH 7.5), 10 mM EDTA and water. In a 50 mL tube, 2.5 mL of 2 M KCl, 1 mL of 1 M Tris-HCl (pH 7.5), 1 mL of 500 mM EDTA were added and topped off with water to reach 50 mL volume.

Part I: isolation of DPC-associated DNA and free DNA

MB buffer was pre-warmed at 55°C. A white precipitate formed when the buffer was kept at 4°C, so the buffer was not used until the precipitate was completely dissolved. At the end of each experiment, cells were counted, and it was preferable to have at least 1×10^6 cells/sample, but the DPC assay was also performed using less cells. Cells were transferred to 2 mL tubes, centrifuged at 1000 rpm for 5 min and the supernatant was removed. 950 μ L of pre-warmed MB buffer was added to each sample by pipetting up and down 6 times to lyse the cells. Lysates were sheared by passing through a 22G syringe 6 times. One syringe was used per sample. 950 μ L of pre-chilled ethanol (stored in -20°C) was added to each sample to precipitate DNA. It was normal to not see a pellet during this step, especially if less cells were used. Samples were then centrifuged at 4°C for approximately 20 min at max speed (14,000 rpm). The supernatant was carefully removed

as a pellet was not always seen. Samples were then washed with 950 μ L of Buffer 2 and centrifuged again at 4°C for approximately 20 min at max speed (14,000 rpm). Supernatant was carefully removed, and samples were used for the next steps or stored in -80°C until ready to proceed with the protocol.

Buffer 3 was pre-warmed, and a heating block was also preset to 42°C for the next steps. 500 μ L of pre-warmed Buffer 3 was added to each sample and then placed in the 42°C heating block for 6 min. The samples were briefly spun and sheared by passing through a 25G syringe 5 times. The samples were centrifuged at full speed for about 45 sec. 500 μ L of Buffer 4 was added to each sample, inverted a few times and placed on ice for 6 min. Samples were then centrifuged at 4°C for 5 min at max speed (14,000 rpm). 15 mL tubes were labeled in the meantime for the collection of free DNA. The supernatant was collected into the pre-labeled 15 mL tubes as it contained the free DNA. 1.5 mL of Buffer 5 was added to each sample, incubated at 55°C for 10 min, then placed on ice for 6 min and centrifuged at 4°C for 5 min at max speed (14,000 rpm). Supernatant was collected and added to 15 mL tubes. This step with Buffer 5 was repeated once more. After repeating this step, there was a total of 4 mL of DNA that was collected in each 15 mL tube. When the free DNA was collected, the next steps were to collect the DPC-associated DNA. Still with the 2 mL tubes, 1 mL of Proteinase K Buffer was added followed by proteinase K (Thermo Fisher Scientific Cat#: EO0492) which was added at a final concentration of 0.2 mg/mL. The samples were then placed in 55°C heating block and incubated for 45 min. After, samples were cooled on ice for 6 min and centrifuged at 4°C for 10 min at max speed (14,000 rpm). In the meantime, 1.5 mL tubes were labelled. After centrifuging, the 1 mL supernatant was added into the newly labelled tubes. This was the DPC-associated DNA. The DPC-associated DNA samples were stored at 4°C (short-term) if not able to immediately continue to the second part of the protocol.

Part II: quantification of DNA by pico green dsDNA assay

To quantify the amount of DPCs in each sample, a DPC/DNA coefficient was determined. First, 1.5 mL tubes were all labeled accordingly and 1X TE Buffer was prepared from 20X TE (Invitrogen Cat#: P7589). For samples containing free DNA, 115 μ L of 1X TE was added to each newly labeled tube. 10 μ L (from the 4 mL recovered) of free DNA was added to each tube, so the total volume per sample was 125 μ L. For the DPC-associated samples, 62.5 μ L of 1X TE was added to each newly labeled tube, followed by 62.5 μ L (from the 1 mL recovered) of the DPC-associated DNA to get a total volume of 125 μ L. All samples were mixed and pipetted into a 96-well black clear bottom plate. Pico Green (Invitrogen Cat#: P7589) was pre-warmed, diluted 1:200 in 1X TE Buffer and kept in the dark. After the Pico was added to the 1X TE Buffer, it was vortexed and 125 μ L was added to each sample. The plate was covered in aluminum foil and brought to the plate reader and fluorescence was measured with excitation and emission wavelengths 485 nm and 528 nm, respectively. Analysis was done in Microsoft Excel.

Detection of DSBs

For the detection of DSBs, 2 methods were utilized for this work: flow cytometry to detect γ H2AX or neutral comet assay. For neutral comet assay, the OxiSelect™ Comet Assay Kit (3-well slides) was purchased (Cell Biolabs Cat# STA351). The assay was performed according to the manufacturer's protocol. For this assay, all buffers were prepared prior to the start of this protocol.

Reagent preparation

1X Lysis Buffer: recipe consisted of 14.6 g NaCl, 20 mL EDTA solution, 10 mL of 10X lysis and water. The solution was mixed thoroughly, and the water volume was adjusted to approximately 90 mL. The pH was slowly adjusted to 10.0 using 10 N NaOH with water added to adjust to a final volume of 100 mL. The lysis buffer was stored in 4°C.

Alkaline Solution (100 mL): recipe consisted of 1.2 g NaOH, 0.2 mL EDTA solution and water. Water was adjusted to reach a volume of 100 mL.

TAE Buffer: For the electrophoresis solution, 1X TAE was chosen. 20 mL of 50X TAE was added to 980 mL of water to get 1X TAE buffer.

Preparation of samples and slides

For neutral comet assay experiments, suspension cells were used. Cells were collected and centrifuged at 1000 rpm for 5 min and the supernatant was discarded. Cell pellets were washed once using chilled PBS, centrifuged and supernatant was discarded. Cells were resuspended at 1×10^6 cells/mL in chilled PBS. The OxiSelect™ Comet Agarose was heated to 95°C in a water bath until the agarose liquified. Each cell sample was combined with the Comet Agarose at a 1:10 ratio (v/v), mixed well by pipetting and immediately transferred 75-85 µL onto the 3-well slides. The suspension was spread over the well with the pipette tip to ensure complete coverage. The slides were then transferred to 4°C in the dark for 15 min. To ensure a dark environment, cells were placed in black containers. Slides were then transferred to a small container that contained pre-chilled 1X lysis buffer. Slides were immersed in the buffer for 1 hr. Enough buffer was used to completely cover the slides. Carefully, the lysis buffer was aspirated and replaced with pre-chilled alkaline solution and the slides were immersed for 30 min at 4°C in the dark.

Assay protocol

The alkaline solution was aspirated from the containers and replaced with pre-chilled 1X TAE Electrophoresis Solution. The slides were immersed for 5 min and then repeated once more. The slides were carefully maintained horizontally and were transferred to an electrophoresis chamber. The electrophoresis chamber was placed in a large bucket/container that was filled with ice to maintain cold conditions. The chamber

was filled with 1X TAE until the buffer just covered the slides. Voltage was applied to the chamber for 20 min at 1 volt/cm. After, the slides were carefully transferred to a clean container containing pre-chilled deionized water and immersed for 2 min. Water was aspirated and the process was repeated 2 more times. After the final rinse with water, it was aspirated and replaced with cold 70% ethanol for 5 min. The slides were then removed from the ethanol and left to air dry for about 30 min. If there was not enough time to do imaging, slides were placed at 4°C in dark containers until ready to use. After air drying, the Vista Green DNA dye was prepared. 1X Vista Green Dye was prepared by diluting the provided stock 1:10000 in 1X TE Buffer. The solution remained protected from the light. 100 µL of dye was added to each well and incubated at room temperature for 15 min in the dark. Slides were then taken to the microscope room and imaged using an inverted Olympus IX70 fluorescence microscope using a FITC filter. The percentage of tail DNA of individual cells was calculated using the OpenComet plugin of ImageJ and represented by percentage of tail DNA \pm standard error of mean (SEM). At least 25 images were taken per sample.

Clonogenic assay

The colony forming assay is a technique that measures the ability of a single cell to grow into a colony. For all clonogenic assays done, suspension cells were used. Methocult™ H4230 Methylcellulose-Based Medium is a semi-solid medium used for all colony forming unit (CFU) assays (StemCell Technologies Cat#: 04230). 1% antibiotic was added to all newly opened methylcellulose bottles. To begin, 14 mL round bottom polystyrene tubes were all pre-labeled (Corning Cat# 352057) and methylcellulose was pre-warmed at 37°C. Using a 5 mL syringe with a blunt end needle, 3.2 mL of methylcellulose was added to each tube. The extra 0.2 mL accounted for some methylcellulose that stuck to the tubes. Cells were prepared by mixing all three triplicate

wells together to get a final volume of 300 μ L (all experiments were performed in triplicate) and were added to the 14 mL tubes. The final volume of media added to the 14 mL tubes was 300 μ L (100 μ L of media containing cells x3 per sample). After the cell mixture was added to methylcellulose, tubes were vortexed vigorously and then set aside until all bubbles disappeared. Next, the contents of the tube were aspirated using a 5 mL syringe with a blunt end needle attached. The syringe was slowly filled and sat for a couple minutes to let the methylcellulose stuck to the sides slide back down. The syringe was filled enough to get 3 mL of the methylcellulose and cell mixture. 1 mL of the methylcellulose and cell mixture was then dispersed on a 35 mm grided plate (SARSTEDT Cat#: 83.3900.002). Plates were rotated while the methylcellulose was added to ensure even coating. Plates were stored inside 150 mm plates that each additionally contained two 35 mm plates filled with PBS to prevent methylcellulose from drying. Plates were placed in the incubator and counted after 7-10 days. Colonies (CFUs) were counted by eye using an inverted light microscope.

Formaldehyde detection

Formaldehyde detection in murine peripheral blood and bone marrow samples was done using a formaldehyde detection kit (Abcam Cat#: ab196997, ab272524). Assay was performed according to the manufacturer's instructions with no modifications.

***In vitro* formaldehyde challenge**

100 μ L of wild-type, *Polq*^{-/-}, *Adh5*^{-/-}, *Aldh2*^{-/-}, *Adh5*^{-/-};*Polq*^{-/-} and *Aldh2*^{-/-};*Polq*^{-/-} Lin-cKit+ mBMCs were seeded in a 96-well plate at 25x10³ cells/well. All experiments were performed in triplicate. Cells were treated with various doses of formaldehyde (Thermo Fisher Scientific Cat#: 119690010) ranging from 5 μ M to 50 μ M for 24 hr. After 24 hr, cells were counted and plated in methylcellulose for clonogenic assay and colonies were counted after 7-10 days. For DPC and γ H2AX analysis, Lin-cKit+ mBMCs were treated

with formaldehyde for 12 hr and cells were subsequently collected for analysis. Wild-type, *Polq*^{-/-}, *Adh5*^{-/-}, and *Adh5*^{-/-};*Polq*^{-/-} Lin-cKit⁺ mBMCs were treated with 20 μM formaldehyde while wild-type, *Polq*^{-/-}, *Aldh2*^{-/-}, and *Aldh2*^{-/-};*Polq*^{-/-} Lin-cKit⁺ mBMCs were treated with 50 μM formaldehyde. These respective doses were selected based on the formaldehyde sensitivity from the initial clonogenic assay described above.

***In vivo* methanol challenge**

The protocol used for this experiment was adapted from a previously published protocol with some modifications (28). Wild-type, *Polq*^{-/-}, *Adh5*^{-/-}, *Aldh2*^{-/-}, *Adh5*^{-/-};*Polq*^{-/-} and *Aldh2*^{-/-};*Polq*^{-/-} mice were used for this experiment. 5 mice/genotype aged between 7-12-weeks-old were selected, and their cages were moved from MERB 541 housing to MERB 545 special housing since they were to be injected with a biohazardous agent. Once the mice were moved to MERB 545, mice were all weighed prior to the start of the experiment. 100% methanol (Thermo Fisher Scientific Cat#: A452SK-4) was diluted to 10% methanol (v/v) in PBS and injected intraperitoneally at 12.6 mL/kg. The first dose was given on day 1 and the second dose was given on day 8. Mice were checked every day and weighed the same days as they were given their intraperitoneal injections of methanol. On day 10, mice were euthanized by CO₂ exposure followed by cervical dislocation. Blood and bone marrow were collected for analysis of hematopoiesis.

Drug treatments

Disulfiram/ALDH_i (SelleckChem Cat# S1680), N6022/ADH_{5i} (SelleckChem Cat#: S7589), ART558/Pol_{θi} (SelleckChem Cat#: S9936), RP-6685/Pol_{θi} (SelleckChem Cat#: E1528), ruxolitinib (SelleckChem Cat#: S1378), and quizartinib (SelleckChem Cat#: S1526) were inhibitors used for all *in vitro* and *in vivo* experiments. Disulfiram is an ALDH inhibitor that is an FDA approved treatment option for people suffering from alcoholism. N6022 is an ADH₅ inhibitor that has been studied in the treatment of pulmonary diseases

such as asthma. ART558 and RP-6685 are both Polθ inhibitors with ART558 used for all *in vitro* experiments while RP-6685 was used for *in vivo* experiments. Ruxolitinib is an FDA approved JAK1/2 kinase inhibitor that competes with ATP for binding to the catalytic site of JAK1 and JAK2 and is used to treat patients with JAK mutations. Quizartinib is an FDA approved FLT3 inhibitor that binds to the ATP binding domain of the FLT3 receptor and is used to treat patients with FLT3 mutations.

mBMC drug treatment

100 μL of wild-type and *Polq*^{-/-} Lin-cKit⁺ mBMCs were seeded in 6-well plates at 1x10⁶ cells/well. Cells were treated with different doses of either ALDHi (0, 1.25, 2.5 or 5 μM) or ADH5i (0, 10, 25, 50, 100 μM) for 48 hr and were then collected for analysis. Wild-type, *Adh5*^{-/-} and *Aldh2*^{-/-} Lin-cKit⁺ BMCs were treated with different doses of ART558 (0, 12.5, 25 μM) for 72 hr.

***Nalm6 parental and RAD54*^{-/-} treatment**

For all *in vitro* experiments, 100 μL of Nalm6 parental and Nalm6-*RAD54*^{-/-} were seeded in 96-well plates at 10x10³ cells/well. Cells were first treated with various concentrations of ADH5i (0, 10, 20, 40 μM) or ALDHi (0, 0.125, 0.250 μM) for 48 hr to determine the optimal dose to use in combination with Polθi. Then, Nalm6 parental and Nalm6-*RAD54*^{-/-} were treated with either Polθi alone (12.5 μM) or Polθi (12.5 μM) in combination with ADH5i (20 μM) or ALDHi (0.125 μM) for 48 hr. After the drug treatment, cells were collected for analysis. For all *in vitro* experiments, one-way ANOVA and Tukey's test for multiple comparisons was used to determine statistical significance between groups with *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001, and ****p ≤ 0.0001.

Patient cell treatment

AML201, AML1, MPN1, and P293 patient cells were thawed 24 hours prior to the start of all experiments. Cells were first counted to make sure viability was between 80-100%. 100 μ L of patient cells were seeded in 96-well plates at 25×10^3 cells/well. All experiments were performed in triplicate. AML 201 and AML1 patient cells were treated with Pol θ i (25 μ M), ALDH i (0.075 μ M), ADH5 i (75 μ M), FLT3 i (0.125 μ M) alone or in combination (Pol θ i 25 μ M + ADH5 i 75 μ M \pm FLT3 i 0.125 μ M and Pol θ i 25 μ M + ALDH i 0.075 μ M \pm FLT3 i 0.125 μ M) for 6 days. MPN1 patient cells were treated with Pol θ i (25 μ M), ALDH i (0.05 μ M), ADH5 i (50 μ M), JAK1/2 i (0.1 μ M) alone or in combination (Pol θ i 25 μ M + ADH5 i 50 μ M \pm JAK1/2 i 0.1 μ M and Pol θ i 25 μ M + ALDH i 0.05 μ M \pm JAK1/2 i 0.1 μ M) for 6 days. MPN P293 patient cells treated with Pol θ i (25 μ M), ALDH i (0.075 μ M), ADH5 i (75 μ M), JAK1/2 i (0.1 μ M) alone or in combination (Pol θ i 25 μ M + ADH5 i 75 μ M \pm JAK1/2 i 0.1 μ M and Pol θ i 25 μ M + ALDH i 0.075 μ M \pm JAK1/2 i 0.1 μ M) for 6 days. On day 3 of treatment, cells were given a second dose of drug(s). After 6 days, cells were counted using Trypan blue staining. The mean number of cells \pm SEM was calculated and one-way ANOVA and Tukey's test for multiple comparisons was used to determine statistical significance between groups with * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, and **** $p \leq 0.0001$.

ADH activity assay

1 mL of wild-type and *Pol θ ^{-/-}* Lin-cKit⁺ mBMCs were seeded in 6-well plates at 1×10^6 cells/mL. Cells were treated with 0, 10, 25, 50 and 100 μ M of ADH5 i for 48 hr. After treatment, the Alcohol Dehydrogenase Activity Colorimetric Assay Kit (BioVision Cat#: K787) was performed to measure ADH activity.

ALDH activity assay

1 mL of wild-type and *Polg*^{-/-} Lin-cKit⁺ mBMCs were seeded in 6-well plates at 1x10⁶ cells/mL. Cells were treated with 0, 1.25, 2.5, and 5 μM of ALDH_i for 48 hr. After treatment, the ALDEFLUOR™ ALDH Detection Kit (StemCell Technologies Cat#: 01700) was performed to measure ALDH activity.

Xenograft model

Prior to injection, Nalm6-*RAD54*^{-/-} cells were kept in cell culture and confirmed to be mycoplasma negative. Nalm6-*RAD54*^{-/-} cells were prepared at a concentration of 2x10⁵ cells/mL in PBS. 2x10⁴ Nalm6-*RAD54*^{-/-} human leukemia cells (in 100 μL) were injected into SCID Beige mice (Charles River, strain 250) via tail vein. Mice were then randomly assigned to 6 groups: control, disulfiram (ALDH_i), N6022 (ADH_{5i}), RP-6685 (Polθ_i), disulfiram (ALDH_i) + RP-6685 (Polθ_i) and N6022 (ADH_{5i}) + RP-6685 (Polθ_i). After 7 days, mice were treated via oral gavage daily for 14 days. In the groups where two treatments were combined, the drugs were administered at least 6 hr apart. Both N6022 and disulfiram were dissolved in DMSO and further diluted using 10% Kolliphor HS 15 (Millipore Cat#: 42966) and were administered to mice at 50 mg/kg/day. RP-6685 was first dissolved in DMSO and then in sterile water + 5% 1-methyl-2-pyrrolidinone (Millipore Cat# 443778) + 10% D-α-tocopherol polyethylene glycol 1000 succinate (Millipore Cat#: 57668). RP-6685 was administered at 80 mg/kg/day. On day 21, the last drug doses were given to the mice. 48 hr after the last doses, blood samples were collected via tail nick and an anti-human CD19 antibody was used to detect the leukemia burden by flow cytometry. Mice were monitored daily and euthanized once they became sick. Paralysis was a common symptom/phenotype observed in these mice that led to euthanasia. The date of death for all mice was recorded, so median survival time (MST) could be determined.

Virus construction and cell infection

Phoenix ECO cells were cultured in DMEM medium using 10 cm plates. The day before transfection, cells were split, so that the cells were about 80% confluent on the day of the viral infection. 10 µg of the FLT3(ITD)-GFP or JAK2(V617F)-GFP plasmid and 5 µg pCL-Eco (Addgene Cat#: 12371) were co-transfected into Phoenix-ECO cells (ATCC Cat#: CRL-3214) in DMEM culture with 30 µL Lipofectamin 2000 Transfection Reagent (Thermo Fisher Scientific Cat#: 1168019). Media containing the virus was collected 24 and 48 hr after transfection. Wild-type, *Polq*^{-/-}, *Adh5*^{-/-}, *Aldh2*^{-/-}, *Adh5*^{-/-};*Polq*^{-/-} and *Aldh2*^{-/-};*Polq*^{-/-} were spin-infected (1800 rpm for 90 min at 32°C) by retrovirus medium containing polybrene (10 µg/mL) (Sigma-Aldrich Cat#: TR-1003). After, cells were left in the incubator at 37°C overnight. The following day, supernatant was removed and fresh mBMC media was added. After 3-4 days, cells were selected for GFP+ using BD FACSAria II cell sorter. GFP+ cells were kept in cell culture for subsequent experiments.

Sensitivity to hydrogen peroxide (H₂O₂)

100 µL of wild-type and *Polq*^{-/-} Lin-cKit+ mBMCs were seeded in 96-well plates at 2x10⁴ cells/well. A 1 mM stock of H₂O₂ was prepared by diluting the H₂O₂ (Thermo Fisher Scientific Cat#: H325-500) in PBS. Cells were treated with 0, 25, 50, 100 or 200 µM of hydrogen peroxide for 4 hr. Cells were then collected for analysis. For all H₂O₂ experiments, statistical analysis was performed using one-way ANOVA and Tukey's test for multiple comparisons when comparing different doses within each genotype and unpaired t-tests were also performed when comparing *wt* vs. *Polq*^{-/-} for a specific dose with *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001, and ****p ≤ 0.0001.

H2DCFDA assay

2',7'-Dichlorodihydrofluorescein diacetate (H2DCFDA) (Sigma Aldrich Cat#: D6883) is a cell-permeable probe used to detect intracellular reactive oxygen species (ROS). It is de-esterified intracellularly and turns to highly 2',7'-dichlorofluorescein upon oxidation. 1 mL of wild-type and *Polq*^{-/-} Lin-cKit⁺ mBMCs were plated in 12 well plates at 1x10⁵ cells/well. Cells were treated with 0, 25, 50, 100 and 200 μM of hydrogen peroxide for 4 hr. 50 mg of H2DCFDA powder was dissolved in 10.261 mL of DMSO to make a 10 mM stock. H2DCFDA was further diluted in PBS to make a 1 mM solution. After treatment, H2DCFDA dye was added to each well at a final concentration of 5 μM. Plates were kept in the dark at 37 °C for 30 min. Cells were then spun down, washed in PBS and transferred to 5 mL polypropylene tubes and subsequently analyzed using flow cytometry where geometric mean was determined for each sample.

TMEJ activity assay

1 mL of wild-type, *Adh5*^{-/-} and *Aldh2*^{-/-} Lin- cKit⁺ mBMCs were seeded in 6-well plates at 1x10⁶ cells/mL. Cells were treated with either 0, 12.5 or 25 μM of ART558 for 72 hr. After 24 hr, cells were co-transfected with pBABE-MMEJ plasmid (83) after digestion with I-Sce-I and dsRED-Mito using the Nucleofector™ kit (Lonza Cat#: VPA-1003). The dsRED-Mito was used as a control for transfection efficiency. 48 hr after that, GFP⁺ and Red⁺ cells were detected by flow cytometry analysis. TMEJ was shown as the ratio of GFP⁺/Red⁺ cells. Statistical significance was calculated using one-way ANOVA and Tukey's test for multiple comparisons with *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001, ****p ≤ 0.0001.

GSH sensitivity

1 mL of *Adh5^{-/-};Polq^{-/-}* and *Aldh2^{-/-};Polq^{-/-}* Lin-cKit⁺ BMCs were seeded in 96-well plates at 10x10³ cells/mL. Cells were treated with various doses of formaldehyde (0, 5, 10, 20, 50 μM) ± 50 μM of L-Buthionine sulfoximine (L-BSO) (MedChem Express Cat#: HY-106376A) for 24 hr. Cells were then counted and plated in methylcellulose.

***In vitro* cigarette smoke exposure**

The smoke from one cigarette was drawn into 12.5 mL of mBMC media without FBS using a peristaltic pump. The pump was set at an optimal speed to allow one cigarette to burn in approximately 15 min and the resulting solution was considered 100% CSE. The solution was filtered through a 0.22 μM pore acrodisc syringe filter and applied immediately to cell cultures. 100 μL of wild-type, *Polq^{-/-}*, *Adh5^{-/-}*, *Aldh2^{-/-}*, *Adh5^{-/-};Polq^{-/-}* and *Aldh2^{-/-};Polq^{-/-}* Lin-cKit⁺ BMCs were seeded in 96-well plates at 10x10³ cells/well. Cells were treated with various concentrations (0, 0.125, 0.250, 0.50, 0.75) of CSE-infused media. After 24 hr, cells were counted and plated in methylcellulose. Obtaining CSE was done in collaboration with Beata Kosmider at the Lewis Katz School of Medicine at Temple University and the CSE protocol has been previously described (131).

***In vivo* cigarette smoke exposure**

3-4-month-old wild-type, *Polq^{-/-}*, *Adh5^{-/-}*, and *Aldh2^{-/-}* male mice were used. Mice were exposed to cigarette smoke for 2 hr per day for 3 months using a Teague TE-10 smoking system (Teague Enterprises, Woodland, CA). After the 3-month exposure, blood samples were collected from mice via tail nick. After the mice reached 1.5 years old, they were euthanized, and blood, bone marrow and organs were collected for analysis. Animal care, handling and experimental procedures were carried out in strict accordance with protocols approved by the IACUC of Temple University. Mouse exposure to cigarette

smoke was done in collaboration with Beata Kosmider at the Lewis Katz School of Medicine at Temple University and the protocol has been previously described (132).

CHAPTER 3

RESULTS

Polθ does not cooperate with ADH5 or ALDH2 to protect hematopoietic cells from physiological levels of formaldehyde

To determine whether ADH5 and/or ALDH2 cooperate with Polθ in protecting hematopoietic cells from physiological levels of formaldehyde and formaldehyde-mediated DNA damage, *Adh5*^{-/-};*Polq*^{-/-} and *Aldh2*^{-/-};*Polq*^{-/-} mice were generated (Figure 8, Appendix Figure 1). *Adh5*^{-/-};*Polq*^{-/-} and *Aldh2*^{-/-};*Polq*^{-/-} were viable and displayed no detectable growth abnormalities or adverse phenotypes. Additionally, no organ dysfunctions were detected in aged mice, all indications of a normal lifespan of *Adh5*^{-/-};*Polq*^{-/-} and *Aldh2*^{-/-};*Polq*^{-/-} double knockout mice (Figure 9). *Adh5*^{-/-};*Aldh2*^{-/-};*Polq*^{-/-} mice could not be generated because most *Adh5*^{-/-};*Aldh2*^{-/-} mice die perinatally (28).

