

UCLA

UCLA Electronic Theses and Dissertations

Title

From Germline to Somatic: ASXL1's Multifaceted Role in Development and Disease

Permalink

<https://escholarship.org/uc/item/2wk1s48s>

Author

Lin, Isabella

Publication Date

2024

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

From Germline to Somatic:

ASXL1's Multifaceted Role in Development and Disease

A dissertation submitted in satisfaction of the
requirements for the degree Doctor of Philosophy
in Human Genetics

by

Isabella Lin

2024

© Copyright by

Isabella Lin

2024

ABSTRACT OF THE DISSERTATION

From Germline to Somatic:
ASXL1's Multifaceted Role in Development and Disease

by

Isabella Lin

Doctor of Philosophy in Human Genetics

University of California, Los Angeles, 2024

Professor Valerie A. Arboleda, Chair

Genetics underpins all of biology, and is the most rapidly evolving field in biomedical research. The *ASXL* gene family are critical players in rare developmental disorders and various cancers, bridging seemingly disparate areas of study. This dissertation provides a comprehensive examination of the role of *ASXL1* mutations in the intersection of rare genetic disorders and cancer, highlighting the multifaceted impact across developmental biology, genetics, and metabolism.

Chapter 1 explores the broad regulatory roles of the *ASXL* gene family in developmental and oncogenic processes, detailing the molecular architecture, clinical and epigenetic consequences of *ASXL* mutations, and identifies potential avenues for research and therapeutic development. **Chapter 2** investigates how *ASXL1* mutations disrupt epigenetic regulation in Bohring-Opitz syndrome (BOS) patient-derived tissues, impacting gene expression through dysregulation of Wnt signaling pathways. This cross-tissue effect offers potential therapeutic targets.

Chapter 3 also utilizes a multi-omics approach to explore the effect of *ASXL1* mutations across development in induced pluripotent stem cells , neural progenitor cells , and neural crest cells, identifying defects in neurodevelopment and cell migration. **Chapter 4** delves into metabolic dysregulations in BOS, highlighting mitochondrial dysfunction and metabolic inflexibility that could underlie clinical manifestations of the syndrome. **Chapter 5** presents a unique familial case with *ASXL1* missense mutation, expanding the phenotypic spectrum associated with *ASXL1* mutations and suggesting distinct metabolic phenotypes.

In **Chapter 6**, a direct comparison of *ASXL1* mutations in acute myeloid leukemia and BOS reveals shared molecular pathways, advocating a unified gene-centric approach to studying these disorders. **Chapter 7** discusses the implications of *BLM* gene mutations in clonal hematopoiesis of indeterminate potential genes, linking germline disorders and cancer predisposition.

By integrating multi-omics analyses, patient-derived cellular models, and comprehensive metabolic characterizations, this work establishes crucial connections between rare genetic disorders, metabolic defects, and cancer, underscoring the potential for cross-disease insights to inform targeted therapies. This dissertation contributes to our understanding of specific disorders like BOS and reveals shared molecular dysregulations between rare germline disorders and cancer. These findings pave the way for developing targeted therapeutic interventions that could benefit a broad spectrum of conditions.

The dissertation of Isabella Lin is approved.

Dinesh S. Rao

Harold J. Pimentel

Julian A. Martinez

Valerie A. Arboleda, Chair

TABLE OF CONTENTS

<u>Abstract</u>	ii
<u>Committee Page</u>	iv
<u>List of Tables and Figures</u>	viii
<u>List of Acronyms and Abbreviations</u>	xiii
<u>Acknowledgments</u>	xxviii
<u>Vita</u>	xxx
<u>Chapter 1. The ASXL Gene Family: Regulating Development and Growth Through Epigenetic Mechanisms</u>	1
<u>Introduction</u>	2
<u>Gene structure of the ASXL gene family</u>	5
<u>Regulatory mechanisms governing ASXL genes</u>	12
<u>Mechanisms of truncating mutation-driven dysfunction</u>	13
<u>Pathogenic ASXL germline mutations: inheritance and mutation types</u>	15
<u>Clinical consequences of somatic ASXL mutations in malignancy</u>	17
<u>Challenges in ASXL research and therapeutic development</u>	19
<u>Key gaps in research</u>	22
<u>Project goals and methodologies</u>	23
<u>References</u>	41
<u>Chapter 2. Multi-omics of Bohring-Opitz syndrome: Identifying Wnt Signaling Dysregulation</u>	57
<u>Introduction</u>	59
<u>Methods</u>	63
<u>Results</u>	71

<u>Discussion</u>	74
<u>Conclusion</u>	91
<u>References</u>	116
<u>Chapter 3. Dissecting the Impact of <i>ASXL1</i> Mutations on Neurodevelopmental Processes in Bohring-Opitz syndrome</u>	126
<u>Introduction</u>	128
<u>Methods</u>	131
<u>Results</u>	134
<u>Discussion</u>	139
<u>References</u>	151
<u>Chapter 4. <i>ASXL1</i> Mutations Drive Metabolic Dysregulation and Mitochondrial Dysfunction in Bohring-Opitz syndrome</u>	154
<u>Introduction</u>	156
<u>Methods</u>	159
<u>Results</u>	165
<u>Discussion</u>	174
<u>Conclusion</u>	179
<u>References</u>	198
<u>Chapter 5. Expanding the <i>ASXL1</i> Mutation Spectrum: Transcriptomic and Metabolic Dysregulation Underlying Overgrowth Phenotype in Inherited <i>ASXL1</i> Missense Mutations</u>	203
<u>Introduction</u>	205
<u>Methods</u>	207
<u>Results</u>	208

<u>Discussion</u>	216
<u>Conclusion</u>	220
<u>References</u>	242
<u>Chapter 6. Comparative Analysis of ASXL1 Mutations in Bohring-Opitz syndrome and Acute Myeloid Leukemia</u>	245
<u>Introduction</u>	247
<u>Methods</u>	250
<u>Results</u>	252
<u>Discussion</u>	260
<u>Conclusion</u>	267
<u>References</u>	299
<u>Chapter 7. Bloom Syndrome and Carriers: Increased Frequency of Somatic Mutations in Clonal Hematopoiesis of Indeterminate Potential</u>	304
<u>Introduction</u>	305
<u>Methods</u>	309
<u>Results</u>	311
<u>Discussion</u>	315
<u>References</u>	332
<u>Chapter 8. Conclusion: Bridging Genes and Diseases: From Insights to Impact</u>	339
<u>Concluding remarks</u>	340
<u>Challenges and limitations</u>	341
<u>Future research directions</u>	341

LIST OF TABLES AND FIGURES