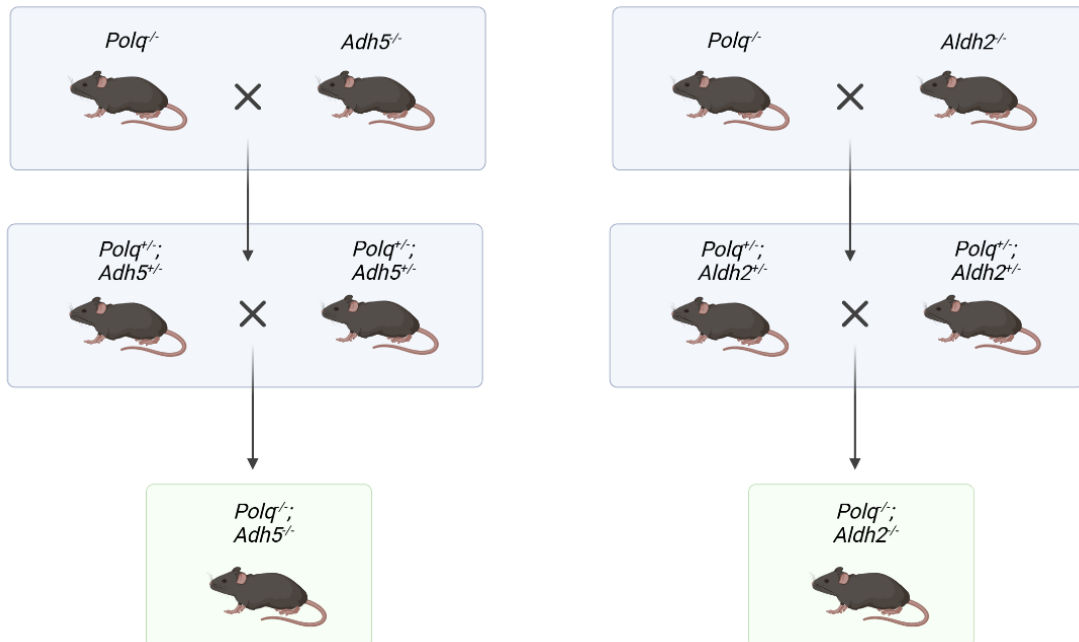


Figure 8. Breeding protocol to obtain double knockout mice.

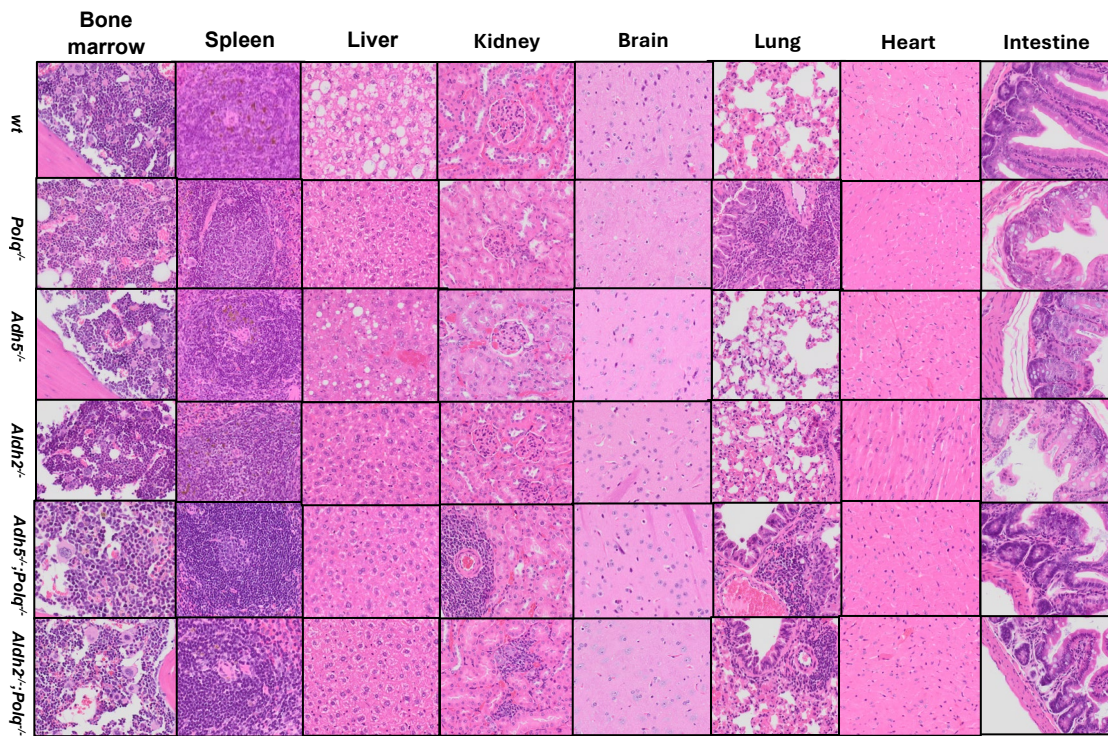


Figure 9. Analysis of organs from aged genetic knockout mice. H&E-staining performed at Fox Chase Cancer Center on tissue sections from the bone marrow, spleen, liver, kidney, brain, lung, heart, and bowel harvested from the aged mice with the indicated genotypes (40x). Aged mice were 18-24 months-old at the time of organ harvesting.

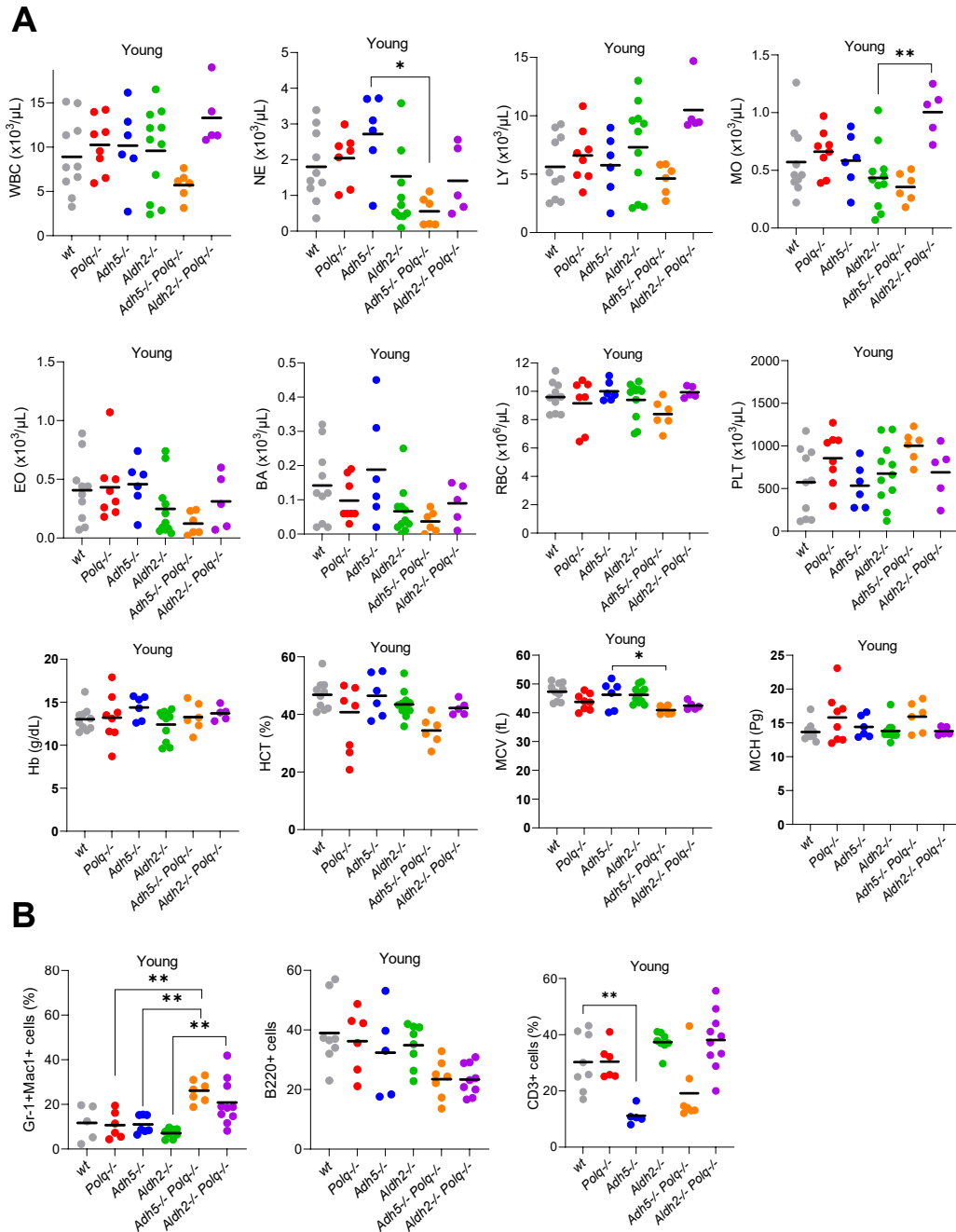


Figure 10. Analysis of peripheral blood from young (3-4-months-old) mice. (A) Hemavet analysis of peripheral blood cell parameters: white blood cells (WBC), neutrophils (NE), lymphocytes (LY), monocytes (MO), eosinophils (EO), basophils (BA), red blood cells (RBC), platelets (PLT), hemoglobin (Hb), hematocrit (HCT), mean corpuscular volume (MCV), and mean corpuscular hemoglobin (MCH). (B) Gr-1+Mac-1+, B220+ and CD3+ cells were detected in peripheral blood by flow cytometry.

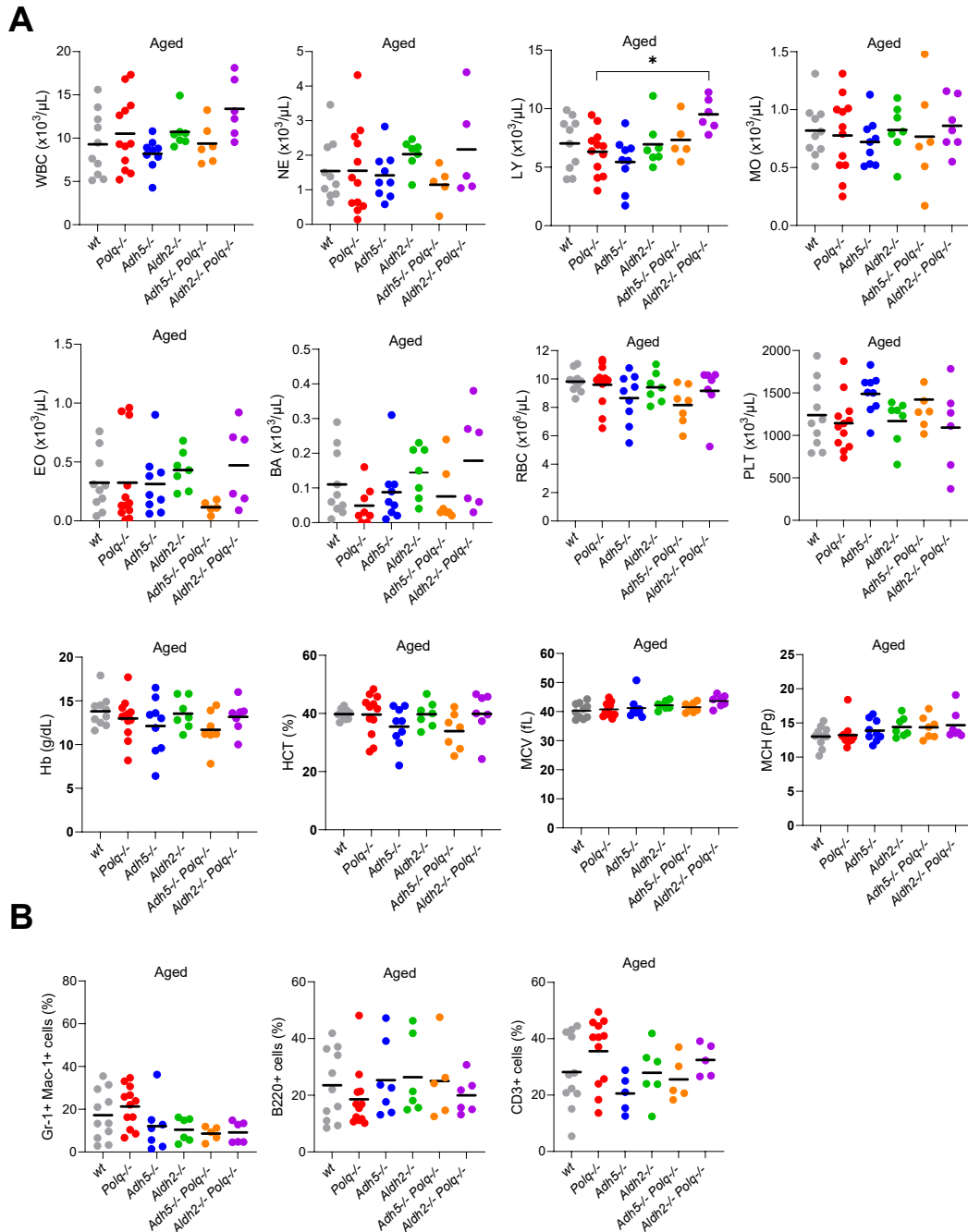


Figure 11. Analysis of peripheral blood from aged (18-24 months-old) mice. (A) Hemavet analysis of peripheral blood cell parameters: white blood cells (WBC), neutrophils (NE), lymphocytes (LY), monocytes (MO), eosinophils (EO), basophils (BA), red blood cells (RBC), platelets (PLT), hemoglobin (Hb), hematocrit (HCT), mean corpuscular volume (MCV), and mean corpuscular hemoglobin (MCH). (B) Gr-1+Mac-1+, B220+ and CD3+ cells were detected in peripheral blood by flow cytometry.

To test the impact of physiological levels of formaldehyde on hematopoiesis, hematological analysis was performed on young (3-4 months-old) and aged (18+ months-old) wild-type, *Polq*^{-/-}, *Adh5*^{-/-}, *Aldh2*^{-/-}, *Adh5*^{-/-};*Polq*^{-/-} and *Aldh2*^{-/-};*Polq*^{-/-} mice. No significant changes were detected in hematological blood parameters of *Adh5*^{-/-};*Polq*^{-/-} and *Aldh2*^{-/-};*Polq*^{-/-} double knockout mice when compared to their respective single knockout counterparts except for modestly elevated levels of Gr-1+Mac-1+ in young *Adh5*^{-/-};*Polq*^{-/-} and *Aldh2*^{-/-};*Polq*^{-/-} mice (Figure 10A, 10B). No significant changes were detected in hematological blood parameters of aged mice (Figure 11A, 11B).

Bone marrow analysis was performed on young and aged mice. No major changes were observed in bone marrow cell lineages of the double knockout mice when compared to their respective single knockout counterparts except a modest decrease of the percentage of Lin-cKit⁺ HSPCs in aged *Adh5*^{-/-};*Polq*^{-/-} mice (Figure 12A, 12B, 13A, 13B). Serial clonogenic assays did not detect significant differences when comparing double knockouts to single knockouts in young and aged cells (Figure 12C, 13C). *Aldh2*^{-/-} and *Aldh2*^{-/-};*Polq*^{-/-} groups did have a decreased clonogenic capacity that was noted. Finally, the absence of Polθ and ADH5 or ALDH2 was not associated with elevated levels of DSBs, detected by γH2AX in Lin-cKit⁺ HSPCs (Figure 12D, 13D). Overall, simultaneous knockout of *Polq* and *Adh5* or *Aldh2* did not result in any abnormalities in the hematological parameters of young and aged mice. Therefore, it seems unlikely that Polθ is needed to support ADH5 or ALDH2 in protecting murine hematopoietic cells from physiological levels of formaldehyde. Polθ-mediated end-joining (TMEJ) has been previously described as a back-up DSB repair pathway (72), so its true biological significance may be masked under physiological conditions.

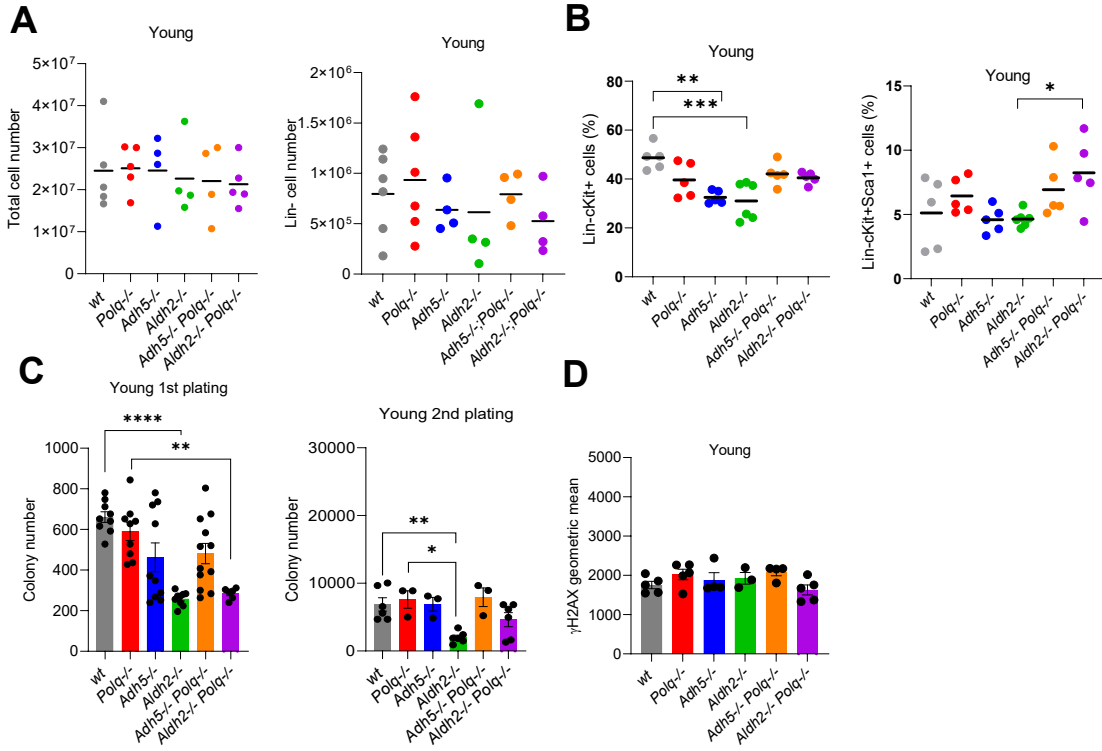


Figure 12. Analysis of bone marrow from young mice. (A) Total cell number. (B) Lin- cell number. (C) Data presented as mean number of colonies ± SEM. (D) Data presented as geometric mean ± SEM.

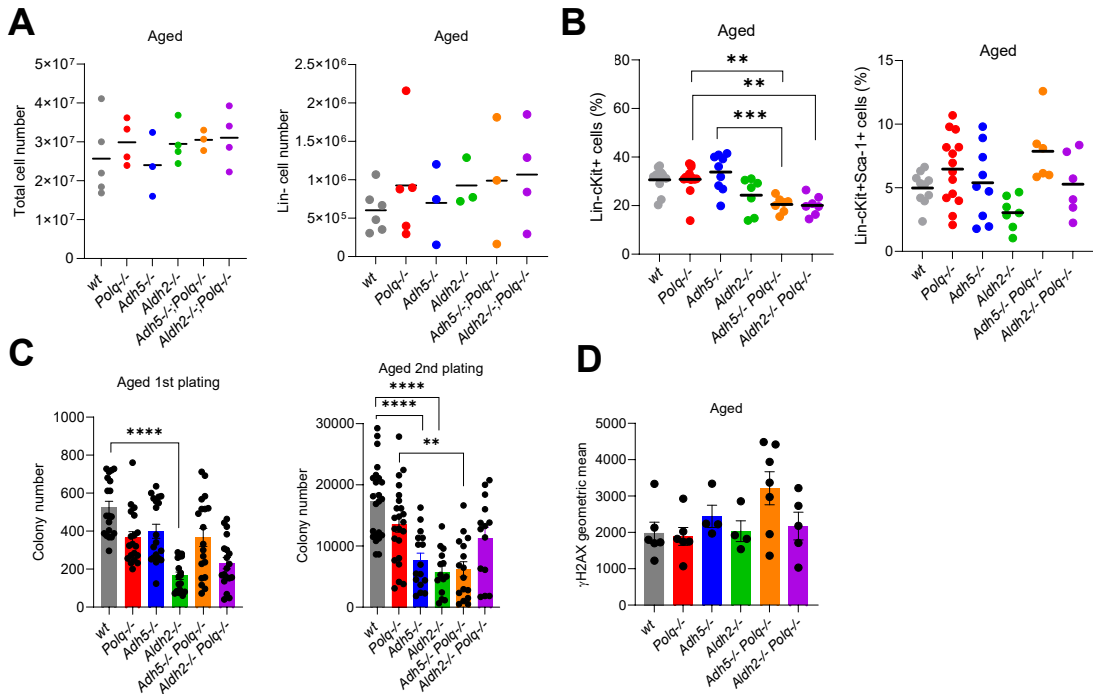


Figure 13. Analysis of bone marrow from aged mice. (A) Total cell number. (B) Lin- cell number. (C) Data presented as mean number of colonies ± SEM. (D) Data presented as geometric mean ± SEM.

Polθ cooperates with ADH5 and ALDH2 to protect hematopoietic cells challenged by high levels of formaldehyde

To take a further look at the cooperativity between ADH5, ALDH2, and Polθ in protecting hematopoietic cells from formaldehyde, our wild-type, *Polq*^{-/-}, *Adh5*^{-/-}, *Aldh2*^{-/-}, *Adh5*^{-/-};*Polq*^{-/-} and *Aldh2*^{-/-};*Polq*^{-/-} mice were challenged by increased levels of formaldehyde. Methanol is metabolized in the liver through a series of oxidative reactions. ADH1b is an alcohol dehydrogenase responsible for the oxidation of methanol to produce formaldehyde which is then oxidized by ADH5 or ALDH2 to produce formate or formic acid which can be broken down into CO₂ and H₂O (8). Therefore, our genetic knockout mice were challenged with methanol as previously described in (28). All mice were given two intraperitoneal (IP) injections of 10% methanol. On day 10, mice were euthanized, and blood and bone marrow were collected for analysis of DNA damage and clonogenic capacity (Figure 14A). Methanol treatment did not cause any major weight loss or adverse side effects in these mice (Figure 14B).

As expected, increased levels of formaldehyde were observed in peripheral blood samples taken from *Adh5*^{-/-}, *Aldh2*^{-/-}, *Adh5*^{-/-};*Polq*^{-/-} and *Aldh2*^{-/-};*Polq*^{-/-} mice prior to methanol injections when compared to peripheral blood samples taken from wild-type and *Polq*^{-/-} mice (Figure 14C). Higher formaldehyde concentrations were observed in *Adh5*^{-/-} and *Aldh2*^{-/-} ± *Polq*^{-/-} mice after administration of methanol (Figure 14C). However, intracellular concentrations of formaldehyde were not measured in the bone marrow prior to and after methanol injections which was a limitation of this study. This data is consistent with other reports, demonstrating that knockout of ADH5 or ALDH2 elevates formaldehyde in the blood (25, 28, 38).

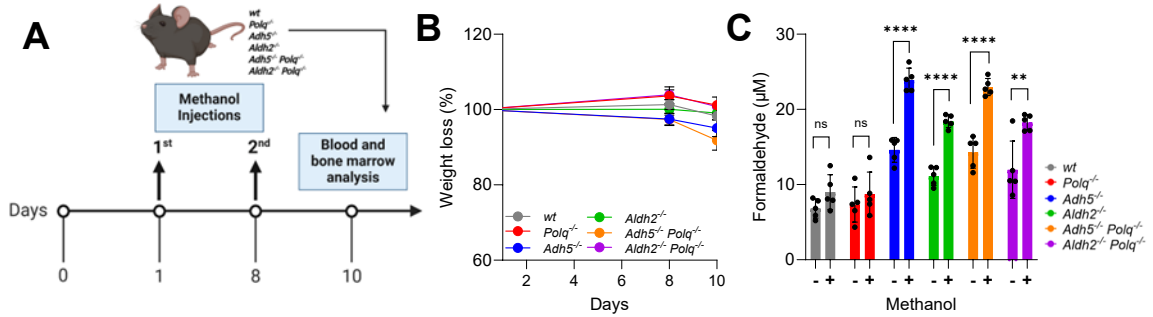


Figure 14. Methanol increases formaldehyde levels in murine blood. (A) Experimental timeline of methanol challenge. (B) Data presented as mean weight \pm SEM (n=5 per group). (C) Data presented as mean formaldehyde levels in peripheral blood of mice treated (+) or not (-) with methanol \pm SEM (n=5 per group).

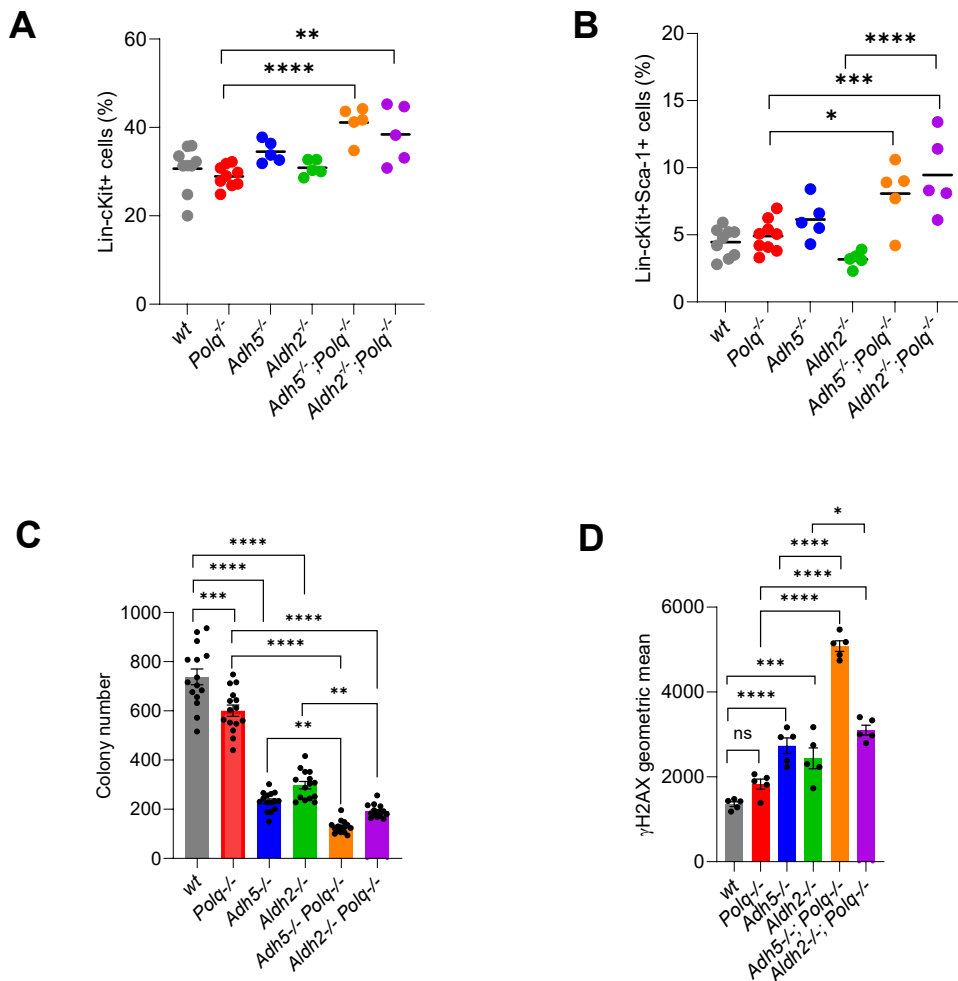


Figure 15. Pol θ and ADH5 or ALDH2 cooperate to protect murine HSPCs from methanol. (A) Data presented as mean Lin-cKit⁺ BMCs \pm SEM. (B) Data presented as mean Lin-cKit⁺Sca-1⁺ BMCs \pm SEM. (C) Data presented as mean number of colonies \pm SEM. (D) Data presented as geometric mean \pm SEM of γ H2AX immunofluorescence.

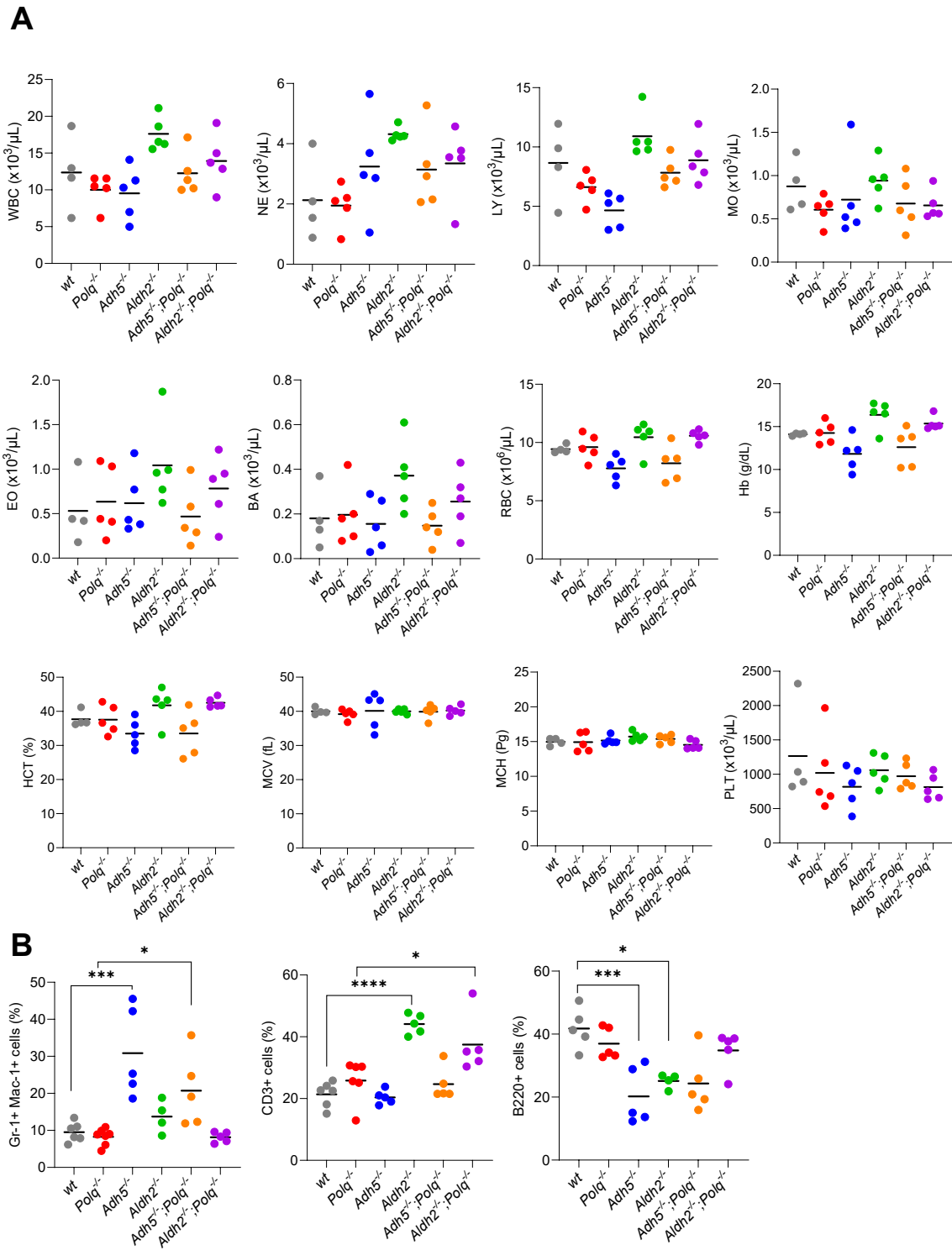


Figure 16. Analysis of peripheral blood from mice after methanol challenge. (A) Data presented as mean \pm SEM of the indicated hematological parameters in peripheral blood. (B-C) Data presented as mean \pm SEM of the indicated hematological parameters in peripheral blood via flow cytometry analysis.

Remarkably, the clonogenic capacity of Lin-cKit⁺ BMCs obtained from methanol treated mice was significantly reduced in *Adh5*^{-/-};*Polq*^{-/-} and *Aldh2*^{-/-};*Polq*^{-/-} mice when compared to their respective single knockout counterparts which was associated with accumulation of DSBs, detected by γ H2AX immunofluorescence (Figure 15C, 15D). Meanwhile, the methanol derived formaldehyde did not significantly alter peripheral blood cell parameters in *Adh5*^{-/-};*Polq*^{-/-} and *Aldh2*^{-/-};*Polq*^{-/-} mice when compared to their respective single knockout counterparts (Figure 16A, 16B). But the percentage of Lin-cKit⁺Sca-1⁺ HSPCs in the bone marrow of *Adh5*^{-/-};*Polq*^{-/-} and *Aldh2*^{-/-};*Polq*^{-/-} mice were moderately enhanced when compared to the single-knockouts (Figure 15A, 15B). These results suggest that ADH5, ALDH2 and Pol θ cooperate in protecting murine hematopoietic cells from formaldehyde challenge.

To validate this, Lin-cKit⁺ BMCs were treated with various doses of formaldehyde *in vitro*. *Adh5*^{-/-};*Polq*^{-/-} HSPCs were extremely sensitive to the 5-20 μ M formaldehyde range when compared to *Adh5*^{-/-} and *Polq*^{-/-} cells while *Aldh2*^{-/-};*Polq*^{-/-} HSPCs demonstrated sensitivity only to 50 μ M of formaldehyde when compared *Aldh2*^{-/-} and *Polq*^{-/-} cells (Figure 17A). Inhibition of the clonogenic activity of *Adh5*^{-/-};*Polq*^{-/-} and *Aldh2*^{-/-};*Polq*^{-/-} cells exposed to formaldehyde was also associated with enhanced accumulation of DSBs and DPCs (Figure 17B, 17C). The higher formaldehyde sensitivity of *Adh5*^{-/-};*Polq*^{-/-} cells compared to *Aldh2*^{-/-};*Polq*^{-/-} likely stems from ADH5 being the primary enzyme in formaldehyde detoxification, while ALDH2 is a backup (25, 28). Formaldehyde reacts with glutathione (GSH) and ADH5 oxidizes the formaldehyde-GSH products (26). GSH synthesis inhibitor (L-BSO) did not increase the sensitivity of *Adh5*^{-/-};*Polq*^{-/-} cells to formaldehyde-induced DPC-mediated DSBs but greatly enhanced the sensitivity of *Aldh2*^{-/-};*Polq*^{-/-} cells (Figure 18). Altogether, these results clearly indicate that Pol θ provides an additional tier of protection against elevated levels of formaldehyde in hematopoietic cells.

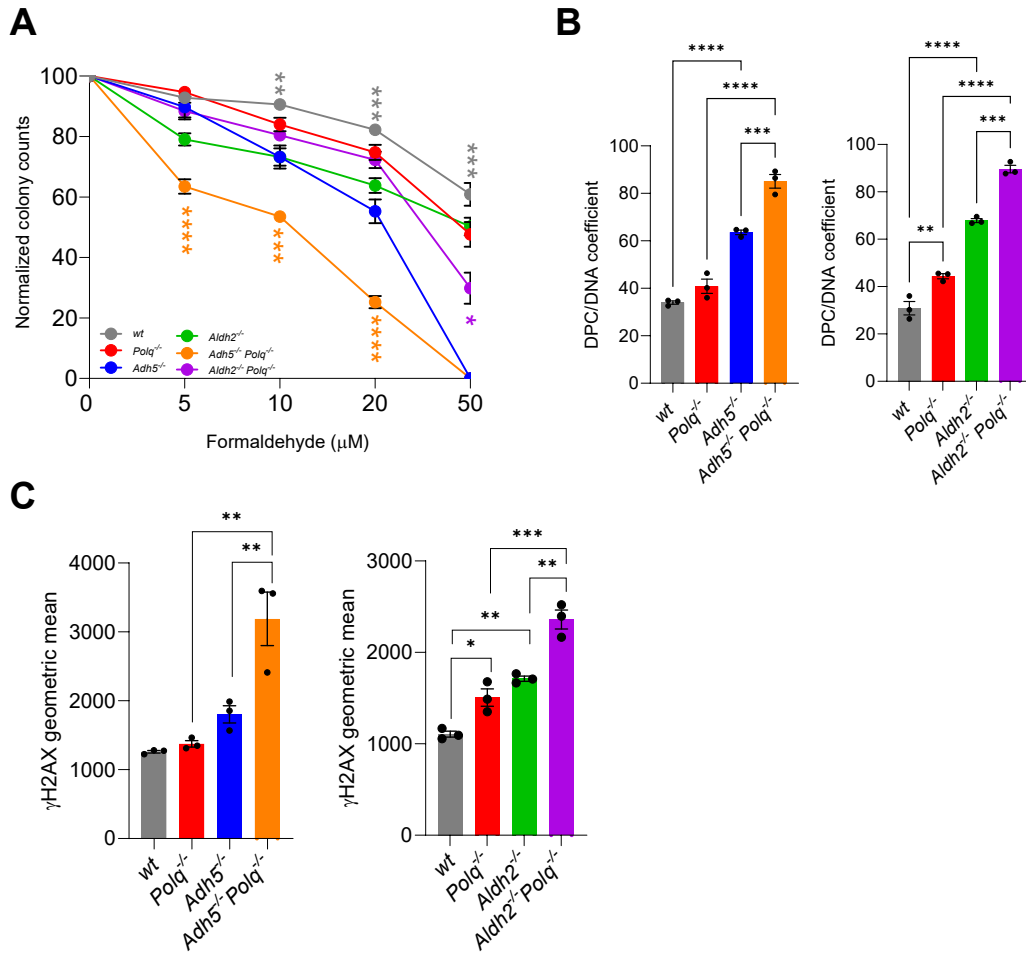


Figure 17. Polθ and ADH5 or ALDH2 cooperate to protect HSPCs from formaldehyde. (A) Data presented as mean colonies \pm SEM (n=3 per group). (B) Data presented as mean DPC/DNA coefficient \pm SEM (n=3 per group). (C) Data presented as geometric mean \pm SEM (n=3 per group).

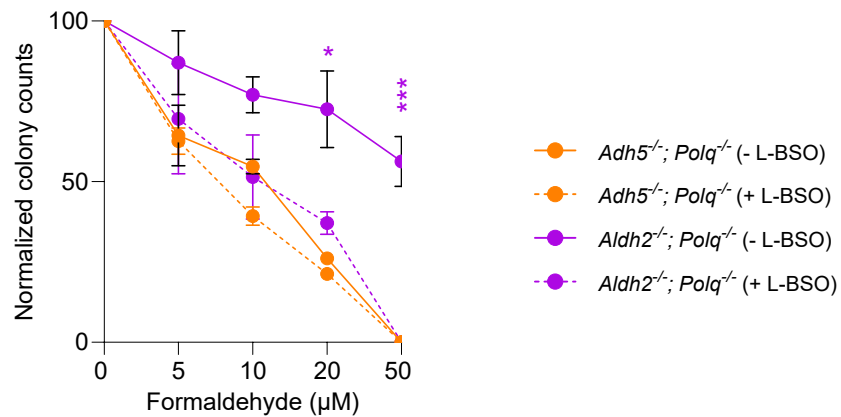


Figure 18. L-BSO enhances synthetic lethality in *Aldh2*^{-/-}; *Polq*^{-/-} HSPCs. Data presented as mean colonies \pm SEM (n=3 per group).

Pharmacological inhibition of ADH5 or ALDH2 increases formaldehyde concentrations and reduces the growth of *Polq*^{-/-} HSPCs

To validate the collaboration between Polθ and ADH5 or ALDH2 in protecting HSPCs from formaldehyde challenge, ADH5 and ALDH2 dehydrogenase activities were inhibited by N6022 (ADH5i) or disulfiram (ALDH2i), in wild-type and *Polq*^{-/-} murine Lin⁻cKit⁺ HSPCs (Figure 19A, 20A) (133, 134). Inhibition of ADH5 or ALDH2 resulted in accumulation of formaldehyde (Figure 19B, 20B) and DPCs (Figure 19D, 20D) in wildtype and *Polq*^{-/-} cells. However, neutral comet assay revealed enhanced DSBs and reduced clonogenic potential only in *Polq*^{-/-} murine HSPCs treated with ADH5i or ALDH2i when compared to wild-type counterparts (Figure 19C, 19E, 20C, 20E).

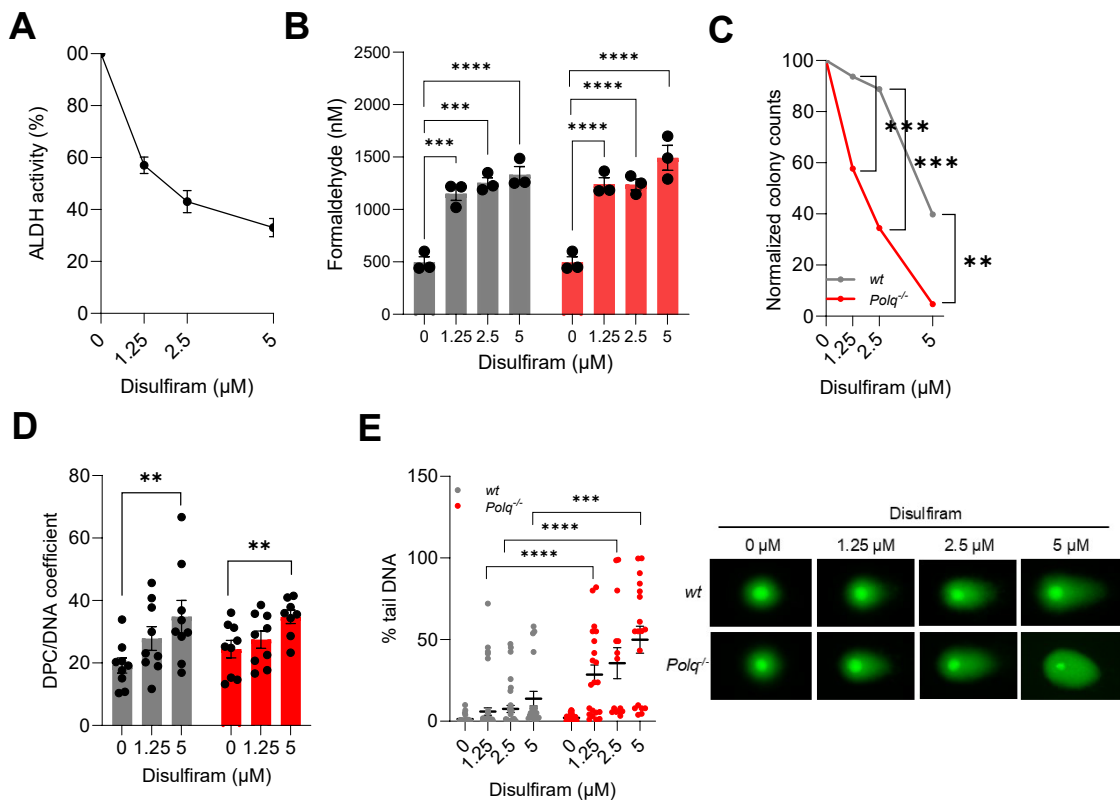


Figure 19. Simultaneous inhibition of Polθ and ALDH2 is toxic for murine HSPCs. (A) ALDH activity after treatment with disulfiram with data presented as mean ± SEM (n=3 per group). (B) Formaldehyde detection with data presented as mean ± SEM. (C) Clonogenic assay with data presented as mean colonies ± SEM (n=3). (D) ARK assay with data presented as mean DPC/DNA coefficient ± SEM. (E) Neutral comet assay with data presented as mean % tail DNA ± SEM (left). Representative comet images (right).

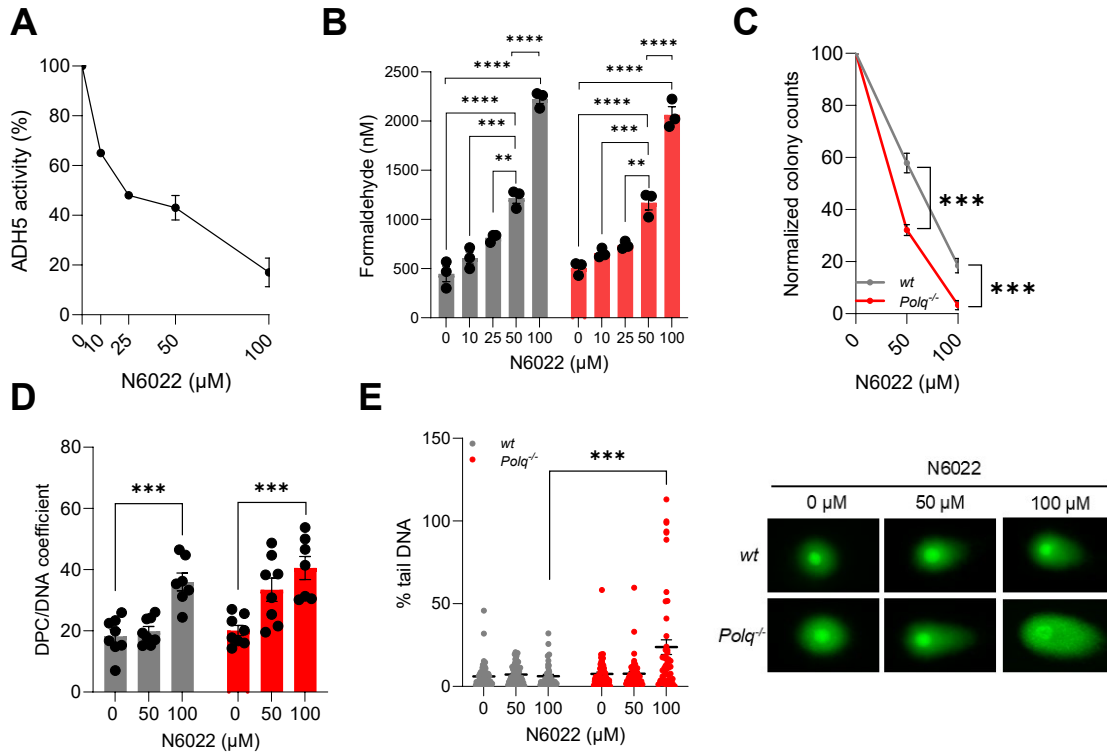


Figure 20. Simultaneous inhibition of Polθ and ADH5 is toxic for murine HSPCs. (A) ADH activity after N6022 treatment with data presented as mean ± SEM (n=3 per group). (B) Formaldehyde detection with data presented as mean ± SEM. (C) Clonogenic assay with data presented as mean colonies ± SEM (n=3 per group). (D) ARK assay with data presented as mean DPC/DNA coefficient ± SEM. (E) Neutral comet assay with data presented as mean % tail DNA ± SEM (left). Representative comet images (right).

Next, Polθ-mediated TMEJ activities were inhibited by the Polθi ART558 in wild-type, *Adh5*^{-/-} and *Aldh2*^{-/-} murine Lin⁺cKit⁺ HSPCs (Figure 21A) (92). The inhibition of TMEJ did not significantly increase DPCs in *Adh5*^{-/-} and *Aldh2*^{-/-} HSPCs but caused accumulation of DSBs detected by anti-γH2AX immunofluorescence, nuclear foci and western blot, consistent with the role of Polθ in repairing DSBs, but not DPCs (Figure 21C, 21D, 21E, 21F). The increase in DSBs was associated with reduced colony formation of *Adh5*^{-/-} and *Aldh2*^{-/-} HSPCs treated with Polθi when compared to wild-type counterparts (Figure 21B).

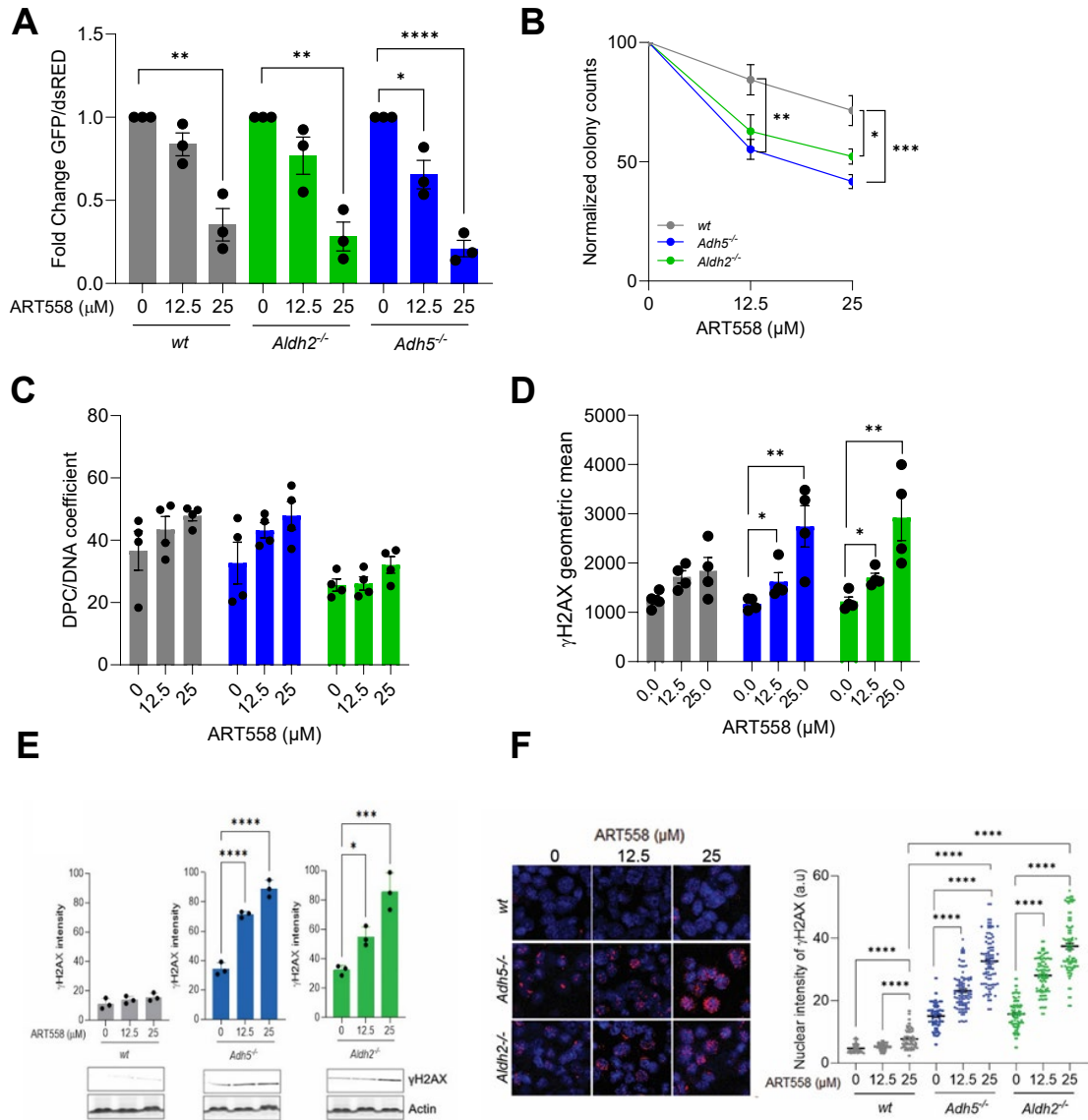


Figure 21. Simultaneous inhibition of Polθ and ADH5 or ALDH2 is toxic to murine HSPCs. (A-F) Wild-type, *Adh5*^{-/-} and *Aldh2*^{-/-} Lin-cKit⁺ BMCs treated with indicated concentrations of Polθi for 72 hours. (A) Data presented as mean TMEJ activity ± SEM. (B) Clonogenic assay with data presented as mean colonies ± SEM (n=3 per group). (C) ARK assay with data presented as average DPCs ± SEM. (D) γH2AX analysis with data presented as geometric mean ± SEM. (E) Representative western blots detecting γH2AX (lower panels). Quantification of γH2AX protein levels (upper panels). (F) Representative γH2AX nuclear foci (red), DNA was counterstained with DAPI (blue) (*left panel*). Mean level ± SEM of the γH2AX nuclear foci intensity (*right panel*).

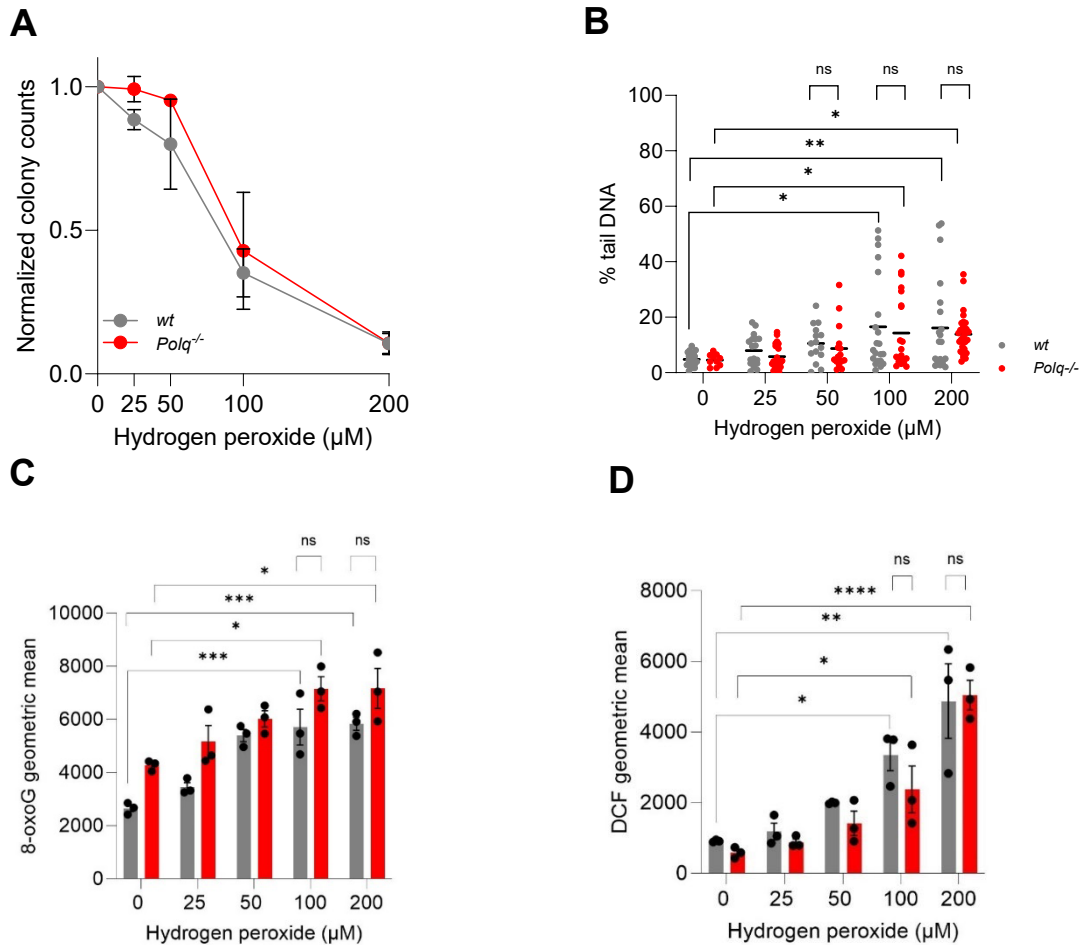


Figure 22. *Polq*^{-/-} HSPCs are not more sensitive to oxidative DNA damage. (A-D) Wild-type and *Polq*^{-/-} cells were treated with various concentrations of hydrogen peroxide for 4 hours. (A) Data presented as mean colonies \pm SEM (n=3). (B) Data presented as mean % tail DNA \pm SEM. (C) Data presented as geometric mean \pm SEM. (D) Data presented as DCF geometric mean \pm SEM.

Since formaldehyde also triggers oxidative DNA lesions, the role of Polθ in responding to oxidative DNA damage was evaluated using hydrogen peroxide. Remarkably, *Polq*^{-/-} murine HSPCs displayed similar sensitivity to hydrogen peroxide and accumulated similar amounts of reactive oxygen species (ROS), oxidative DNA damage (8-oxoG) and DSBs as wild-type cells (Figure 22A, 22B, 22C, 22D). This observation highlights the collaborative role of Polθ and ADH5 or ALDH2 in selective protection of HSPCs against the toxic effect of formaldehyde-mediated DSBs, but not hydrogen peroxide-mediated DSBs in murine HSPCs.

Polθ and ADH5 or ALDH2 protect AML and MPN cells from the genotoxic effects of formaldehyde induced by OTKs

Recently, we reported that OTKs such as FLT3(ITD) and JAK2(V617F) stimulate one-carbon metabolism, resulting in accumulation of high levels of formaldehyde and Polθ plays a critical role in protecting AML and MPN cells from genotoxic effects of formaldehyde (124). To test if Polθ collaborates with ADH5 or ALDH2 in protecting leukemic cells from formaldehyde generated by OTKs, FLT3(ITD) or JAK2(V617F) was expressed in wild-type, *Polq*^{-/-}, *Adh5*^{-/-}, *Aldh2*^{-/-}, *Adh5*^{-/-};*Polq*^{-/-} and *Aldh2*^{-/-};*Polq*^{-/-} Lin-cKit⁺ BMCs. Cell proliferation and colony formation by FLT3(ITD)-positive double-knockout *Adh5*^{-/-};*Polq*^{-/-} and *Aldh2*^{-/-};*Polq*^{-/-} cells was severely reduced when compared to FLT3(ITD)-positive *Polq*^{-/-}, *Adh5*^{-/-}, and *Aldh2*^{-/-} single-knockouts, indicating a combined protective effect of Polθ with either of these dehydrogenases (Figure 23A, 23C). Cell proliferation and clonogenic capacity of JAK2(V617F)-positive *Aldh2*^{-/-};*Polq*^{-/-} cells was also suppressed when compared to *Aldh2*^{-/-} and *Polq*^{-/-} counterparts (Figure 23B, 23D). In contrast to FLT3(ITD), proliferation and clonogenicity of JAK2(V617F)-positive *Adh5*^{-/-};*Polq*^{-/-} cells was like that of *Adh5*^{-/-} and *Polq*^{-/-} counterparts suggesting that Polθ and ADH5 do not functionally cooperate in JAK2(V617F)-positive HSPCs.

In addition to genetic knockout data, FLT3(ITD)-positive AML 201 and AML1 patient cells (Figure 24A, 24B) and JAK2(V617F)-positive MPN1 and P293 MPN patient cells (Figure 25A, 25B) were more sensitive to the combination of ALDH_i and Polθ_i when compared to treatment with these drugs used as single agents. As with murine HSPCs, only FLT3(ITD)-positive AML primary cells, but not JAK2(V617F)-positive MPN primary cells were more sensitive to ADH5_i and Polθ_i. Noteworthy, FLT3_i quizartinib and JAK1/2_i ruxolitinib enhanced the effect of the combinations of Polθ_i + ALDH_i and Polθ_i + ADH5_i in AML and MPN patient cells (Figure 24A, 24B, 25A, 25B).

These data indicate Polθ cooperates with ADH5 or ALDH2 in protecting AML and MPN cells from the genotoxic effects of formaldehyde. This is indirectly supported by TCGA data, showing the mRNA expression of *ALDH2* and *ADH5* to be relatively low in AML when compared to other tumors, but seems to be compensated by high *POLQ* mRNA expression (Figure 26A). In addition, *POLQ* and *ADH5/ALDH2* genetic aberrations appeared to be mutually exclusive in tumors in the TCGA datasets, suggesting that they belong to cooperating oncogenic pathways (Figure 26B).

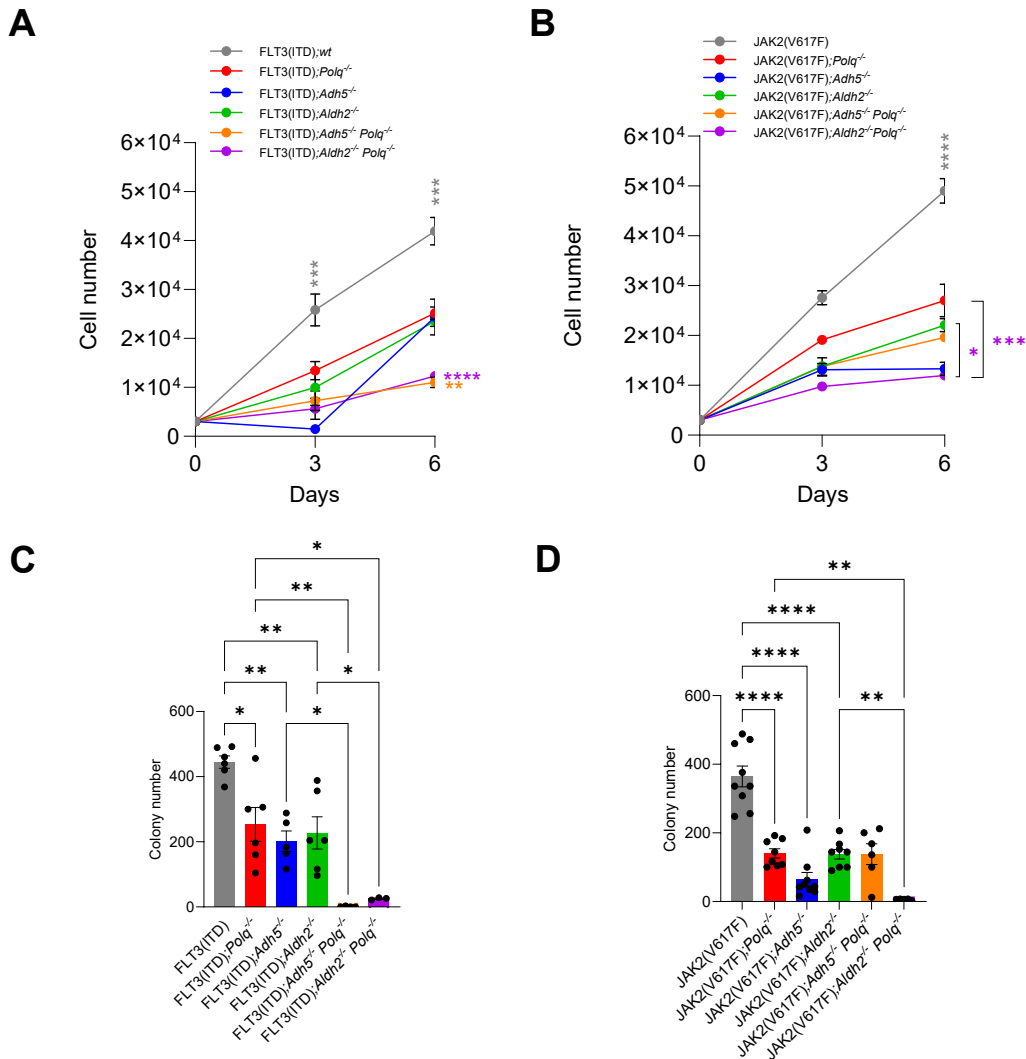


Figure 23. Polθ and ADH5 or ALDH2 protect murine HSPCs from formaldehyde challenge triggered by FLT3(ITD) or JAK2(V617F) OTKs. (A-B) Cells grown in culture over a period of 6 days. Data presented as mean \pm SEM. (C-D) Clonogenic assay with data presented as mean number of colonies \pm SEM.

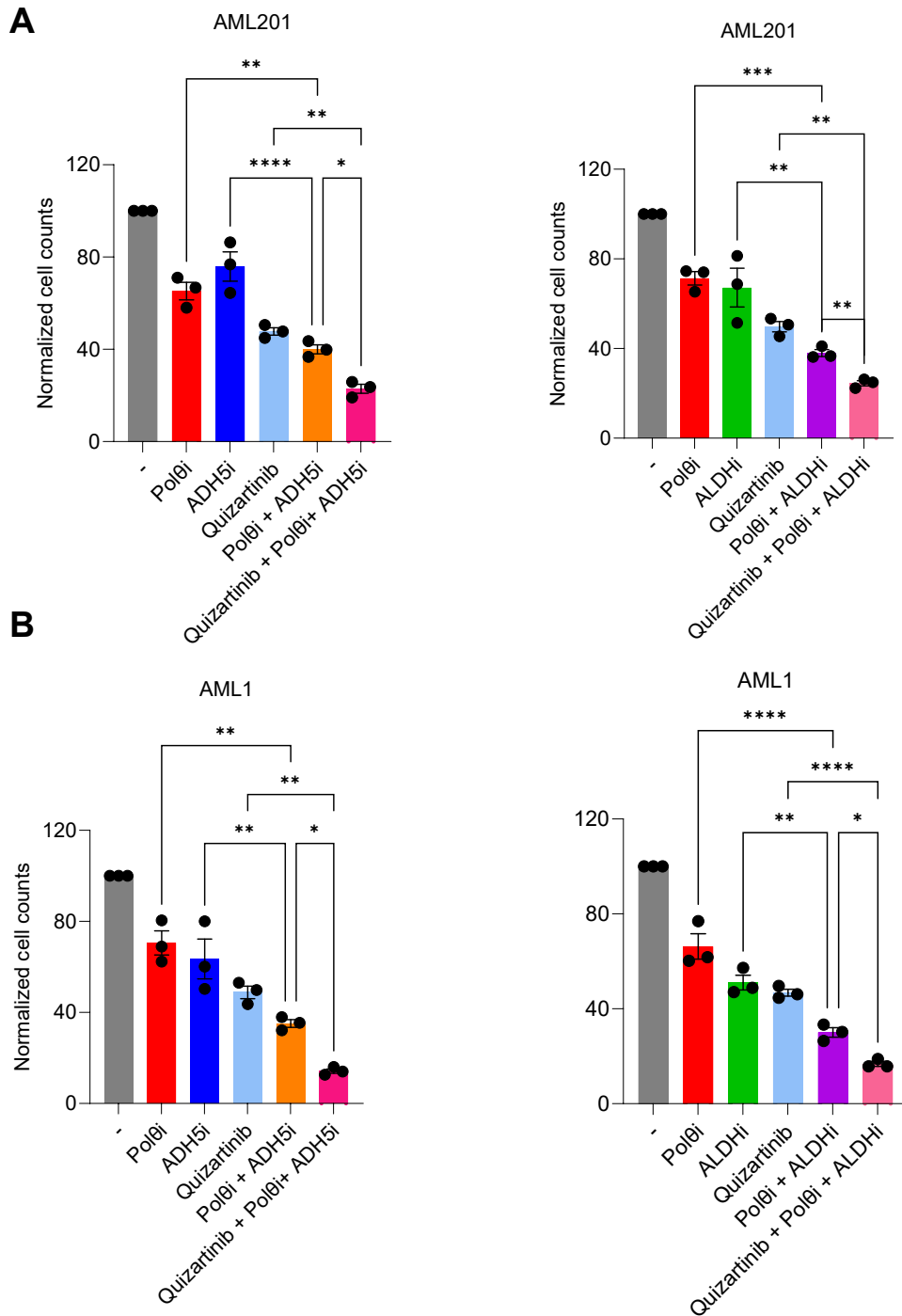


Figure 24. Polθ and ADH5 or ALDH2 protect AML patient cells from formaldehyde challenge triggered by FLT3(ITD). (A-B) FLT3(ITD)-positive AML 201 and AML1 patient cells treated with 25 μM Polθi, 0.075 μM ALDH2i, 75 μM ADH5i, 0.125 μM FLT3i alone or in combinations for 6 days. Data presented as mean cell number ± SEM.

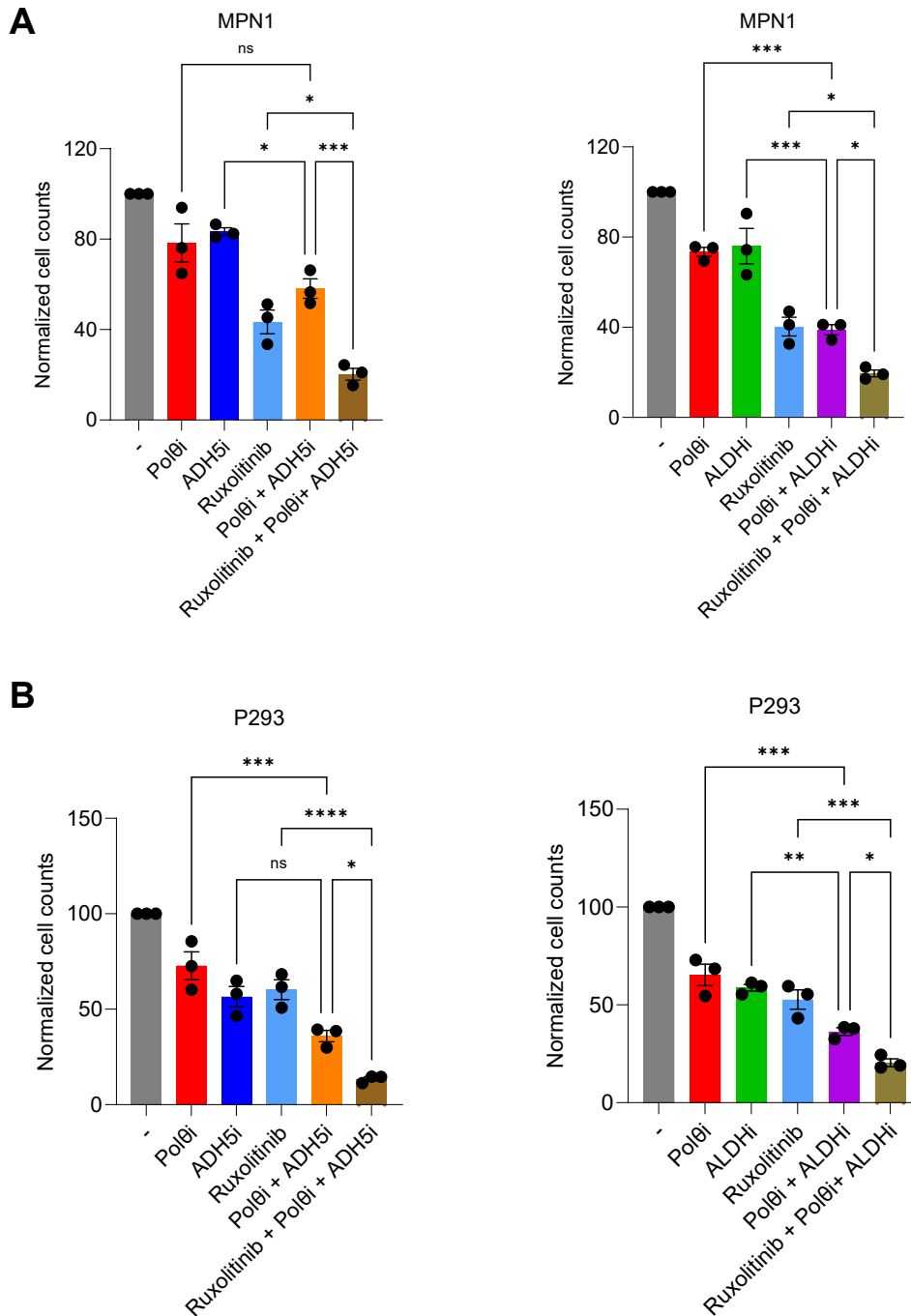
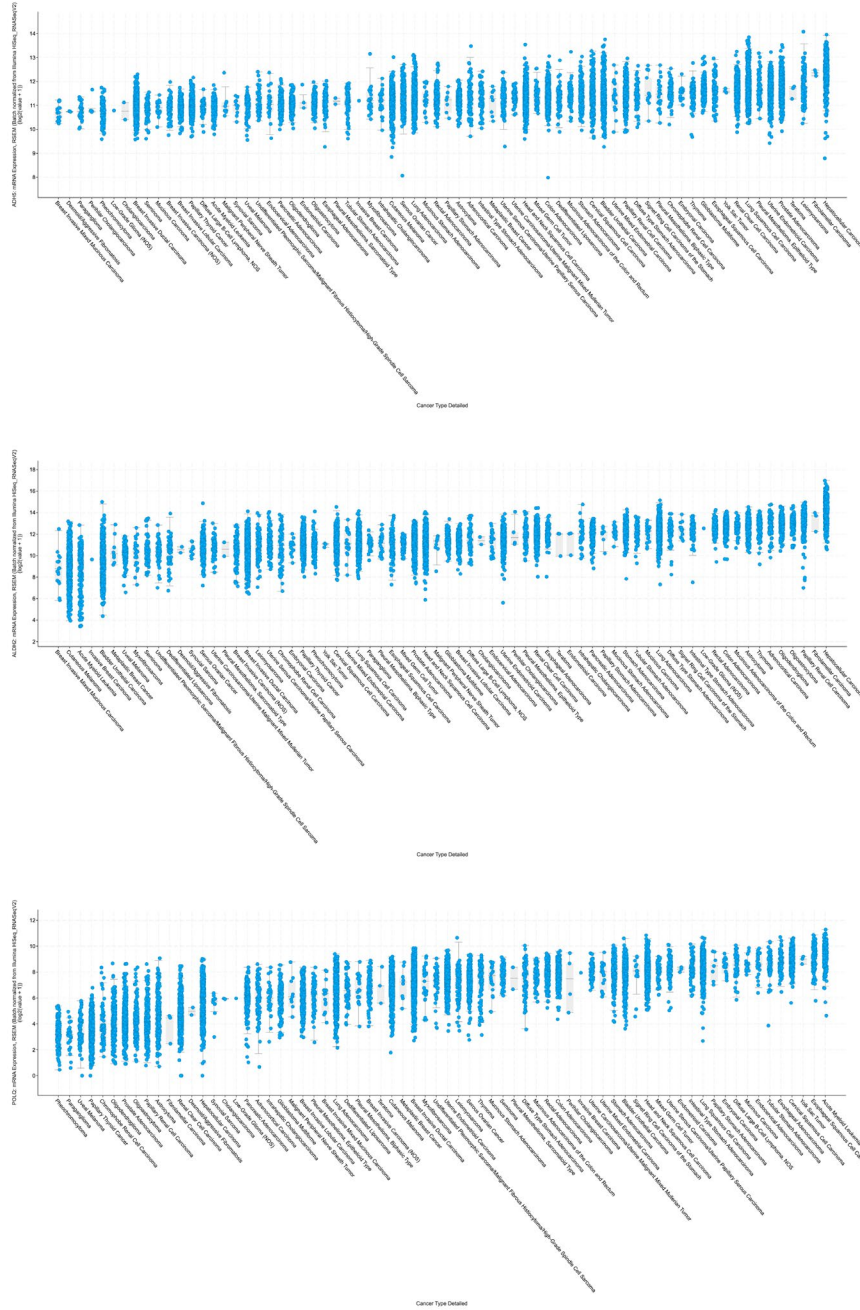


Figure 25. Polθ and ADH5 or ALDH2 protect MPN patient cells from formaldehyde challenge triggered by JAK2(V617F). (A) JAK2(V617F)-positive MPN1 patient cells treated with 25 μM Polθi, 0.05 μM ALDH1, 50 μM ADH5i, 0.1 μM JAK1/2i alone or in combinations for 6 days. Data presented as mean ± SEM. (B) JAK2(V617F)-positive MPN P293 patient cells treated with 25 μM Polθi, 0.075 μM ALDH1, 75 μM ADH5i, 0.1 μM JAK1/2i alone or in combinations for 6 days. Data presented as mean ± SEM.

A



B

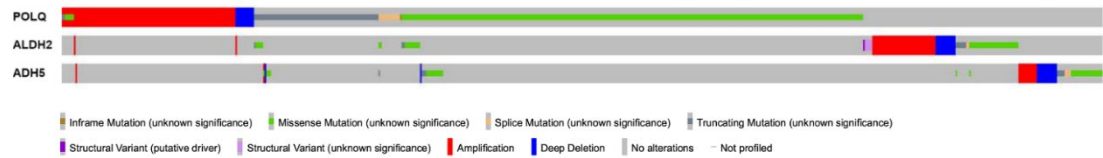


Figure 26. Global analysis of *ADH5*, *ALDH2* and *POLQ* expression in cancer. (A) mRNA expression of *ADH5*, *ALDH2* and *POLQ* in various tumors in TCGA datasets. (B) Co-occurrence of *ADH5*, *ALDH2*, and *POLQ* alterations in 10967 tumor samples from TCGA datasets.

Simultaneous inhibition of Polθ and ADH5 or ALDH2 exerts an anti-leukemia effect

Because inhibition of Polθ is synthetically lethal in HR-deficient cancer cells, we hypothesized that inhibition of ADH5 or ALDH2 may enhance this anti-tumor effect (81, 135). To test this, we utilized Nalm6 parental leukemic cells and their isogenic HR-deficient RAD54 knockout (HR-deficient) counterparts; the latter displaying hypersensitivity to Polθi (124, 136). Cells were treated with Polθi, ADH5i, and ALDHi alone or in combination. Combined with Polθi, either ADH5i or ALDHi enhanced accumulation of DSBs selectively in HR-deficient Nalm6-*RAD54*^{-/-} cells (Figure 27A, 27B). Polθi + ALDHi and Polθi + ADH5i combinations increased the percentage of apoptotic cells by 3-4-fold in Nalm6-*RAD54*^{-/-} cells, while exerting very modest effect in Nalm6 parental cells (Figure 27C, Appendix Figure 2).

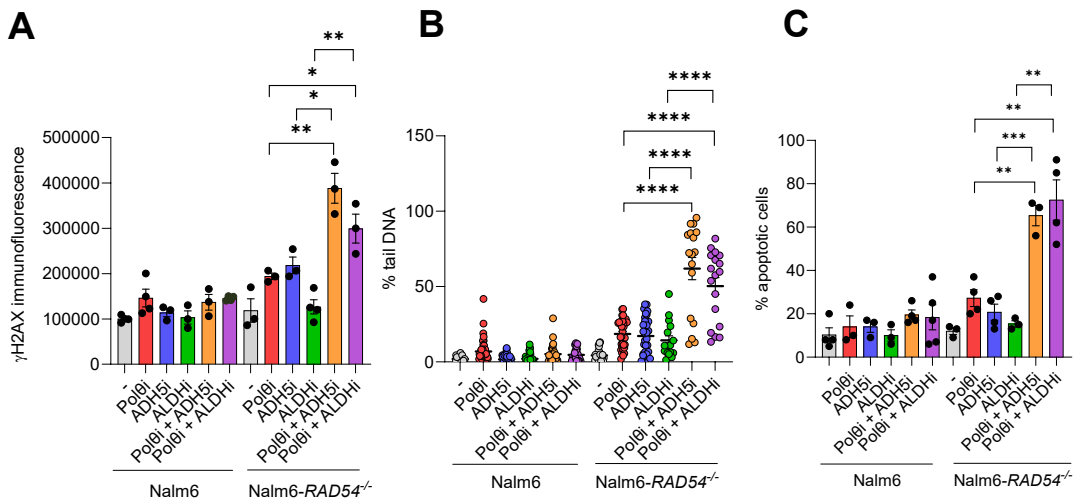


Figure 27. Combination of Polθi + ALDHi and Polθi + ADH5i increased DNA damage in HR-deficient leukemia cells. (A-C) Cells were treated with 12.5 μ M Polθi, 0.125 μ M ALDHi, 20 μ M ADH5i, and the indicated combinations for 48 hours. (A) γ H2AX analysis with data presented as mean γ H2AX immunofluorescence \pm SEM. (B) Neutral comet assay with data presented as mean % tail DNA \pm SEM. (C) Apoptosis analysis with data presented as mean % \pm SEM of the annexin V/propidium iodide double-stained cells.

Accumulation of DSBs and induction of apoptosis in Nalm6-*RAD54*^{-/-} cells treated with Polθi + ALDH_i or Polθi + ADH5_i was accompanied by reduction of their clonogenic capacity when compared to parental counterparts (Figure 28). For example, 0.125 μM ALDH_i enhanced the effect of 12.5 μM Polθi by >6-fold and of 25 μM Polθi by >10-fold in Nalm6-*RAD54*^{-/-} cells, while exerting very modest effect in Nalm6 counterparts (Figure 24A). Similarly, 20 μM ADH5_i enhanced the effect of 12.5 μM Polθi by >4-fold and of 25 μM Polθi by >2-fold in Nalm6-*RAD54*^{-/-} cells. No effect was observed in Nalm6 parental cells (Figure 28B). Based on these findings, we hypothesized that ADH5 and ALDH inhibitors may enhance the synthetically lethal effect of Polθi against HR-deficient leukemia cells.

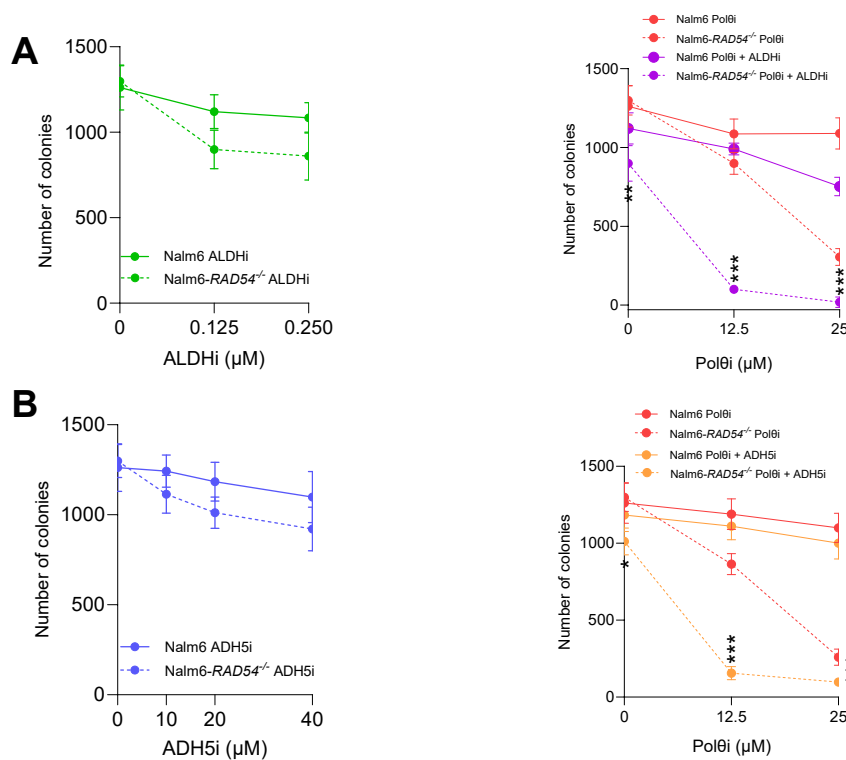


Figure 28. Combination of Polθi + ALDH_i and Polθi + ADH5_i decreased proliferation of HR-deficient leukemia cells. (A) Cells treated with ALDH_i (*left*), and 0.125 μM ALDH_i + 12.5 μM or 25 μM of Polθi (*right*) for 48 hours followed by clonogenic assay. Data presented as mean colonies ± SEM (n=3 per group). (B) Cells treated with ADH5_i (*left*), and 20 μM ADH5_i + 12.5 μM or 25 μM of Polθi (*right*) for 48 hours followed by clonogenic assay. Results represent mean colonies ± SEM (n=3 per group) when compared to untreated counterparts.

To validate our *in vitro* findings and determine if there is any clinical relevance, SCID beige mice were injected with Nalm6-*RAD54*^{-/-} cells. Mice were then randomized into six different groups: control, RP-6685, disulfiram, N6022, RP-6685 + disulfiram and RP-6685 + N6022. 7 days after induction, mice were treated with their respective inhibitor(s) for 14 consecutive days via oral gavage. 2 days after the last dose was given, a blood sample was taken via tail nick to run flow cytometry analysis for human CD19⁺ cells to determine leukemic burden (Figure 29). Remarkably, there was almost no leukemic burden in the RP-6685 + disulfiram and RP-6685 + N6022 groups when compared to the single inhibitor and control mice (Figure 30A). Control, Polθi, ADH5i, and ALDH_i-treated mice died from leukemia after 28.7 ± 0.2, 30.8 ± 0.8, 30.2 ± 0.9 and 30.3 ± 0.8 days, respectively (Figure 30B). In contrast, the combination of Polθi + ALDH_i resulted in a significant increase of MST of 40.0 ± 1.3 days while the Polθi + ADH5i combination resulted in the longest MST of 47.8 ± 1.7 days when compared to individual treatments (Figure 30B). Polθi + ADH5i exerted significantly stronger anti-leukemia effect when compared to Polθi + ALDH_i (p<0.02). All mice died because of leukemia and harbored 16.4 ± 6.7% and 27.0 ± 15.0% of CD19⁺ cells in peripheral blood and bone marrow. This data combined with our *in vitro* data indicate that Polθi + ALDH_i and Polθi + ADH5i combines exert a strong anti-leukemia effect in HR-deficient leukemia cells.

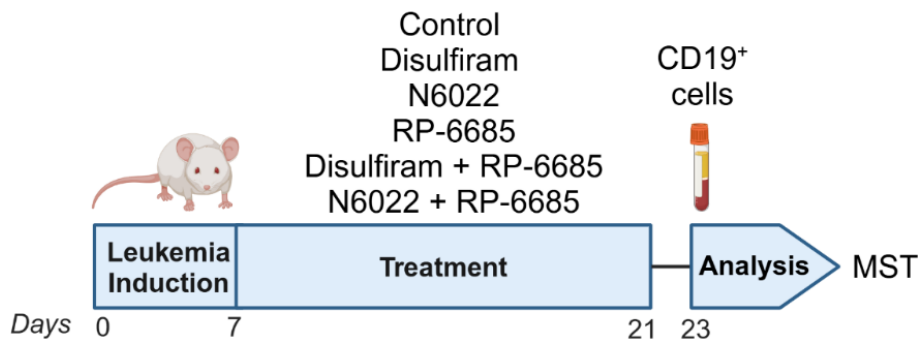


Figure 29. Experimental timeline of mice injected with HR-deficient leukemia cells.

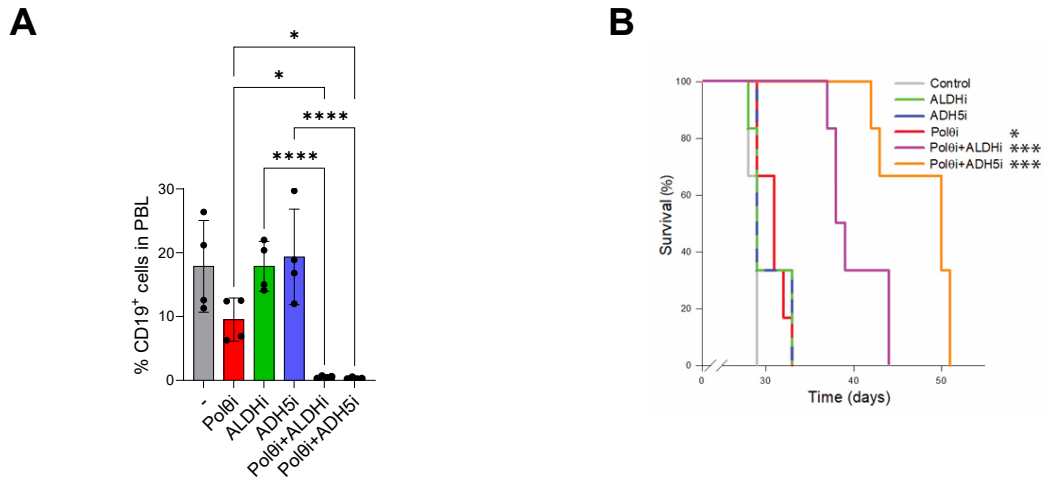


Figure 30. Combination of Polθi + ALDHⁱ and Polθi + ADH⁵ⁱ decreased leukemia burden and increased survival time in mice injected with HR-deficient leukemia cells. (A) Percentage of human CD19⁺ cells detected in the peripheral blood of mice on day 23. Data presented as the mean ± SEM (n=4 per group). (B) Median survival time (MST) of mice treated with the indicated inhibitors (n=5 per group).

Genetic knockout mice/cells challenged by cigarette smoke

Tobacco smoke contains a multitude of carcinogens, including highly reactive aldehydes such as formaldehyde. Formaldehyde is a component of cigarette smoke and is produced when additives such as sugars and sorbitol are burnt. Smokers inhale formaldehyde when they take a puff from a cigarette, commonly referred to as first-hand smoke while non-smokers can inhale it as well from second-hand smoke. It has been found that aldehydes are the predominant forces that induce DNA damage and inhibit DNA repair in tobacco smoke carcinogenesis (10). There are some studies that suggest there is a correlation between cigarette smoking and risk of AML (137). Therefore, we decided to challenge our genetic knockout mice with cigarette smoke as an additional formaldehyde challenge to see whether ADH5, ALDH2, and Polθ cooperate in protecting HSPCs from formaldehyde-containing cigarette smoke.

3-4-month-old wild-type, *Polq*^{-/-}, *Adh5*^{-/-}, and *Aldh2*^{-/-} mice were exposed to cigarette smoke for 3 months using a smoking system at Temple University in collaboration with Beata Kosmider. *Adh5*^{-/-};*Polq*^{-/-}, and *Aldh2*^{-/-};*Polq*^{-/-} double knockout mice were not

used for this experiment because there were not any age matched mice at the time of this experiment. Each day, the smoking system exposed mice to cigarette smoke that mimicked the exposure a primary smoker would encounter. At the end of the 3-month smoking period, blood samples were collected from mice via tail nick. Hematological analysis was performed to determine whether there were any hematological differences or abnormalities when comparing single knockouts to wild-type counterparts. There were no significant differences observed between the groups (Figure 31A, 31B). All mice were monitored daily, and no adverse effects were detected in any mice, so they were left undisturbed in their cages until they reached 18-24 months old to see if any abnormalities developed over time after cigarette smoke exposure.

When the mice reached 18+ months old, blood and bone marrow were collected for further hematological analysis. Again, there were no significant differences in any of the blood parameters when comparing single knockouts to wild-type counterparts (Figure 32A, 32B). There was a difference in the B-cell population (B220+ marker) when comparing *Adh5*^{-/-} vs. *Aldh2*^{-/-} mice. For bone marrow analysis, total BMCs were isolated from femur bones and counted using Trypan blue staining. There was a modest decrease in total cell population in *Adh5*^{-/-} mice, but nothing statistically significant when comparing all single knockouts to the wild-type counterparts (Figure 33A). Flow cytometry analysis of the murine HSPCs showed no differences (Figure 33B). Remarkably, there was a significant decrease in *Polq*^{-/-}, *Adh5*^{-/-}, and *Aldh2*^{-/-} colonies compared to wild-type, (Figure 33C). Increased levels of DSBs were not detected in these cells, so we were not able to confirm whether the decrease in clonogenic capacity is the result of the accumulation of DSBs.

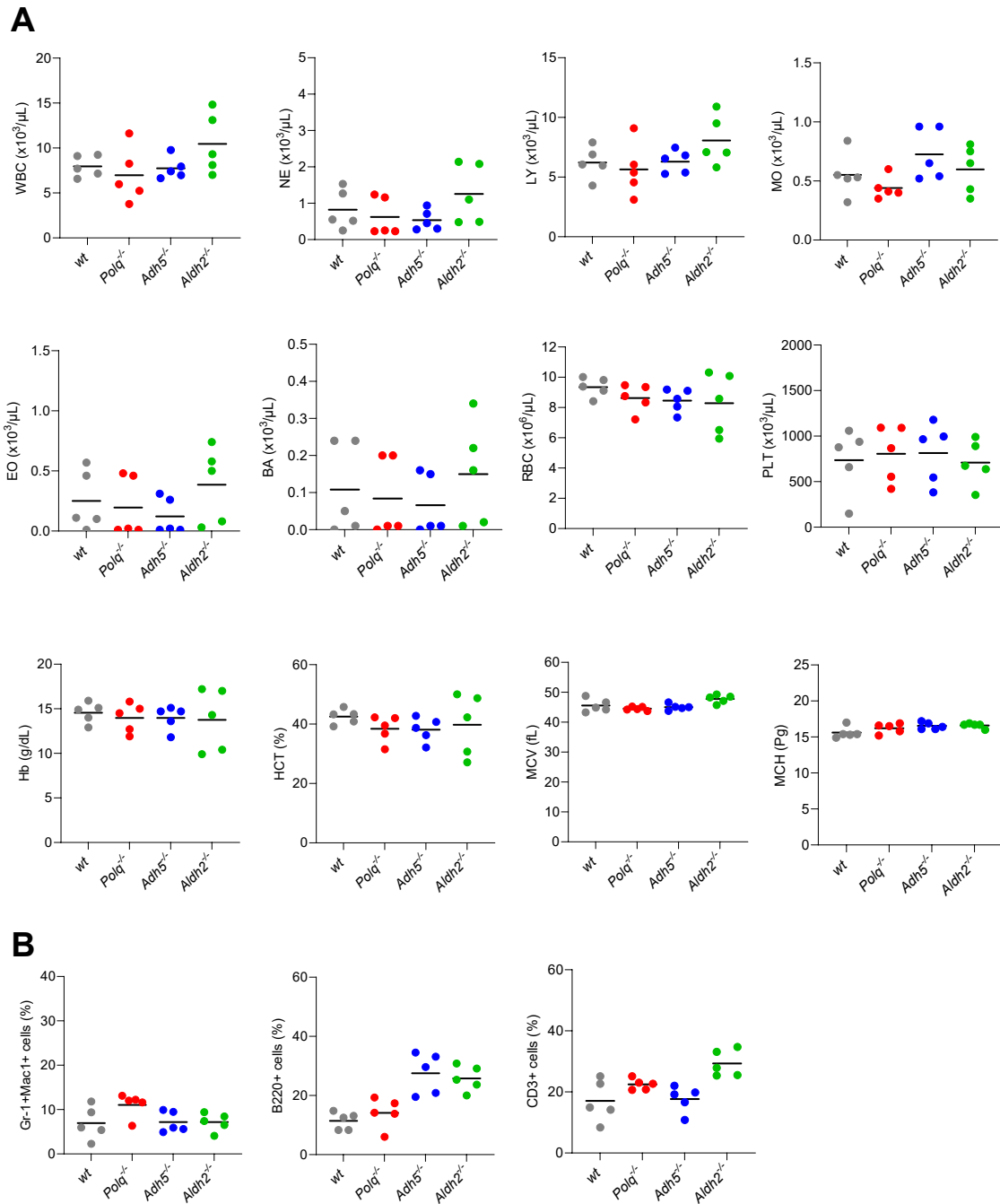


Figure 31. Analysis of peripheral blood from mice exposed to CSE for 3 months. (A) Hemavet analysis of peripheral blood cell parameters: white blood cells (WBC), neutrophils (NE), lymphocytes (LY), monocytes (MO), eosinophils (EO), basophils (BA), red blood cells (RBC), platelets (PLT), hemoglobin (Hb), hematocrit (HCT), mean corpuscular volume (MCV), and mean corpuscular hemoglobin (MCH). (B) Gr-1+Mac1+ cells, B220+ and CD3+ cells were detected in peripheral blood of the indicated genotypes

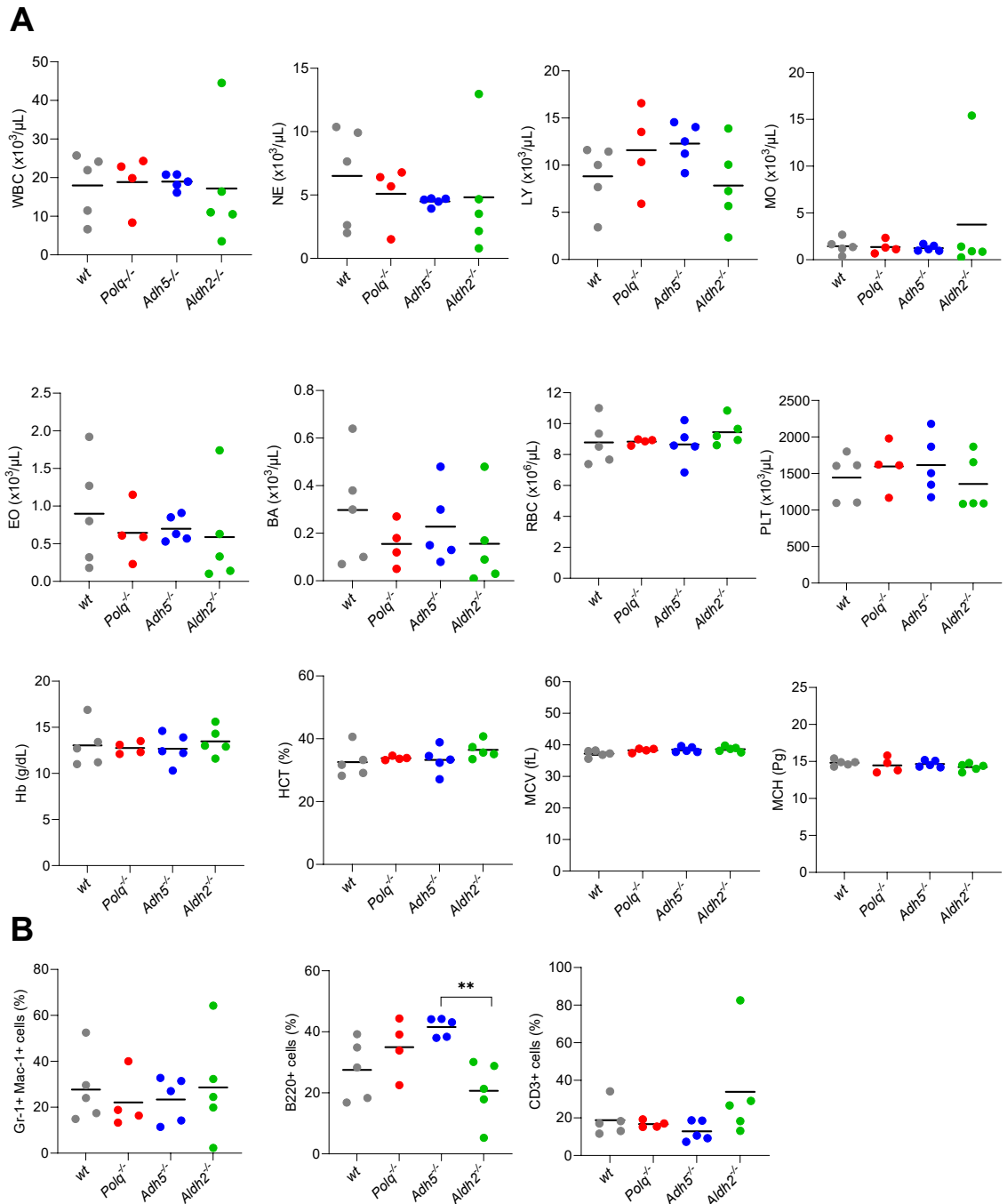


Figure 32. Analysis of peripheral blood from aged mice initially exposed to CSE. (A) Hemavet analysis of peripheral blood cell parameters: white blood cells (WBC), neutrophils (NE), lymphocytes (LY), monocytes (MO), eosinophils (EO), basophils (BA), red blood cells (RBC), platelets (PLT), hemoglobin (Hb), hematocrit (HCT), mean corpuscular volume (MCV), and mean corpuscular hemoglobin (MCH). (B) Gr-1+Mac-1+ cells, B220+ and CD3+ cells were detected in peripheral blood of the indicated genotypes.

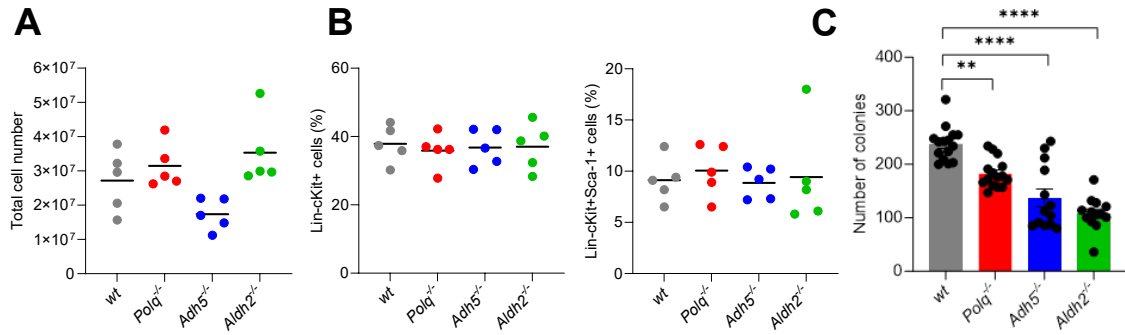


Figure 33. Analysis of bone marrow from aged mice initially exposed to CSE. (A) Total BMC number after isolation from femur bones. (B) HSPC populations using flow cytometry. (C) Data presented as mean number of colonies \pm SEM.

To investigate this further, wild-type, *Polq*^{-/-}, *Adh5*^{-/-}, *Aldh2*^{-/-}, *Adh5*^{-/-};*Polq*^{-/-}, and *Aldh2*^{-/-}; *Polq*^{-/-} Lin-cKit⁺ BMCs were exposed to cigarette smoke *in vitro*. Cells were exposed to different concentrations of cigarette smoke extract (CSE) for 24 hours and then plated in methylcellulose to assess their clonogenic capacity. Surprisingly, all genetic knockout groups were more sensitive to the CSE compared to the wild type (Figure 34). Levels of DSBs were not assessed in these cells either.

While we did not see any abnormal hematological phenotypes develop in our wild-type and single knockout mice, we wanted to determine whether there may be any clinical relevance in studying ADH5, ALDH2, and Polθ in the context of lung cancer as smoking is a major risk factor for lung cancer, accounting for about 80-90% of cases. According to TCGA data, there was a statistically significant negative correlation between *ADH5* and *POLQ* as well as *ALDH2* and *POLQ* in lung adenocarcinoma (LUAD), indicating a similar cooperation and correlative relationship also seen in AML (Figure 35). Furthermore, it may be clinically relevant to investigate the relationship between ADH5, ALDH2, and Polθ in the context of solid tumors such as LUAD.

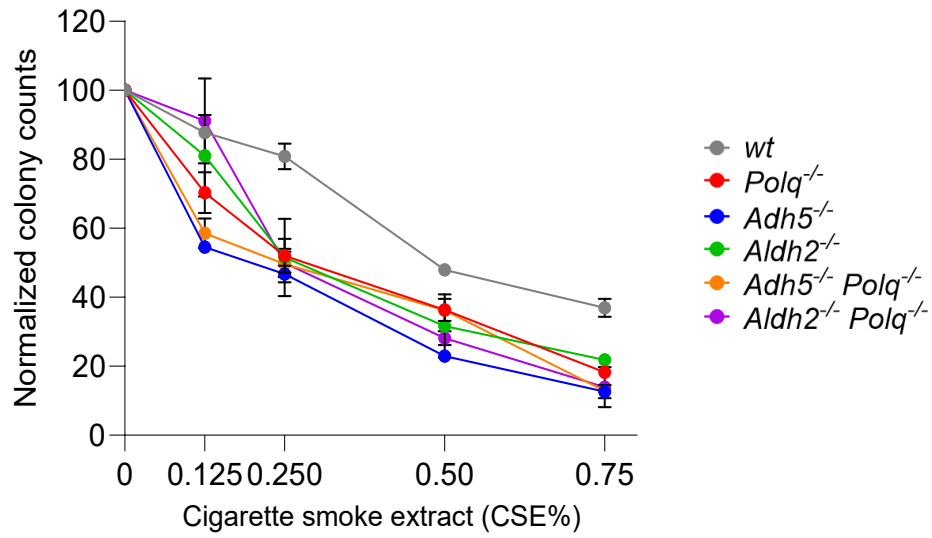


Figure 34. Genetic knockout mice are more sensitive to cigarette smoke *in vitro*.

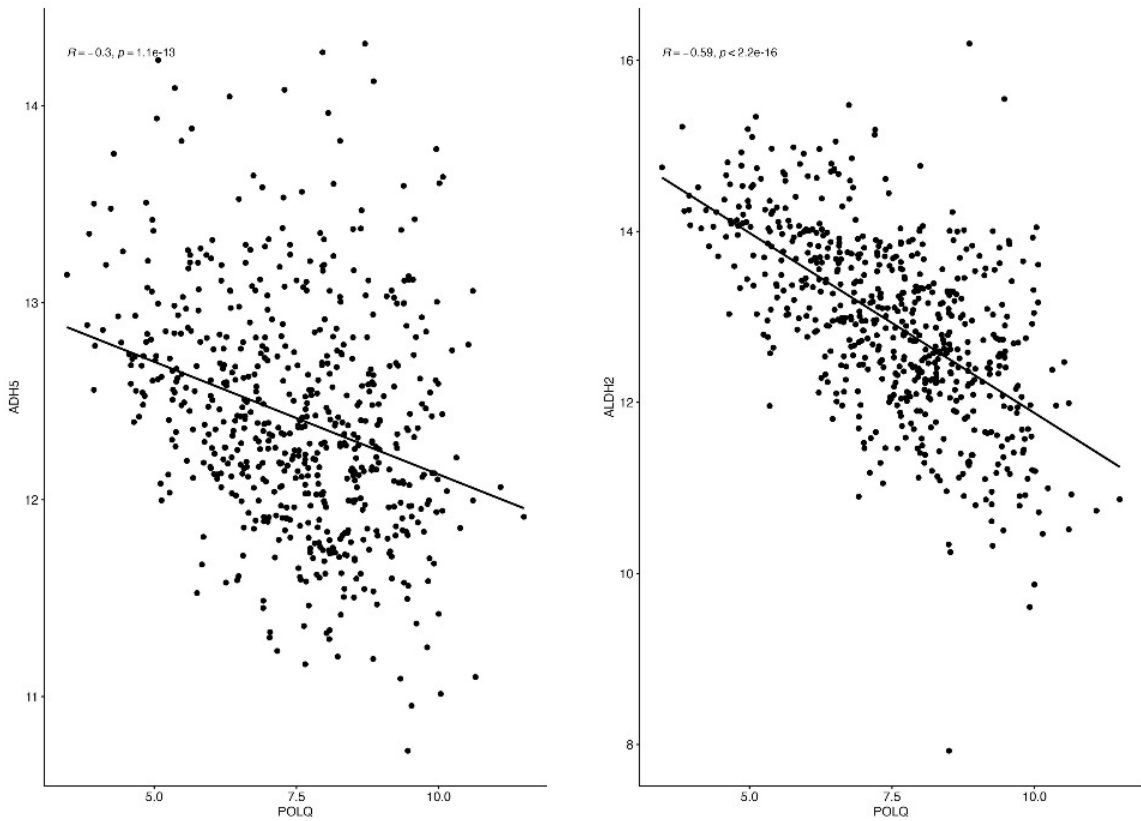


Figure 35. Correlation between *ADH5/POLQ* and *ALDH2/POLQ* expression in LUAD. Data generated from TCGA datasets.

CHAPTER 4

DISCUSSION

Dual targeting of Pol θ and formaldehyde detoxifiers (ADH5 or ALDH2)

Formaldehyde is ubiquitously present within the environment and our cells, where it can easily react with DNA, generating DNA damage. The accumulation of formaldehyde-mediated DNA damage can overwhelm DNA repair pathways, leading to genomic instability. Therefore, formaldehyde is a constant physiological threat to HSPCs. To maintain an intracellular balance of formaldehyde, cells must effectively break down formaldehyde and efficiently repair formaldehyde-mediated DNA damage. ADH5 and ALDH2 function as formaldehyde detoxifiers while the FA pathway is crucial in repairing formaldehyde-mediated ICLs, providing “2-Tier” protection in hematopoietic cells (21, 23, 25, 28, 37). Formaldehyde-mediated DNA lesions are not equally genotoxic and mutagenic, so it’s important to understand which lesions are responsible for formaldehyde genotoxicity and how they play a role in diseases such as leukemia.

In this study, we identified TMEJ as a DNA repair pathway that cooperates with ADH5 and ALDH2 to repair genotoxic formaldehyde-mediated DPC-associated DSBs, providing an additional layer (Tier-3) of protection in normal HSPCs challenged by elevated levels of formaldehyde (Figure 36). This protective mechanism was altered in malignant HSPCs and can be therapeutically exploited. We discovered ADH5 and ALDH2 inhibitors enhance the anti-leukemia effect of Pol θ inhibitors by inducing DNA damage and apoptosis in cell-based assays. We confirmed this anti-cancer activity by treating OTK-positive patient samples *in vitro* and using a xenograft mouse model of HR-deficient leukemia. Together these data suggest that targeting formaldehyde protective pathways may be a viable therapeutic strategy in treating cancers such as AML.

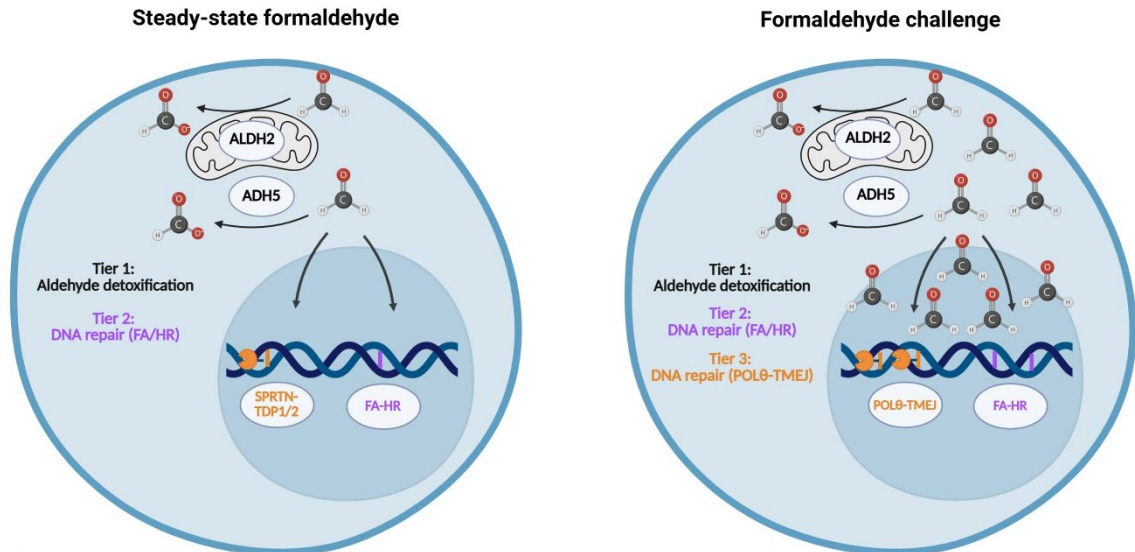


Figure 36. Newly discovered “3-Tier” protective mechanism in hematopoietic cells. Previously described “2-Tier” system (left). HSPCs challenged by elevated levels of formaldehyde can accumulate DPC-associated DSBs, another genotoxic lesion, repaired by TMEJ (3rd tier of protection) (right).

To fully understand the cooperativity between ADH5, ALDH2 and Polθ in normal HSPCs, *Adh5^{-/-};Polq^{-/-}* and *Aldh2^{-/-};Polq^{-/-}* double knockout mice were generated. With a focus on the hematopoietic system, a comprehensive hematological analysis was performed to determine the hematological landscape of our *Adh5^{-/-};Polq^{-/-}* and *Aldh2^{-/-};Polq^{-/-}* mice. Overall, no striking differences were found in the different blood cell populations when comparing double knockouts to single knockouts and wild-type counterparts, in both young and aged mice (Figures 10, 11). H&E staining of organs from aged mice showed no abnormalities in hematopoietic organs such as the spleen and bone marrow, but increased infiltrates in the lungs and kidney were observed in both double knockout groups (Figure 9). There were no defects in the clonogenic capacity of double knockout cells, corresponding with no significant changes in levels of DSBs in young and aged mice (Figures 12, 13). Mouse models have been an important tool to study formaldehyde genotoxicity and how formaldehyde contributes to bone marrow failure syndromes such as FA and ADDS. Generation of *Adh5^{-/-};Fancd2^{-/-}* and *Aldh2^{-/-};Fancd2^{-/-}*

mice led to the discovery that formaldehyde is a hematopoietic genotoxin and HSPCs rely on ADH5, ALDH2 and the FA DNA repair pathway to maintain normal hematopoietic functions (23, 37, 45). *Adh5^{-/-};Aldh2^{-/-}* mice demonstrated hematopoietic cells endogenously generate lethal amounts of formaldehyde, further revealing the genotoxic nature of formaldehyde and that ADH5 and ALDH2 both prevent formaldehyde genotoxicity (25, 28). In contrast, *Adh5^{-/-};Polq^{-/-}* and *Aldh2^{-/-};Polq^{-/-}* mice did not develop cancer or exhibit signs of bone marrow failure. The lifespan of *Adh5^{-/-};Polq^{-/-}* and *Aldh2^{-/-};Polq^{-/-}* mice was comparable to single knockout and wild-type. Therefore, it seems unlikely that Polθ is needed to support ADH5 and ALDH2 in protecting murine HSPCs from physiological levels of formaldehyde, especially with HR and FA DNA repair pathways intact. Simultaneous loss of Polθ and ADH5 or ALDH2 did not adversely affect normal HSPCs, making them ideal targets for synthetic lethality in a cancer-specific context. Furthermore, the work here established a mouse model system that allowed us to better understand the cooperation between ADH5, ALDH2, and Polθ in normal HSPCs.

Mice carrying constitutive *Polq* knockout combined with *Adh5* or *Aldh2* knockout might develop compensatory mechanisms, masking cooperativity in protecting HSPCs from formaldehyde. Therefore, we decided to challenge our genetic knockout mice with increased levels of formaldehyde. Wild-type, *Polq^{-/-}*, *Adh5^{-/-}*, *Aldh2^{-/-}*, *Adh5^{-/-};Polq^{-/-}* and *Aldh2^{-/-};Polq^{-/-}* mice were treated with methanol. After methanol treatment, elevated levels of DSBs were observed in *Adh5^{-/-};Polq^{-/-}* and *Aldh2^{-/-};Polq^{-/-}* Lin-cKit⁺ BMCs, directly correlating with a significant decrease in clonogenic capacity (Figure 15). *In vivo* data was confirmed by treating Lin-cKit⁺ BMCs with various concentrations of formaldehyde *in vitro*. High levels of DPCs and DSBs were observed along with severely compromised clonogenic capacity in *Adh5^{-/-};Polq^{-/-}* and *Aldh2^{-/-};Polq^{-/-}* Lin-cKit⁺ BMCs (Figure 17). Together, this data show that ADH5 and ALDH2 do cooperate with Polθ to protect HSPCs

from high levels of formaldehyde. One limitation of our *in vivo* methanol challenge is that formaldehyde levels were not measured in the bone marrow prior to or after methanol injections, so it cannot be said with certainty that high levels of formaldehyde made it to the bone marrow. Formaldehyde is a small, highly reactive molecule with an estimated half-life of around 90 seconds, and its rapid metabolism makes it a difficult molecule to accurately quantify in cells (138). However, we were able to confirm increased levels of formaldehyde in the blood before and after methanol injections using a previously established and commercially available formaldehyde detection kit (124).

The causal relationship between formaldehyde and leukemia has remained unclear, due to the lack of evidence surrounding the biological mechanisms underlying formaldehyde toxicity in the bone marrow, the site of leukemia induction (115). The most common route of formaldehyde exposure is inhalation and there is uncertainty whether inhaled formaldehyde can lead to toxicity in distant organs such as the bone marrow. Recent studies have shown that inhaled formaldehyde in rodents led to the accumulation of DPCs in the bone marrow, giving biological plausibility to the idea that inhaled formaldehyde can reach the bone marrow (115). Moreover, formaldehyde-mediated DPCs and DPC-associated DSBs may be potential genotoxic lesions and provide a plausible biological mechanism for how formaldehyde can target the bone marrow, leading to genome instability and the development of myeloid malignancies such as myeloproliferative neoplasms (MPN) or acute myeloid leukemia (AML). Our study confirms the findings of other research groups that show exposure to formaldehyde can impact bone marrow and cause the accumulation of DPCs and DSBs. Whether the accumulation of these lethal lesions contributes to or is the result of malignant transformation of leukemia cells still needs to be understood.

In leukemia cells, OTKs upregulate the 1C cycle, increasing intracellular levels of formaldehyde and DPC-associated DSBs. To combat this, OTK-positive leukemia cells increase the expression of Polθ, counteracting formaldehyde genotoxicity by repairing formaldehyde-mediated DPC-associated DSBs through its TMEJ activities (124). Mechanistically, Polθ is upregulated through the ERK1/2 serine/threonine kinase-dependent inhibition of c-CBL E3 ligase-mediated ubiquitination of Polθ, preventing its proteasomal degradation (124). This protective mechanism in leukemia cells allows for the efficient repair of DPC-associated DSBs, avoiding cell death. There are multiple metabolic processes that produce intracellular formaldehyde, but it is unknown which processes contribute the most to the endogenous formaldehyde burden. It is well known the 1C cycle produces formaldehyde, but the work by Vekariya et. al establishes the 1C cycle as a significant source of formaldehyde in leukemic cells and Polθ is an important cellular defense against the toxic effects of metabolically induced DNA damage. Additionally, TCGA analysis found low mRNA expression of *ADH5* and *ALDH2* in AML while *POLQ* expression was very high (Figure 26). This finding indicates that reduced detoxification of formaldehyde likely contributes to increased levels of formaldehyde and the upregulation of *POLQ* in AML indicates a functional requirement for *POLQ* to limit formaldehyde toxicity, making it a viable therapeutic target in treating leukemia. Therefore, we wanted to know if we could exploit this AML-specific vulnerability by targeting Polθ in combination with ADH5 or ALDH2, generating formaldehyde toxicity in leukemia cells.

Since OTKs increase intracellular levels of formaldehyde in leukemia cells, wild-type, *Polq*^{-/-}, *Adh5*^{-/-}, *Aldh2*^{-/-} *Adh5*^{-/-}; *Polq*^{-/-} and *Aldh2*^{-/-}; *Polq*^{-/-} Lin-cKit⁺ BMCs were transfected with either FLT3(ITD) or JAK2(V617F) to determine whether the presence of an OTK induces formaldehyde toxicity. *Adh5*^{-/-}; *Polq*^{-/-} and *Aldh2*^{-/-}; *Polq*^{-/-} Lin-cKit⁺ BMCs infected with FLT3(ITD) exhibited slow growth and decreased clonogenic capacity

compared to infected single knockout and wild-type counterparts (Figure 23A, 23C). Meanwhile, only *Aldh2*^{-/-};*Polq*^{-/-} Lin-cKit⁺ BMCs infected with JAK2(V617F) showed slower growth and decreased clonogenic capacity compared to the respective infected single knockout and wild-type counterparts (Figure 23B, 23D). This data suggest the addition of an OTK increases the formaldehyde burden in cells, enhancing the synthetic lethality, like what was observed when our genetic knockout mice and Lin-cKit⁺ BMCs were treated with methanol and formaldehyde, respectively. To explore this in a clinically relevant context, two AML patient samples with leukemic clones harboring FLT3(ITD) and two MPN patient samples with leukemic clones harboring JAK2(V617F) were treated with ADH5i + Polθi or ALDH2i + Polθi combinations. AML patient cells with FLT3(ITD) showed increased sensitivity to both ADH5i + Polθi and ALDH2i + Polθi and the anti-leukemic effect was enhanced with the addition of an FDA approved FLT3(ITD) inhibitor, quizartinib (Figure 24A, 24B). MPN patients treated with ADH5i + Polθi and ALDH2i + Polθi showed increased sensitivity to ALDH2i + Polθi while showing only moderate sensitivity to the ADH5i + Polθi, corroborating our mBMC data. However, with the addition of ruxolitinib, an FDA approved JAK1/2i, both combinations exhibited enhanced sensitivity (Figure 25A, 25B). All together, these data indicate the addition of ADH5i or ALDH2i enhance the anti-leukemic effects of Polθi, making ADH5 and ALDH2 additional therapeutic options for OTK-positive MPN and AML patients, especially when used with an OTK inhibitor. Additional *in vitro* and *in vivo* testing of patient samples will be essential to further validate their anti-cancer effects in leukemia. The 1C cycle is a significant source of formaldehyde in OTK-positive leukemias and should be explored in other cancers as well.

Polθ exhibits synthetic lethality with various HR repair genes, so we wanted to determine whether ADH5 or ALDH2 inhibitors would enhance the anti-cancer effects of Polθ inhibitors in HR-deficient cells. Nalm6 parental and Nalm6-*RAD54*^{-/-} are isogenic

acute lymphoblastic leukemia (ALL) cell lines where Nalm6-*RAD54*^{-/-} are HR-deficient, making them an ideal cell line pair to investigate the therapeutic relevance of ADH5 and ALDH2 inhibitors (92, 93, 133, 134). *In vitro* drug treatment showed strong anti-cancer effects such as elevated levels of DSBs and increased levels of apoptosis in Nalm6-*RAD54*^{-/-} cells treated with ADH5i + Polθi and ALDH2i + Polθi when compared to respective single inhibitor and control treatments (Figure 27, 28). To verify this, SCID-beige mice were injected with Nalm6-*RAD54*^{-/-} cells and subsequently treated with vehicle, Polθi, ADH5i, ALDH2i, ADH5i + Polθi or ALDH2i + Polθi. After drug treatment was completed, analysis of human-CD19⁺ showed no leukemic burden in mice treated with both ADH5i + Polθi or ALDH2i + Polθi (Figure 30A). The median survival time (MST) of mice treated with both drug combinations was significantly longer than mice receiving single inhibitor or vehicle treatment (Figure 30B). Overall, the work outlined here demonstrates that the simultaneous targeting of Polθ and ADH5 or ALDH2 is an effective anti-cancer strategy, inducing formaldehyde toxicity in a highly selective manner, taking advantage of cancer-specific alterations in aldehyde protective mechanisms.

Disulfiram, N6022, ART558 and RP-6685 were the four inhibitors used for all experiments. Disulfiram is an ALDH inhibitor that is FDA approved to treat alcoholism. Taking disulfiram does not cure alcoholism but is meant to discourage people from drinking as it will cause extreme sickness if taken while consuming alcohol due to a build-up of acetaldehyde (133). Currently, disulfiram is being investigated as a potential anti-cancer agent with many pre-clinical studies showing it induces oxidative stress, DNA damage, and apoptosis in leukemia cells (139, 140). In one study, disulfiram inhibited the viability of many acute lymphoblastic and myeloid leukemia cell lines as well as patient-derived cells from patients with treatment-resistant disease. These leukemia cells also exhibited greater sensitivity to disulfiram when compared to solid cancer cell lines. (139). Our work

also demonstrated a strong anti-cancer effect in leukemia cells treated with disulfiram, but in combination with a Pol θ inhibitor. Currently, there are multiple ongoing clinical trials investigating disulfiram for the treatment of cancers including breast, lung, prostate, melanoma, and liver (133). N6022 is an ADH5 inhibitor that is not an FDA approved drug, but pre-clinical and clinical testing have shown its safety and efficacy in treating patients with asthma, a respiratory illness (134). While it is not approved to treat any type of cancer, we showed that N6022 exhibits a strong anti-cancer effect in combination with a Pol θ inhibitor in leukemia cells. There are many active clinical trials testing different Pol θ inhibitors in the treatment of various solid tumors, specifically those with HR-deficiency or DDR alterations that are resistant to PARP inhibitors. Based on the work outlined here, ADH5 or ALDH2 inhibitors in combination with a Pol θ inhibitor should be considered for clinical studies in the treatment of leukemia.

ADH5 and ALDH2 are the main formaldehyde detoxifiers in hematopoietic cells (25, 28). However, in humans, there are 19 ALDH and 7 ADH genes that have a range of functions, including the detoxification of many alcohols and aldehydes (21, 141, 142). Members of the ALDH family can metabolize highly reactive aldehydes, protecting cells from aldehyde-related toxicity. The ALDH superfamily may also play a role in other vital cellular functions such as cell proliferation, differentiation and survival, and the synthesis of retinoic acid (142). *Adh5*^{-/-} and *Aldh2*^{-/-} single knockout mice exhibit no abnormal phenotypes, suggesting some functional redundancy. While ALDH and ADH members have some overlapping functions, each enzyme has distinct substrate specificities with differences in gene expression within tissues. Whether there are other ALDH or ADH enzymes that protect blood cells from formaldehyde or acetaldehyde has not been studied. ALDH1A1 is an ALDH that helps convert retinal to retinoic acid (RA) and is known to be upregulated in cancer stem cells and associated with drug resistance (143). A recent

study found ALDH1A1 enhances DNA repair and PARPi resistance by upregulating Polθ through the retinoic acid (RA) pathway in *BRCA2*-mutated ovarian cancer cells. The production of RA activates the RA pathway which regulates gene expression and is responsible for the transcriptional activation of the *POLQ* gene (143). This study indicates ALDH1A1 may be a viable therapeutic target for overcoming PARPi resistance in solid tumors such as ovarian cancer. There may be other ALDH or ADH-mediated protective mechanisms that influence the expression of DNA repair proteins such as Polθ, promoting drug resistance and cell survival.

ADH5 and ALDH2 work in parallel to detoxify formaldehyde with ADH5 located in the cytosol and ALDH2 in the mitochondria. ALDH2 oxidizes formaldehyde to produce formate whereas ADH5 cannot directly oxidize free formaldehyde (25, 28, 38). Instead, formaldehyde reacts with glutathione (GSH), resulting in the formation of the intermediate product S-hydroxymethyl-GSH (HSMGSH) (26). HSMGSH is a substrate for ADH5, and this intermediate molecule is oxidized by ADH5 to yield formate (26). Since GSH is a key player in ADH5-mediated formaldehyde detoxification and protects cells from formaldehyde-induced toxicity, we wanted to determine whether GSH impacts the synthetically lethal effects seen in *Adh5^{-/-};Polq^{-/-}* and *Aldh2^{-/-};Polq^{-/-}* cells challenged by high levels of formaldehyde (methanol, formaldehyde, OTKs, inhibitors). To do this, *Adh5^{-/-};Polq^{-/-}* and *Aldh2^{-/-};Polq^{-/-}* Lin-cKit⁺ BMCs were treated with various concentrations of formaldehyde ± L-BSO, a potent GSH synthesis inhibitor. *Aldh2^{-/-};Polq^{-/-}* Lin-cKit⁺ BMCs were more sensitive to formaldehyde with the addition of L-BSO while *Adh5^{-/-};Polq^{-/-}* exhibited modest sensitivity with the addition of L-BSO (Figure 18). Inhibiting GSH synthesis was redundant in *Adh5^{-/-};Polq^{-/-}* cells whereas in *Aldh2^{-/-};Polq^{-/-}*, it shut off ADH5-mediated formaldehyde detoxification which is why the effect was much greater. Loss of GSH exacerbated the synthetically lethal effects of formaldehyde, especially in *Aldh2^{-/-};Polq^{-/-}*;

Polq^{-/-} cells with intracellular levels of formaldehyde likely reaching toxic levels, resulting in cell death. GSH is important not only in maintaining intracellular formaldehyde levels but also in maintaining the redox balance as an antioxidant, helping neutralize harmful reactive oxygen species (ROS) and free radicals (26). Proper formaldehyde detoxification requires a strict redox balance achieved by GSH, minimizing aldehyde and oxidative-related DNA damage.

Formaldehyde and lung cancer

There are many intra- and extracellular sources of formaldehyde, but how much each of them contributes to the total formaldehyde burden in our cells remains to be seen. It is unknown whether the formaldehyde burden primarily comes from intracellular metabolism or if formaldehyde derived from other tissues is responsible for most of the intracellular formaldehyde. We have shown here and in previous work that the 1C cycle is a meaningful source of formaldehyde in leukemia cells and should be explored further in other cancer types (124). Additional formaldehyde challenges would provide more understanding into whether exogenously generated formaldehyde contributes to carcinogenesis and if ADH5, ALDH2 and Polθ cooperate to protect cells from these environmental threats. Specific *in vivo* formaldehyde challenges could involve exposure to cigarette smoke or adhering to a specialized diet (i.e. serine/glycine/glucose rich diet).

Cigarette smoking is a known risk factor for AML (144). Carcinogens in cigarette smoke include formaldehyde and acetaldehyde which are found in high abundance and are deposited within the upper respiratory tract (145, 146). Absorption of these aldehydes happens in the lungs and can diffuse through the bloodstream, potentially affecting blood cells and distant tissues. However, whether aldehydes present in cigarette smoke make it to the bone marrow and lead to aldehyde-mediated toxicity remains unknown. The contribution of cigarette smoke to the intracellular formaldehyde burden in smokers vs.

non-smokers is also an important question. Cigarette smoke could be another biological mechanism for aldehyde-induced toxicity, specifically in HSPCs, contributing to carcinogenesis. Now that we have established TMEJ as a DNA repair pathway that cooperates with ADH5 and ALDH2 in protecting normal and malignant HSPCs from various formaldehyde challenges, we wanted to simulate real-world conditions and expose our genetic knockout mice and cells to cigarette smoke.

To understand whether this “3-Tier” system protects HSPCs from the deleterious effects of cigarette smoke, wild-type, *Polq*^{-/-}, *Adh5*^{-/-} and *Aldh2*^{-/-} were exposed to cigarette smoke 2 hours a day, 5-days per week for a total of 3 months. This protocol mimics the cigarette smoke exposure of a primary smoker *in vivo*. There was not any age matched *Adh5*^{-/-};*Polq*^{-/-} and *Aldh2*^{-/-};*Polq*^{-/-} mice to use for this specific experiment. After 3 months, blood samples were taken, and a full hematological analysis was performed. There were no significant differences in the different blood cell populations when comparing single knockout mice to wild-type (Figure 31). With no hematological abnormalities or adverse phenotypes observed in the mice, they were kept under normal housing conditions. Leukemia is more common in older people, so another hematological analysis was performed when the mice reached 18+ months old which would translate to a 56–69-year-old human. Again, there were no significant changes observed in the different blood cell parameters or in the HSPC/HSC populations (Figure 32, 33A-B). Remarkably, the clonogenic capacity was severely compromised in *Polq*^{-/-}, *Adh5*^{-/-}, *Aldh2*^{-/-} mice when compared to wild-type Lin-cKit⁺ BMCs (Figure 33C). Future studies should include double knockout mice to fully understand if cigarette smoke contributes to aldehyde-induced toxicity in HSCs. In the meantime, wild-type, *Polq*^{-/-}, *Adh5*^{-/-}, *Aldh2*^{-/-}, *Adh5*^{-/-};*Polq*^{-/-} and *Aldh2*^{-/-};*Polq*^{-/-} Lin-cKit⁺ BMCs were treated with cigarette smoke *in vitro*. Preliminary results showed *Polq*^{-/-}, *Adh5*^{-/-}, *Aldh2*^{-/-}, *Adh5*^{-/-};*Polq*^{-/-} and *Aldh2*^{-/-};*Polq*^{-/-} Lin-cKit⁺ BMCs

were more sensitive to cigarette smoke in a dose dependent manner when compared to wild-type (Figure 34). Biological replicates are needed, but the data suggests *ADH5*, *ALDH2* and/or *Polθ* may play a role in protecting HSPCs from cigarette smoke. The primary focus of this study was to better understand cigarette smoke and its contribution to aldehyde-mediated toxicity in HSPCs. We did not perform any analysis on other tissues, but we did collect organs from the aged mice exposed to cigarette smoke. Histopathological analysis of organs, including the lungs, could give more insight into the role of this aldehyde protective mechanism in protecting lung cells from cigarette smoke and/or formaldehyde genotoxicity.

POLQ is overexpressed in lung adenocarcinoma (LUAD) and is associated with increased mutational load (147). TCGA data showed a statistically significant negative correlation between the mRNA expression of *ADH5* and *POLQ* as well as *ALDH2* and *POLQ* in LUAD (Figure 35). For example, the low expression of *ADH5* and *ALDH2* coincides with high expression of *POLQ*, like what we uncovered in AML, indicating a potential cooperative relationship, but this time in LUAD. Patients with lung cancer harbored higher levels of formaldehyde in their lungs when compared to normal tissues and a significant increase in formaldehyde and acetaldehyde were detected in the exhaled air of patients with lung cancer when compared to healthy people (29). Based on this data, the role of formaldehyde genotoxicity in lung cancer should be studied more in depth.

Acetaldehyde and cancer

Modifiable risk factors for cancer are lifestyle choices that can be altered to reduce the risk of developing cancer. Modifiable risk factors include smoking, alcohol consumption, physical activity, poor diet and obesity. Cigarette smoking is one of the most prevalent environmental causes of cancer. About 1.3 billion people use tobacco worldwide and it is the leading cause of preventable deaths globally. Cigarette smoking alone causes

approximately 30% of cancer deaths in the United States while alcohol use accounts for approximately 5% of all cancers and 4% of all cancer deaths. Alcohol consumption is the leading risk factor for cancers of the upper aerodigestive tract, liver and colon (148). Most of the toxic effects of alcohol are likely caused by the oxidation of ethanol which produces the highly reactive metabolite acetaldehyde (45). Acetaldehyde is a byproduct of alcohol metabolism and is also present in tobacco smoke. It is also classified as a human carcinogen due to its ability to react with DNA, creating DNA adducts and other deleterious DNA lesions like DPCs and DSBs (45).

One proposed mechanism for alcohol-induced cancer includes the formation of acetaldehyde-mediated DNA damage, creating genomic instability. The liver is the main organ responsible for metabolizing ethanol where the oxidation of ethanol to acetaldehyde involves cytosolic alcohol dehydrogenase 1 (ADH1). Microsomal cytochrome P450 isoenzymes CYP2E1, 1A2 and 3A4 also contribute to ethanol oxidation, primarily after chronic ethanol consumption (149). Catalase may be the primary enzyme responsible for alcohol detoxification in the brain, kidney and heart (149). Acetaldehyde is then metabolized to acetate by ALDH2 and acetate escapes to the blood and is further metabolized to CO₂ via tricarboxylic acid cycle in tissues (21, 28, 149).

ALDH2 plays a crucial role in acetaldehyde metabolism and allelic variations in the gene encoding ALDH2 results in defective enzymatic activity. People with the *ALDH2*2* polymorphism have a functionally inactive ALDH2, so their bodies cannot break down acetaldehyde, leading to its accumulation and formation of acetaldehyde-mediated DNA damage (150). Many Japanese studies show a correlation between the *ALDH2*2* polymorphism and increased risk of upper aerodigestive cancers (47, 48, 151). It has also been shown that lymphocytes from drinkers with the inactive *ALDH2*2* have higher frequencies of sister chromatid exchanges than those with the active form *ALDH2*1*,

indicating DNA damage and chromosome instability (150). Like formaldehyde, it has the capacity to bind to DNA and proteins, forming DNA adducts and DNA-protein crosslinks (152-154). Overall, *ALDH2**2 has been associated with an increased risk of developing cancers including esophageal, liver, colorectal and head and neck (151). ALDH2 dysfunction has also been shown to contribute to other human diseases including cardiovascular diseases, diabetes and neurodegenerative disorders (49, 155). Together, the evidence suggests that ALDH2 dysfunction contributes to a variety of human diseases including cancer. However, like formaldehyde, acetaldehyde protective mechanisms and how they contribute to carcinogenesis remains largely unknown. The generation of acetaldehyde-mediated DNA damage creates genome instability and may be one of the mechanisms that drive alcohol-related cancers.

Lung cancer is the leading cause of cancer-related deaths worldwide. It is characterized by genome instability which can be induced by factors such as cigarette smoke and reactive metabolites (156). ALDH2 expression was found to be low in lung tumor tissues and ALDH2 repression in lung cancer is associated with poor prognosis, implicating ALDH2 in the progression of lung cancer (156). ALDH2 deficiency in lung cancer cells led to the accumulation of acetaldehyde and increased DNA damage both *in vitro* and *in vivo* (156). Acetaldehyde can lead to various DNA lesions, including DPCs and DSBs that may be repaired by the FA DNA repair pathway (37, 45, 157). Acetaldehyde can also generate DNA base modifications, requiring base excision repair (BER) (158). BER is initiated by a DNA glycosylase that recognizes and removes the damaged DNA base. APE1, an AP endonuclease, cleaves the phosphodiester bond at the abasic site, generating a DNA strand break containing a hydroxyl residue at the 3'-end and deoxyribose phosphate at the 5'-end. DNA polymerase β fills the gap and adds the correct nucleotide and DNA repair is completed by XRCC1 and DNA ligase III α , sealing the ends

(158). Cells deficient in XRCC1 are characterized by reduced activity of BER and increased genomic instability due to an increase in unrepaired SSBs (159). ALDH2 is downregulated while XRCC1 expression is high in lung and liver cancers, a possible mechanism cancer cells employ to counteract the accumulation of acetaldehyde-mediated DNA damage brought on by the improper clearance of acetaldehyde (158). Loss of XRCC1 led to an increase in expression of ALDH2, suggesting these two proteins may cooperate in protecting lung cells from the accumulation of SSBs (158). This study uncovers a potential acetaldehyde protective mechanism that becomes dysregulated in lung cancer cells, providing a novel strategy in the treatment of lung cancer.

Loss of BER capabilities not only leads to genome instability, but also to metabolic changes which include alternations in serine biosynthesis and an increase in the activity of the 1C cycle, a known endogenous source of formaldehyde (160). BER deficiency may contribute to increased levels of intracellular formaldehyde, possibly leading to cell transformation and carcinogenesis (160). BER or more specifically XRCC1 could be a potential driver of not only acetaldehyde-mediated, but also formaldehyde-mediated genotoxicity.

Primary protection against acetaldehyde in hematopoietic cells is provided by ALDH2 while the FA DNA repair pathway is responsible for repairing acetaldehyde-mediated DNA damage, specifically ICLs (45). NHEJ and HR can also repair acetaldehyde-mediated DNA lesions (45). However, it remains unknown whether Pol θ and TMEJ cooperate with ALDH2 to protect HSPCs or other cells and tissues from acetaldehyde-mediated DPC-associated DSBs. If Pol θ does cooperate with ALDH2 in protecting HSPCs from acetaldehyde-mediated DPC-associated DSBs or other genotoxic lesions, it will establish an acetaldehyde specific protective mechanism that may help us identify novel therapeutic strategies in treating alcohol and tobacco-mediated cancers.

Identifying formaldehyde response genes

Synthetic lethality is one approach in identifying novel therapeutic targets for cancer. In this study, we identified cooperativity between the DNA repair protein Pol θ and formaldehyde detoxifying enzymes ADH5 and ALDH2 in protecting normal and malignant hematopoietic cells from formaldehyde-mediated DNA damage. This discovery uncovered a “3-Tier” protective mechanism that exists within hematopoietic cells. Tier-1 (ADH5/ALDH2), Tier-2 (FA pathway) and Tier-3 (TMEJ) protect against basal and/or high levels of formaldehyde and are supported by genetic studies described here. For example, the simultaneous knockout of Tier-1 (*Adh5*^{-/-}) and Tier-2 (*Fancd2*^{-/-}) members were lethal in mice, supporting the statement that Tier-1 + Tier-2 protects cells from the basal levels of formaldehyde. However, mice bearing the genetic knockout of Tier-1 (*Adh5*^{-/-}) and Tier-3 (*Polq*^{-/-}) members were viable, and toxicity was only detected when hematopoietic cells were challenged with increased levels of formaldehyde. So, TMEJ becomes essential when cells are exposed to increased levels of formaldehyde. We also found ADH5, ALDH2, and Pol θ to be promising therapeutic targets in treating OTK-positive and/or HR-deficient leukemias. The interdependence between Pol θ , ADH5, and ALDH2 in protecting leukemia cells from formaldehyde genotoxicity creates a unique translational opportunity. Identifying other DNA repair pathways that work together with ADH5 and ALDH2 may provide more insight into other aldehyde protective mechanisms.

HR repair proteins BRCA1/2 cooperate with FA DNA repair proteins such as FANCD2 to maintain genome stability through replication fork stabilization (81, 161). Loss of BRCA1/2 results in defective HR repair and replication stress, promoting genome instability. It is known that BRCA1/2 deficient tumors upregulate Pol θ and FANCD2 as a survival mechanism and simultaneous genetic inactivation of *Fancd2* and *Polq* in mice results in embryonic lethality (81). FANCD2 has been shown to promote TMEJ by

recruiting Pol θ to sites of DNA damage (161). This further demonstrates the cooperation between Tier-2 (FA DNA pathway) and Tier-3 (TMEJ) in protecting cells from genome instability.

In addition to FA proteins and Pol θ , exonuclease 1 (EXO1) has been identified as formaldehyde response gene and an important regulator of formaldehyde-mediated DPCs and ICLs (162). Loss of EXO1 leads to formaldehyde sensitivity, the accumulation of DPCs and DSBs and replication stress (162). Its exonuclease activity is sufficient in removing formaldehyde-mediated DPCs, making EXO1 another potential mechanism for DPC repair (162). EXO1 limits DNA replication stress and fork degradation induced by formaldehyde and may even work in parallel to the FA pathway at formaldehyde-stalled replication forks (162).

Transcriptional stress, brought on by DNA damage, activates transcription-coupled nucleotide excision repair (TC-NER) (163). Activation of TC-NER halts transcription while the DNA damage can be repaired. Cockayne syndrome is a rare, inherited genetic disorder that develops in patients that have mutations in *ERCC6* or *ERCC8* genes that encode Cockayne syndrome proteins (CSA, CSB) that are responsible for the activation of the TC-NER repair pathway (164, 165). Endogenous formaldehyde causes transcriptional stress and TC-NER becomes activated to protect cells from formaldehyde-induced transcriptional stress (164). An increase in formaldehyde-mediated DNA adducts were found in the brains and kidneys in mice lacking both *Adh5* and *Csb* while severe kidney dysfunction and neurodegeneration were also observed and the causes of death of these mice (164). This adverse phenotype in mice mimics the most prominent clinical manifestations in patients of this syndrome including growth failure and short stature, premature aging and neurological problems. This indicates that a 2-Tier system exists that protects brain and kidney cells against formaldehyde-mediated DPCs and ICLs (164, 165).

This also shows us that brain and kidney cells are more susceptible to formaldehyde-induced transcriptional stress, depending more on TC-NER than hematopoietic cells which rely on FA DNA repair pathway. It seems cells could be potentially equipped with organ-specific, multi-tiered mechanisms protecting them from formaldehyde toxicity.

The impact of formaldehyde on global cellular pathways has not been fully investigated and identifying cooperating metabolic enzymes and DNA repair pathways are crucial so we can better elucidate mechanisms in which our cells protect themselves from aldehyde-mediated DNA damage and toxicity. Identifying these highly efficient protection systems may be important not only in the context of severe genetic conditions, bone marrow failure syndromes and cancer, but also in neurological and cardiovascular diseases.

Conclusions

Formaldehyde is a highly reactive, genotoxic and carcinogenic molecule that causes more cancer than any other toxic air pollutant (94). In addition to the various environmental sources, our cells continuously produce formaldehyde as a natural byproduct of various metabolic processes. Over 40 years ago, formaldehyde was identified as a “possible human carcinogen,” but was later established as a “known human carcinogen.” It has been associated with an increased risk of multiple cancers including leukemia, but despite decades of research, the mechanisms in which formaldehyde contributes to carcinogenesis are still largely unknown.

Due to its high reactivity, formaldehyde can easily react with DNA and/or proteins to produce some of the most deleterious DNA lesions, including ICLs, DPCs and DSBs. To protect themselves from formaldehyde, hematopoietic cells rely on the strict coordination between metabolic enzymes and DNA repair pathways to maintain a homeostatic, intracellular balance of formaldehyde, limiting formaldehyde-mediated DNA

damage. Failure to efficiently break down formaldehyde and/or repair formaldehyde-mediated DNA damage results in genome instability. A “2-Tier” protective mechanism within hematopoietic cells was previously described where ADH5 and ALDH2 are responsible for the breakdown of formaldehyde (Tier-1) while the FA pathway repairs formaldehyde-mediated lesions such as ICLs (Tier-2). The work done by Atkins et. al. identified TMEJ as a DNA repair pathway that cooperates with the formaldehyde detoxifying enzymes ADH5 and ALDH2 in protecting HSPCs from lethal formaldehyde-mediated DPC-associated DSBs. Formaldehyde-mediated DNA damage is a serious threat to genome stability which is why hematopoietic cells rely on multiple DNA repair pathways to protect themselves with TMEJ providing an extra layer of protection (Tier-3).

The discovery of Pol θ cooperating with ADH5 and ALDH2 in protecting normal HSPCs from formaldehyde uncovered a potential clinical application in the treatment of leukemia. We found that ADH5 and ALDH inhibitors enhance the anti-leukemia effect of Pol θ inhibitors in both OTK-positive and HR-deficient leukemias. These specific leukemic cells rewire their metabolic and DNA repair networks with cells expressing high levels of Pol θ to ensure the repair of formaldehyde-mediated DNA damage that accumulates because of high levels of endogenously produced formaldehyde. Our work demonstrates that this reprogramming can be exploited therapeutically and could be a viable treatment option for patients with leukemia. Understanding the molecular mechanisms responsible for formaldehyde genotoxicity in hematopoietic cells gives us a novel strategy in selectively targeting cancer cells while sparing normal ones. With formaldehyde ubiquitously present in our lives, it has become a threat to our health and well-being. Therefore, it is essential for us to have a deeper insight into the cellular networks responsible for protecting cells from the toxic effects of formaldehyde. We are the first to establish a cooperative relationship between ADH5, ALDH2 and Pol θ in protecting both

normal and malignant HSPCs from high levels of formaldehyde-mediated DPC-associated DSBs and we anticipate many follow-up studies that will expand the current knowledge of formaldehyde and its contribution to carcinogenesis.

REFERENCES

1. Salthammer T, Mentese S, Marutzky R. Formaldehyde in the indoor environment. *Chem Rev.* 2010;110(4):2536-72.
2. Rosenkranz HS. Formaldehyde as a possible carcinogen. *Bull Environ Contam Toxicol.* 1972;8(4):242-4.
3. Salthammer T. The formaldehyde dilemma. *Int J Hyg Environ Health.* 2015;218(4):433-6.
4. Smith AE. Formaldehyde. *Occup Med (Lond).* 1992;42(2):83-8.
5. Salk JE, Krech U, Youngner JS, Bennett BL, Lewis LJ, Bazeley PL. Formaldehyde treatment and safety testing of experimental poliomyelitis vaccines. *Am J Public Health Nations Health.* 1954;44(5):563-70.
6. Mitkus RJ, Hess MA, Schwartz SL. Pharmacokinetic modeling as an approach to assessing the safety of residual formaldehyde in infant vaccines. *Vaccine.* 2013;31(25):2738-43.
7. Valverde-Santiago M, Pontel LB. Emerging mechanisms underlying formaldehyde toxicity and response. *Mol Cell.* 2025;85(11):2068-79.
8. Dorokhov YL, Shindyapina AV, Sheshukova EV, Komarova TV. Metabolic methanol: molecular pathways and physiological roles. *Physiol Rev.* 2015;95(2):603-44.
9. Skrzydlewska E. Toxicological and metabolic consequences of methanol poisoning. *Toxicol Mech Methods.* 2003;13(4):277-93.
10. Weng MW, Lee HW, Park SH, Hu Y, Wang HT, Chen LC, et al. Aldehydes are the predominant forces inducing DNA damage and inhibiting DNA repair in tobacco smoke carcinogenesis. *Proc Natl Acad Sci U S A.* 2018;115(27):E6152-E61.
11. Godish T. Formaldehyde exposures from tobacco smoke: a review. *Am J Public Health.* 1989;79(8):1044-5.
12. Baker RR. The generation of formaldehyde in cigarettes--Overview and recent experiments. *Food Chem Toxicol.* 2006;44(11):1799-822.

13. Ridpath JR, Nakamura A, Tano K, Luke AM, Sonoda E, Arakawa H, et al. Cells deficient in the FANC/BRCA pathway are hypersensitive to plasma levels of formaldehyde. *Cancer Res.* 2007;67(23):11117-22.
14. Nadalutti CA, Prasad R, Wilson SH. Perspectives on formaldehyde dysregulation: Mitochondrial DNA damage and repair in mammalian cells. *DNA Repair (Amst).* 2021;105:103134.
15. Lee ES, Chen H, Hardman C, Simm A, Charlton C. Excessive S-adenosyl-L-methionine-dependent methylation increases levels of methanol, formaldehyde and formic acid in rat brain striatal homogenates: possible role in S-adenosyl-L-methionine-induced Parkinson's disease-like disorders. *Life Sci.* 2008;83(25-26):821-7.
16. Hopkinson RJ, Hamed RB, Rose NR, Claridge TD, Schofield CJ. Monitoring the activity of 2-oxoglutarate dependent histone demethylases by NMR spectroscopy: direct observation of formaldehyde. *Chembiochem.* 2010;11(4):506-10.
17. Walport LJ, Hopkinson RJ, Schofield CJ. Mechanisms of human histone and nucleic acid demethylases. *Curr Opin Chem Biol.* 2012;16(5-6):525-34.
18. Tan W, Deans AJ. Formaldehyde Causes Bone Marrow Failure Linked to Transcriptional Reprogramming or Metabolic Deficiency. *Mol Cell.* 2020;80(6):935-7.
19. Shen X, Wang R, Kim MJ, Hu Q, Hsu CC, Yao J, et al. A Surge of DNA Damage Links Transcriptional Reprogramming and Hematopoietic Deficit in Fanconi Anemia. *Mol Cell.* 2020;80(6):1013-24.e6.
20. Porter DH, Cook RJ, Wagner C. Enzymatic properties of dimethylglycine dehydrogenase and sarcosine dehydrogenase from rat liver. *Arch Biochem Biophys.* 1985;243(2):396-407.
21. Wang M, Dingler FA, Patel KJ. Genotoxic aldehydes in the hematopoietic system. *Blood.* 2022;139(14):2119-29.
22. Burgos-Barragan G, Wit N, Meiser J, Dingler FA, Pietzke M, Mulderrig L, et al. Erratum: Mammals divert endogenous genotoxic formaldehyde into one-carbon metabolism. *Nature.* 2017;548(7669):612.
23. Pontel LB, Rosado IV, Burgos-Barragan G, Garaycochea JI, Yu R, Arends MJ, et al. Endogenous Formaldehyde Is a Hematopoietic Stem Cell Genotoxin and Metabolic Carcinogen. *Mol Cell.* 2015;60(1):177-88.

24. Cammalleri V, Pocino RN, Marotta D, Protano C, Sinibaldi F, Simonazzi S, et al. Occupational scenarios and exposure assessment to formaldehyde: A systematic review. *Indoor Air*. 2022;32(1):e12949.
25. Nakamura J, Holley DW, Kawamoto T, Bultman SJ. The failure of two major formaldehyde catabolism enzymes (ADH5 and ALDH2) leads to partial synthetic lethality in C57BL/6 mice. *Genes Environ*. 2020;42:21.
26. Umansky C, Morellato AE, Rieckher M, Scheidegger MA, Martinefski MR, Fernandez GA, et al. Endogenous formaldehyde scavenges cellular glutathione resulting in redox disruption and cytotoxicity. *Nat Commun*. 2022;13(1):745.
27. Li B, Sun C, Lin X, Busch W. The Emerging Role of GSNOR in Oxidative Stress Regulation. *Trends Plant Sci*. 2021;26(2):156-68.
28. Dingler FA, Wang M, Mu A, Millington CL, Oberbeck N, Watcham S, et al. Two Aldehyde Clearance Systems Are Essential to Prevent Lethal Formaldehyde Accumulation in Mice and Humans. *Mol Cell*. 2020;80(6):996-1012.e9.
29. Dorokhov YL, Sheshukova EV, Bialik TE, Komarova TV. Human Endogenous Formaldehyde as an Anticancer Metabolite: Its Oxidation Downregulation May Be a Means of Improving Therapy. *Bioessays*. 2018;40(12):e1800136.
30. Xu M, Tang H, Rong Q, Zhang Y, Li Y, Zhao L, et al. The Effects of Formaldehyde on Cytochrome P450 Isoform Activity in Rats. *Biomed Res Int*. 2017;2017:6525474.
31. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell*. 2000;100(1):57-70.
32. Mu A, Hira A, Mori M, Okamoto Y, Takata M. Fanconi anemia and Aldehyde Degradation Deficiency Syndrome: Metabolism and DNA repair protect the genome and hematopoiesis from endogenous DNA damage. *DNA Repair (Amst)*. 2023;130:103546.
33. Niraj J, Farkkila A, D'Andrea AD. The Fanconi Anemia Pathway in Cancer. *Annu Rev Cancer Biol*. 2019;3:457-78.
34. Rodriguez A, D'Andrea A. Fanconi anemia pathway. *Curr Biol*. 2017;27(18):R986-R8.
35. Stingle J, Bellelli R, Boulton SJ. Mechanisms of DNA-protein crosslink repair. *Nat Rev Mol Cell Biol*. 2017;18(9):563-73.

36. Langevin F, Crossan GP, Rosado IV, Arends MJ, Patel KJ. Fancd2 counteracts the toxic effects of naturally produced aldehydes in mice. *Nature*. 2011;475(7354):53-8.
37. Garaycochea JI, Crossan GP, Langevin F, Daly M, Arends MJ, Patel KJ. Genotoxic consequences of endogenous aldehydes on mouse haematopoietic stem cell function. *Nature*. 2012;489(7417):571-5.
38. Mu A, Hira A, Niwa A, Osawa M, Yoshida K, Mori M, et al. Analysis of disease model iPSCs derived from patients with a novel Fanconi anemia-like IBMFS ADH5/ALDH2 deficiency. *Blood*. 2021;137(15):2021-32.
39. Jin S, Chen J, Chen L, Histen G, Lin Z, Gross S, et al. ALDH2(E487K) mutation increases protein turnover and promotes murine hepatocarcinogenesis. *Proc Natl Acad Sci U S A*. 2015;112(29):9088-93.
40. Crabb DW, Edenberg HJ, Bosron WF, Li TK. Genotypes for aldehyde dehydrogenase deficiency and alcohol sensitivity. The inactive ALDH2(2) allele is dominant. *J Clin Invest*. 1989;83(1):314-6.
41. Harada S, Agarwal DP, Goedde HW. Aldehyde dehydrogenase deficiency as cause of facial flushing reaction to alcohol in Japanese. *Lancet*. 1981;2(8253):982.
42. Hira A, Yabe H, Yoshida K, Okuno Y, Shiraishi Y, Chiba K, et al. Variant ALDH2 is associated with accelerated progression of bone marrow failure in Japanese Fanconi anemia patients. *Blood*. 2013;122(18):3206-9.
43. Oka Y, Hamada M, Nakazawa Y, Muramatsu H, Okuno Y, Higasa K, et al. Digenic mutations in ALDH2 and ADH5 impair formaldehyde clearance and cause a multisystem disorder, AMeD syndrome. *Sci Adv*. 2020;6(51).
44. Mu A, Hira A, Matsuo K, Takata M. [Aldehyde degradation deficiency (ADD) syndrome: discovery of a novel fanconi anemia-like inherited BMF syndrome due to combined ADH5/ALDH2 deficiency]. *Rinsho Ketsueki*. 2021;62(6):547-53.
45. Garaycochea JI, Crossan GP, Langevin F, Mulderrig L, Louzada S, Yang F, et al. Alcohol and endogenous aldehydes damage chromosomes and mutate stem cells. *Nature*. 2018;553(7687):171-7.
46. Rosado IV, Langevin F, Crossan GP, Takata M, Patel KJ. Formaldehyde catabolism is essential in cells deficient for the Fanconi anemia DNA-repair pathway. *Nat Struct Mol Biol*. 2011;18(12):1432-4.

47. Chang JS, Hsiao JR, Chen CH. ALDH2 polymorphism and alcohol-related cancers in Asians: a public health perspective. *J Biomed Sci.* 2017;24(1):19.
48. Chen CH, Kraemer BR, Mochly-Rosen D. ALDH2 variance in disease and populations. *Dis Model Mech.* 2022;15(6).
49. Takeuchi F, Yokota M, Yamamoto K, Nakashima E, Katsuya T, Asano H, et al. Genome-wide association study of coronary artery disease in the Japanese. *Eur J Hum Genet.* 2012;20(3):333-40.
50. Brooks PJ, Enoch MA, Goldman D, Li TK, Yokoyama A. The alcohol flushing response: an unrecognized risk factor for esophageal cancer from alcohol consumption. *PLoS Med.* 2009;6(3):e50.
51. Masaoka H, Ito H, Soga N, Hosono S, Oze I, Watanabe M, et al. Aldehyde dehydrogenase 2 (ALDH2) and alcohol dehydrogenase 1B (ADH1B) polymorphisms exacerbate bladder cancer risk associated with alcohol drinking: gene-environment interaction. *Carcinogenesis.* 2016;37(6):583-8.
52. Barnett SD, Buxton ILO. The role of S-nitrosoglutathione reductase (GSNOR) in human disease and therapy. *Crit Rev Biochem Mol Biol.* 2017;52(3):340-54.
53. Voulgaridou GP, Anastopoulos I, Franco R, Panayiotidis MI, Pappa A. DNA damage induced by endogenous aldehydes: current state of knowledge. *Mutat Res.* 2011;711(1-2):13-27.
54. Chaw YF, Crane LE, Lange P, Shapiro R. Isolation and identification of cross-links from formaldehyde-treated nucleic acids. *Biochemistry.* 1980;19(24):5525-31.
55. Swenberg JA, Lu K, Moeller BC, Gao L, Upton PB, Nakamura J, et al. Endogenous versus exogenous DNA adducts: their role in carcinogenesis, epidemiology, and risk assessment. *Toxicol Sci.* 2011;120 Suppl 1(Suppl 1):S130-45.
56. Cheng G, Shi Y, Sturla SJ, Jalas JR, McIntee EJ, Villalta PW, et al. Reactions of formaldehyde plus acetaldehyde with deoxyguanosine and DNA: formation of cyclic deoxyguanosine adducts and formaldehyde cross-links. *Chem Res Toxicol.* 2003;16(2):145-52.
57. Chandramouly G, Liao S, Rusanov T, Borisonnik N, Calbert ML, Kent T, et al. Polθ promotes the repair of 5'-DNA-protein crosslinks by microhomology-mediated end-joining. *Cell Rep.* 2021;34(10):108820.

58. Aparicio T, Baer R, Gottesman M, Gautier J. MRN, CtIP, and BRCA1 mediate repair of topoisomerase II-DNA adducts. *J Cell Biol.* 2016;212(4):399-408.
59. Deshpande RA, Lee JH, Arora S, Paull TT. Nbs1 Converts the Human Mre11/Rad50 Nuclease Complex into an Endo/Exonuclease Machine Specific for Protein-DNA Adducts. *Mol Cell.* 2016;64(3):593-606.
60. de Graaf B, Clore A, McCullough AK. Cellular pathways for DNA repair and damage tolerance of formaldehyde-induced DNA-protein crosslinks. *DNA Repair (Amst).* 2009;8(10):1207-14.
61. Nakano T, Katafuchi A, Matsubara M, Terato H, Tsuboi T, Masuda T, et al. Homologous recombination but not nucleotide excision repair plays a pivotal role in tolerance of DNA-protein cross-links in mammalian cells. *J Biol Chem.* 2009;284(40):27065-76.
62. Juarez E, Chambwe N, Tang W, Mitchell AD, Owen N, Kumari A, et al. An RNAi screen in human cell lines reveals conserved DNA damage repair pathways that mitigate formaldehyde sensitivity. *DNA Repair (Amst).* 2018;72:1-9.
63. Chapman JR, Taylor MR, Boulton SJ. Playing the end game: DNA double-strand break repair pathway choice. *Mol Cell.* 2012;47(4):497-510.
64. Scully R, Panday A, Elango R, Willis NA. DNA double-strand break repair-pathway choice in somatic mammalian cells. *Nat Rev Mol Cell Biol.* 2019;20(11):698-714.
65. Oh JM, Myung K. Crosstalk between different DNA repair pathways for DNA double strand break repairs. *Mutat Res Genet Toxicol Environ Mutagen.* 2022;873:503438.
66. Kass EM, Jasin M. Collaboration and competition between DNA double-strand break repair pathways. *FEBS Lett.* 2010;584(17):3703-8.
67. Drouet J, Delteil C, Lefrancois J, Concannon P, Salles B, Calsou P. DNA-dependent protein kinase and XRCC4-DNA ligase IV mobilization in the cell in response to DNA double strand breaks. *J Biol Chem.* 2005;280(8):7060-9.
68. Lavin MF, Kozlov S, Gatei M, Kijas AW. ATM-Dependent Phosphorylation of All Three Members of the MRN Complex: From Sensor to Adaptor. *Biomolecules.* 2015;5(4):2877-902.

69. Burma S, Chen BP, Murphy M, Kurimasa A, Chen DJ. ATM phosphorylates histone H2AX in response to DNA double-strand breaks. *J Biol Chem.* 2001;276(45):42462-7.
70. Irminger-Finger I, Ratajska M, Pilyugin M. New concepts on BARD1: Regulator of BRCA pathways and beyond. *Int J Biochem Cell Biol.* 2016;72:1-17.
71. Chen L, Nievera CJ, Lee AY, Wu X. Cell cycle-dependent complex formation of BRCA1.CtIP.MRN is important for DNA double-strand break repair. *J Biol Chem.* 2008;283(12):7713-20.
72. Schrempf A, Slysokova J, Loizou JI. Targeting the DNA Repair Enzyme Polymerase θ in Cancer Therapy. *Trends Cancer.* 2021;7(2):98-111.
73. Zhao W, Steinfeld JB, Liang F, Chen X, Maranon DG, Jian Ma C, et al. BRCA1-BARD1 promotes RAD51-mediated homologous DNA pairing. *Nature.* 2017;550(7676):360-5.
74. Zhang F, Ma J, Wu J, Ye L, Cai H, Xia B, et al. PALB2 links BRCA1 and BRCA2 in the DNA-damage response. *Curr Biol.* 2009;19(6):524-9.
75. Mazin AV, Mazina OM, Bugreev DV, Rossi MJ. Rad54, the motor of homologous recombination. *DNA Repair (Amst).* 2010;9(3):286-302.
76. Wang Z, Song Y, Li S, Kurian S, Xiang R, Chiba T, et al. DNA polymerase theta (POLQ) is important for repair of DNA double-strand breaks caused by fork collapse. *J Biol Chem.* 2019;294(11):3909-19.
77. Kent T, Chandramouly G, McDevitt SM, Ozdemir AY, Pomerantz RT. Mechanism of microhomology-mediated end-joining promoted by human DNA polymerase theta. *Nat Struct Mol Biol.* 2015;22(3):230-7.
78. Luedeman ME, Stroik S, Feng W, Luthman AJ, Gupta GP, Ramsden DA. Poly(ADP) ribose polymerase promotes DNA polymerase theta-mediated end joining by activation of end resection. *Nat Commun.* 2022;13(1):4547.
79. Seki M, Marini F, Wood RD. POLQ (Pol theta), a DNA polymerase and DNA-dependent ATPase in human cells. *Nucleic Acids Res.* 2003;31(21):6117-26.
80. Newman JA, Cooper CDO, Aitkenhead H, Gileadi O. Structure of the Helicase Domain of DNA Polymerase Theta Reveals a Possible Role in the Microhomology-Mediated End-Joining Pathway. *Structure.* 2015;23(12):2319-30.

81. Ceccaldi R, Liu JC, Amunugama R, Hajdu I, Primack B, Petalcorin MI, et al. Homologous-recombination-deficient tumours are dependent on Poltheta-mediated repair. *Nature*. 2015;518(7538):258-62.
82. Mateos-Gomez PA, Kent T, Deng SK, McDevitt S, Kashkina E, Hoang TM, et al. The helicase domain of Poltheta counteracts RPA to promote alt-NHEJ. *Nat Struct Mol Biol*. 2017;24(12):1116-23.
83. Truong LN, Li Y, Shi LZ, Hwang PY, He J, Wang H, et al. Microhomology-mediated End Joining and Homologous Recombination share the initial end resection step to repair DNA double-strand breaks in mammalian cells. *Proc Natl Acad Sci U S A*. 2013;110(19):7720-5. doi: 10.1073/pnas.1213431110. Epub 2013 Apr 22.
84. Goulet de Rugy T, Bashkurov M, Datti A, Betous R, Guitton-Sert L, Cazaux C, et al. Excess Poltheta functions in response to replicative stress in homologous recombination-proficient cancer cells. *Biol Open*. 2016;5(10):1485-92.
85. Pilie PG, Tang C, Mills GB, Yap TA. State-of-the-art strategies for targeting the DNA damage response in cancer. *Nat Rev Clin Oncol*. 2019;16(2):81-104.
86. Ngoi NYL, Gallo D, Torrado C, Nardo M, Durocher D, Yap TA. Synthetic lethal strategies for the development of cancer therapeutics. *Nat Rev Clin Oncol*. 2025;22(1):46-64.
87. Bryant HE, Schultz N, Thomas HD, Parker KM, Flower D, Lopez E, et al. Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. *Nature*. 2005;434(7035):913-7.
88. Li Q, Qian W, Zhang Y, Hu L, Chen S, Xia Y. A new wave of innovations within the DNA damage response. *Signal Transduct Target Ther*. 2023;8(1):338.
89. Kawamura K, Bahar R, Seimiya M, Chiyo M, Wada A, Okada S, et al. DNA polymerase theta is preferentially expressed in lymphoid tissues and upregulated in human cancers. *Int J Cancer*. 2004;109(1):9-16.
90. Lemee F, Bergoglio V, Fernandez-Vidal A, Machado-Silva A, Pillaire MJ, Bieth A, et al. DNA polymerase theta up-regulation is associated with poor survival in breast cancer, perturbs DNA replication, and promotes genetic instability. *Proc Natl Acad Sci U S A*. 2010;107(30):13390-5.

91. Feng W, Simpson DA, Carvajal-Garcia J, Price BA, Kumar RJ, Mose LE, et al. Genetic determinants of cellular addiction to DNA polymerase theta. *Nat Commun.* 2019;10(1):4286.
92. Zatreanu D, Robinson HMR, Alkhatib O, Boursier M, Finch H, Geo L, et al. Polθ inhibitors elicit BRCA-gene synthetic lethality and target PARP inhibitor resistance. *Nat Commun.* 2021;12(1):3636.
93. Zhou J, Gelot C, Pantelidou C, Li A, Yücel H, Davis RE, et al. A first-in-class Polymerase Theta Inhibitor selectively targets Homologous-Recombination-Deficient Tumors. *Nature cancer.* 2021;2(6):598-610.
94. Formaldehyde Causes More Cancer Than Any Other Toxic Air Pollutant. Little Is Being Done to Curb the Risk. [Internet]. ProPublica. 2024. Available from: <https://www.propublica.org/article/formaldehyde-epa-trump-public-health-danger>
95. Hauptmann M, Stewart PA, Lubin JH, Beane Freeman LE, Hornung RW, Herrick RF, et al. Mortality from lymphohematopoietic malignancies and brain cancer among embalmers exposed to formaldehyde. *J Natl Cancer Inst.* 2009;101(24):1696-708.
96. Zhang L, Tang X, Rothman N, Vermeulen R, Ji Z, Shen M, et al. Occupational exposure to formaldehyde, hematotoxicity, and leukemia-specific chromosome changes in cultured myeloid progenitor cells. *Cancer Epidemiol Biomarkers Prev.* 2010;19(1):80-8.
97. Protano C, Buomprisco G, Cammalleri V, Pocino RN, Marotta D, Simonazzi S, et al. The Carcinogenic Effects of Formaldehyde Occupational Exposure: A Systematic Review. *Cancers (Basel).* 2021;14(1).
98. Stayner LT, Elliott L, Blade L, Keenlyside R, Halperin W. A retrospective cohort mortality study of workers exposed to formaldehyde in the garment industry. *Am J Ind Med.* 1988;13(6):667-81.
99. Beane Freeman LE, Blair A, Lubin JH, Stewart PA, Hayes RB, Hoover RN, et al. Mortality from lymphohematopoietic malignancies among workers in formaldehyde industries: the National Cancer Institute Cohort. *J Natl Cancer Inst.* 2009;101(10):751-61.
100. Blackwell M, Kang H, Thomas A, Infante P. Formaldehyde: evidence of carcinogenicity. *Am Ind Hyg Assoc J.* 1981;42(7):A34, A6, A8, passim.
101. Kerns WD, Pavkov KL, Donofrio DJ, Gralla EJ, Swenberg JA. Carcinogenicity of formaldehyde in rats and mice after long-term inhalation exposure. *Cancer Res.* 1983;43(9):4382-92.

102. Nishikawa A, Nagano K, Kojima H, Ogawa K. A comprehensive review of mechanistic insights into formaldehyde-induced nasal cavity carcinogenicity. *Regul Toxicol Pharmacol.* 2021;123:104937.
103. Swenberg JA, Moeller BC, Lu K, Rager JE, Fry RC, Starr TB. Formaldehyde carcinogenicity research: 30 years and counting for mode of action, epidemiology, and cancer risk assessment. *Toxicol Pathol.* 2013;41(2):181-9.
104. Tobe M, Naito K, Kurokawa Y. Chronic toxicity study on formaldehyde administered orally to rats. *Toxicology.* 1989;56(1):79-86. doi: 10.1016/0300-483x(89)90213-8.
105. Swenberg JA, Kerns WD, Mitchell RI, Gralla EJ, Pavkov KL. Induction of squamous cell carcinomas of the rat nasal cavity by inhalation exposure to formaldehyde vapor. *Cancer Res.* 1980;40(9):3398-402.
106. Zhang L, Steinmaus C, Eastmond DA, Xin XK, Smith MT. Formaldehyde exposure and leukemia: a new meta-analysis and potential mechanisms. *Mutat Res.* 2009;681(2-3):150-68.
107. Pinkerton LE, Hein MJ, Stayner LT. Mortality among a cohort of garment workers exposed to formaldehyde: an update. *Occup Environ Med.* 2004;61(3):193-200.
108. Kitaeva LV, Kitaev EM, Pimenova MN. [The cytopathic and cytogenetic sequelae of chronic inhalational exposure to formaldehyde on female germ cells and bone marrow cells in rats]. *Tsitologiya.* 1990;32(12):1212-6.
109. Orsiere T, Sari-Minodier I, Iarmarcovai G, Botta A. Genotoxic risk assessment of pathology and anatomy laboratory workers exposed to formaldehyde by use of personal air sampling and analysis of DNA damage in peripheral lymphocytes. *Mutat Res.* 2006;605(1-2):30-41.
110. Ye X, Yan W, Xie H, Zhao M, Ying C. Cytogenetic analysis of nasal mucosa cells and lymphocytes from high-level long-term formaldehyde exposed workers and low-level short-term exposed waiters. *Mutat Res.* 2005;588(1):22-7.
111. Yu LQ, Jiang SF, Leng SG, He FS, Zheng YX. [Early genetic effects on workers occupationally exposed to formaldehyde]. *Zhonghua Yu Fang Yi Xue Za Zhi.* 2005;39(6):392-5.
112. Suruda A, Schulte P, Boeniger M, Hayes RB, Livingston GK, Steenland K, et al. Cytogenetic effects of formaldehyde exposure in students of mortuary science. *Cancer Epidemiol Biomarkers Prev.* 1993;2(5):453-60.

113. Zhang Y, Liu X, McHale C, Li R, Zhang L, Wu Y, et al. Bone marrow injury induced via oxidative stress in mice by inhalation exposure to formaldehyde. *PLoS One*. 2013;8(9):e74974. doi: 10.1371/journal.pone.0074974. eCollection 2013.
114. Zhang L, Freeman LE, Nakamura J, Hecht SS, Vandenberg JJ, Smith MT, et al. Formaldehyde and leukemia: epidemiology, potential mechanisms, and implications for risk assessment. *Environ Mol Mutagen*. 2010;51(3):181-91. doi: 10.1002/em.20534.
115. Ye X, Ji Z, Wei C, McHale CM, Ding S, Thomas R, et al. Inhaled formaldehyde induces DNA-protein crosslinks and oxidative stress in bone marrow and other distant organs of exposed mice. *Environ Mol Mutagen*. 2013;54(9):705-18.
116. Wei C, Wen H, Yuan L, McHale CM, Li H, Wang K, et al. Formaldehyde induces toxicity in mouse bone marrow and hematopoietic stem/progenitor cells and enhances benzene-induced adverse effects. *Arch Toxicol*. 2017;91(2):921-33. doi: 10.1007/s00204-016-1760-5. Epub 2016 Jun 23.
117. Barker S, Weinfeld M, Murray D. DNA-protein crosslinks: their induction, repair, and biological consequences. *Mutat Res*. 2005;589(2):111-35.
118. Nakamura J, Nakamura M. DNA-protein crosslink formation by endogenous aldehydes and AP sites. *DNA Repair (Amst)*. 2020;88:102806.
119. Yousefzadeh MJ, Wyatt DW, Takata K, Mu Y, Hensley SC, Tomida J, et al. Mechanism of suppression of chromosomal instability by DNA polymerase POLQ. *PLoS Genet*. 2014;10(10):e1004654.
120. Higgins GS, Prevo R, Lee YF, Helleday T, Muschel RJ, Taylor S, et al. A small interfering RNA screen of genes involved in DNA repair identifies tumor-specific radiosensitization by POLQ knockdown. *Cancer Res*. 2010;70(7):2984-93.
121. Wyatt DW, Feng W, Conlin MP, Yousefzadeh MJ, Roberts SA, Mieczkowski P, et al. Essential Roles for Polymerase theta-Mediated End Joining in the Repair of Chromosome Breaks. *Mol Cell*. 2016;63(4):662-73.
122. Belan O, Sebald M, Adamowicz M, Anand R, Vancevska A, Neves J, et al. POLQ seals post-replicative ssDNA gaps to maintain genome stability in BRCA-deficient cancer cells. *Mol Cell*. 2022;82(24):4664-80 e9.
123. Mann A, Ramirez-Otero MA, De Antoni A, Hanthi YW, Sannino V, Baldi G, et al. POLtheta prevents MRE11-NBS1-CtIP-dependent fork breakage in the absence of BRCA2/RAD51 by filling lagging-strand gaps. *Mol Cell*. 2022;82(22):4218-31 e8.

124. Vekariya U, Toma M, Nieborowska-Skorska M, Le BV, Caron MC, Kukuyan AM, et al. DNA polymerase θ protects leukemia cells from metabolically induced DNA damage. *Blood*. 2023;141(19):2372-89.
125. Skorski T. Oncogenic tyrosine kinases and the DNA-damage response. *Nat Rev Cancer*. 2002;2(5):351-60.
126. Nieborowska-Skorska M, Sullivan K, Dasgupta Y, Podszywalow-Bartnicka P, Hoser G, Maifrede S, et al. Gene expression and mutation-guided synthetic lethality eradicates proliferating and quiescent leukemia cells. *J Clin Invest*. 2017;127(6):2392-406.
127. Sullivan-Reed K, Bolton-Gillespie E, Dasgupta Y, Langer S, Siciliano M, Nieborowska-Skorska M, et al. Simultaneous Targeting of PARP1 and RAD52 Triggers Dual Synthetic Lethality in BRCA-Deficient Tumor Cells. *Cell Rep*. 2018;23(11):3127-36.
128. Young L, Sung J, Stacey G, Masters JR. Detection of Mycoplasma in cell cultures. *Nat Protoc*. 2010;5(5):929-34.
129. Vekariya U, Minakhin L, Chandramouly G, Tyagi M, Kent T, Sullivan-Reed K, et al. PARG is essential for Poltheta-mediated DNA end-joining by removing repressive poly-ADP-ribose marks. *Nat Commun*. 2024;15(1):5822.
130. Hu Q, Klages-Mundt N, Wang R, Lynn E, Kuma Saha L, Zhang H, et al. The ARK Assay Is a Sensitive and Versatile Method for the Global Detection of DNA-Protein Crosslinks. *Cell Rep*. 2020;30(4):1235-45.e4.
131. Kosmider B, Messier EM, Chu HW, Mason RJ. Human alveolar epithelial cell injury induced by cigarette smoke. *PloS one*. 2011;6(12):e26059.
132. Kosmider B, Lin CR, Vlasenko L, Marchetti N, Bolla S, Criner GJ, et al. Impaired non-homologous end joining in human primary alveolar type II cells in emphysema. *Sci Rep*. 2019;9(1):920.
133. Lanz J, Biniiaz-Harris N, Kuvaldina M, Jain S, Lewis K, Fallon BA. Disulfiram: Mechanisms, Applications, and Challenges. *Antibiotics (Basel)*. 2023;12(3).
134. Green LS, Chun LE, Patton AK, Sun X, Rosenthal GJ, Richards JP. Mechanism of inhibition for N6022, a first-in-class drug targeting S-nitrosoglutathione reductase. *Biochemistry*. 2012;51(10):2157-68.

135. Mateos-Gomez PA, Gong F, Nair N, Miller KM, Lazzerini-Denchi E, Sfeir A. Mammalian polymerase theta promotes alternative NHEJ and suppresses recombination. *Nature*. 2015;518(7538):254-7.
136. Sullivan-Reed K, Toma MM, Drzewiecka M, Nieborowska-Skorska M, Nejati R, Karami A, et al. Simultaneous Targeting of DNA Polymerase Theta and PARP1 or RAD52 Triggers Dual Synthetic Lethality in Homologous Recombination-Deficient Leukemia Cells. *Mol Cancer Res*. 2023;21(10):1017-22.
137. Fircanis S, Merriam P, Khan N, Castillo JJ. The relation between cigarette smoking and risk of acute myeloid leukemia: an updated meta-analysis of epidemiological studies. *Am J Hematol*. 2014;89(8):E125-32.
138. (EFSA) EFSA. Endogenous formaldehyde turnover in humans compared with exogenous contribution from food sources. *EFSA Journal* 2014;12(2):3550.
139. Karsa M, Xiao L, Ronca E, Bongers A, Spurling D, Karsa A, et al. FDA-approved disulfiram as a novel treatment for aggressive leukemia. *J Mol Med (Berl)*. 2024;102(4):507-19.
140. Yang W, Xie J, Hou R, Chen X, Xu Z, Tan Y, et al. Disulfiram/cytarabine eradicates a subset of acute myeloid leukemia stem cells with high aldehyde dehydrogenase expression. *Leuk Res*. 2020;92:106351.
141. Edenberg HJ. The genetics of alcohol metabolism: role of alcohol dehydrogenase and aldehyde dehydrogenase variants. *Alcohol Res Health*. 2007;30(1):5-13.
142. Jackson B, Brocker C, Thompson DC, Black W, Vasiliou K, Nebert DW, et al. Update on the aldehyde dehydrogenase gene (ALDH) superfamily. *Hum Genomics*. 2011;5(4):283-303.
143. Lavudi K, Banerjee A, Li N, Yang Y, Cai S, Bai X, et al. ALDH1A1 promotes PARP inhibitor resistance by enhancing retinoic acid receptor-mediated DNA polymerase theta expression. *NPJ Precis Oncol*. 2023;7(1):66.
144. Chelghoum Y, Danaila C, Belhabri A, Charrin C, Le QH, Michallet M, et al. Influence of cigarette smoking on the presentation and course of acute myeloid leukemia. *Ann Oncol*. 2002;13(10):1621-7.

145. Smith CJ, Hansch C. The relative toxicity of compounds in mainstream cigarette smoke condensate. *Food Chem Toxicol.* 2000;38(7):637-46.
146. Seeman JI, Dixon M, Hausmann HJ. Acetaldehyde in mainstream tobacco smoke: formation and occurrence in smoke and bioavailability in the smoker. *Chem Res Toxicol.* 2002;15(11):1331-50.
147. Shinmura K, Kato H, Kawanishi Y, Yoshimura K, Tsuchiya K, Takahara Y, et al. POLQ Overexpression Is Associated with an Increased Somatic Mutation Load and PLK4 Overexpression in Lung Adenocarcinoma. *Cancers (Basel).* 2019;11(5).
148. Roswall N, Weiderpass E. Alcohol as a risk factor for cancer: existing evidence in a global perspective. *J Prev Med Public Health.* 2015;48(1):1-9.
149. Wilson DF, Matschinsky FM. Ethanol metabolism: The good, the bad, and the ugly. *Med Hypotheses.* 2020;140:109638.
150. Morimoto K, Takeshita T. Low Km aldehyde dehydrogenase (ALDH2) polymorphism, alcohol-drinking behavior, and chromosome alterations in peripheral lymphocytes. *Environ Health Perspect.* 1996;104 Suppl 3(Suppl 3):563-7.
151. Yokoyama A, Omori T. Genetic polymorphisms of alcohol and aldehyde dehydrogenases and risk for esophageal and head and neck cancers. *Jpn J Clin Oncol.* 2003;33(3):111-21.
152. Dellarco VL. A mutagenicity assessment of acetaldehyde. *Mutat Res.* 1988;195(1):1-20.
153. Brooks PJ, Theruvathu JA. DNA adducts from acetaldehyde: implications for alcohol-related carcinogenesis. *Alcohol.* 2005;35(3):187-93.
154. Wang M, McIntee EJ, Cheng G, Shi Y, Villalta PW, Hecht SS. Identification of DNA adducts of acetaldehyde. *Chem Res Toxicol.* 2000;13(11):1149-57.
155. Chen CH, Ferreira JC, Gross ER, Mochly-Rosen D. Targeting aldehyde dehydrogenase 2: new therapeutic opportunities. *Physiol Rev.* 2014;94(1):1-34.
156. Li K, Guo W, Li Z, Wang Y, Sun B, Xu D, et al. ALDH2 Repression Promotes Lung Tumor Progression via Accumulated Acetaldehyde and DNA Damage. *Neoplasia.* 2019;21(6):602-14.

157. Abraham J, Balbo S, Crabb D, Brooks PJ. Alcohol metabolism in human cells causes DNA damage and activates the Fanconi anemia-breast cancer susceptibility (FA-BRCA) DNA damage response network. *Alcohol Clin Exp Res.* 2011;35(12):2113-20.
158. Chen X, Legrand AJ, Cunniffe S, Hume S, Poletto M, Vaz B, et al. Interplay between base excision repair protein XRCC1 and ALDH2 predicts overall survival in lung and liver cancer patients. *Cell Oncol (Dordr).* 2018;41(5):527-39.
159. Caldecott KW. XRCC1 and DNA strand break repair. *DNA Repair (Amst).* 2003;2(9):955-69.
160. Markkanen E, Fischer R, Ledentcova M, Kessler BM, Dianov GL. Cells deficient in base-excision repair reveal cancer hallmarks originating from adjustments to genetic instability. *Nucleic Acids Res.* 2015;43(7):3667-79.
161. Kais Z, Rondinelli B, Holmes A, O'Leary C, Kozono D, D'Andrea AD, et al. FANCD2 Maintains Fork Stability in BRCA1/2-Deficient Tumors and Promotes Alternative End-Joining DNA Repair. *Cell Rep.* 2016;15(11):2488-99.
162. Gao Y, Guitton-Sert L, Dessapt J, Coulombe Y, Rodrigue A, Milano L, et al. A CRISPR-Cas9 screen identifies EXO1 as a formaldehyde resistance gene. *Nat Commun.* 2023;14(1):381.
163. Gregersen LH, Svejstrup JQ. The Cellular Response to Transcription-Blocking DNA Damage. *Trends Biochem Sci.* 2018;43(5):327-41.
164. Mulderrig L, Garaycoechea JI, Tuong ZK, Millington CL, Dingler FA, Ferdinand JR, et al. Aldehyde-driven transcriptional stress triggers an anorexic DNA damage response. *Nature.* 2021;600(7887):158-63.
165. Rieckher M, Gallrein C, Alquezar-Artieda N, Bourached-Silva N, Vaddavalli PL, Mares D, et al. Distinct DNA repair mechanisms prevent formaldehyde toxicity during development, reproduction and aging. *Nucleic Acids Res.* 2024;52(14):8271-85.

APPENDIX A

SUPPLEMENTARY DATA

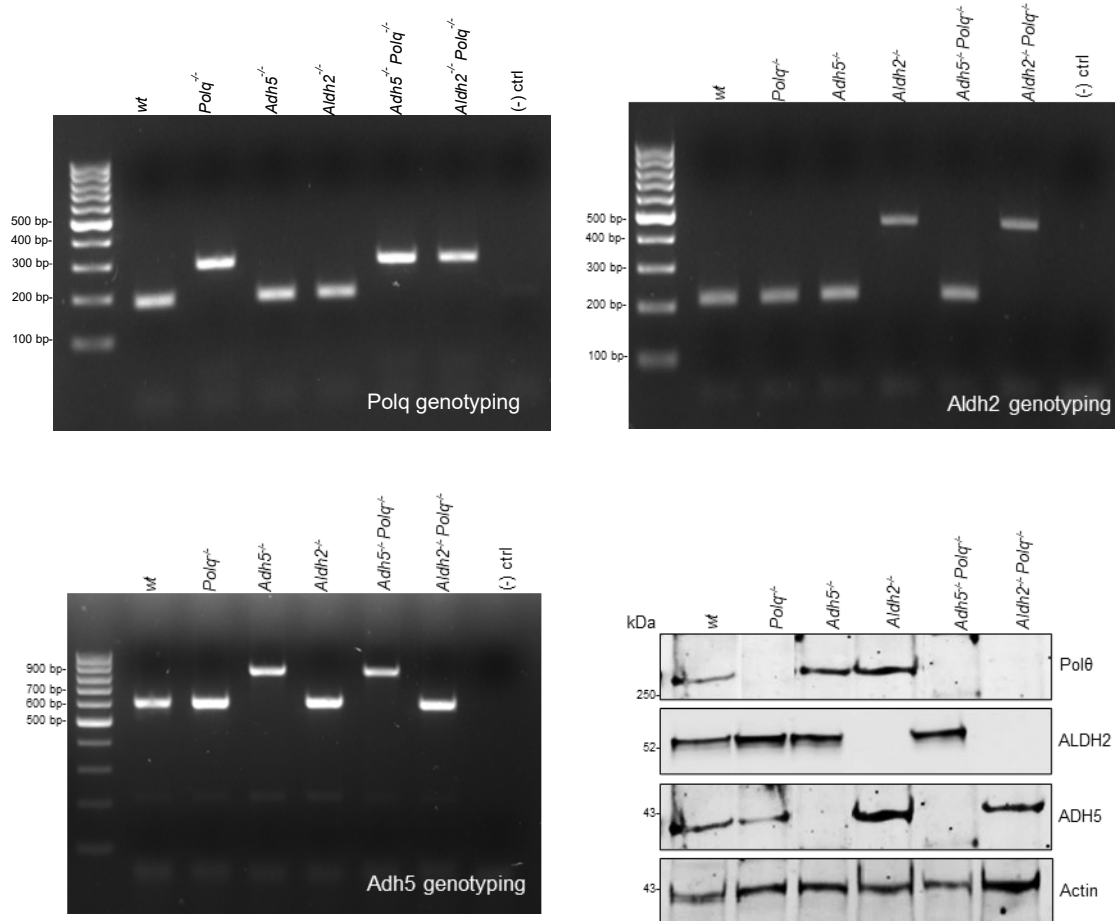


Figure 37. Validation of genetic knockout mice. (A) PCR using Polq primers. Wild-type band is located at 190 bp while mutant band is located at 300 bp. (B) PCR using Aldh2 primers. Wild-type band is at 212 bp while the mutant band is at 414 bp. (C) PCR using Adh5 primers. Wild-type band is at 600 bp while the mutant band is at 900 bp. (D) Western blot analysis of Polq (290 kDa), ADH5 (40 kDa), ALDH2 (56 kDa), and Actin (42kDa) detected in mouse HSPCs.

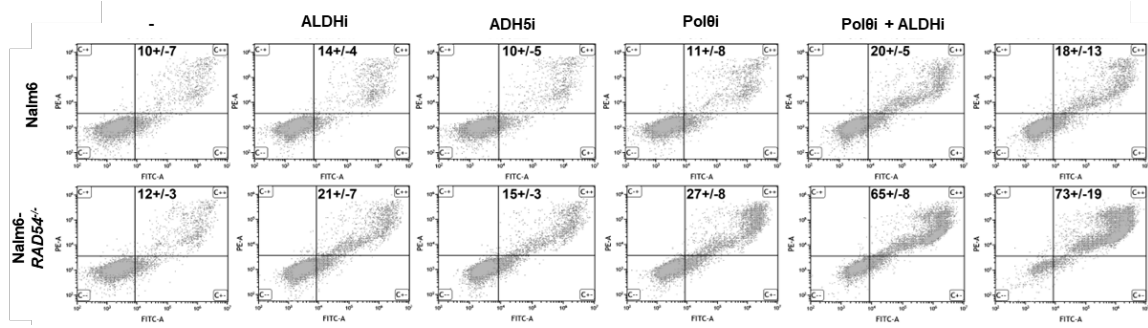


Figure 38. Combination of Polθi + ALDHi and Polθi + ADH5i induced apoptosis in HR-deficient Nalm6-RAD54^{-/-} cells. Visual representation of Nalm6 and Nalm6-RAD54^{-/-} cells treated with 12.5 μM Polθi, 0.125 μM ALDHi, 20 μM ADH5i, and indicated combinations for 48 hours followed by Annexin V/propidium iodide staining. Results represent mean % ± SEM of the double-stained cells.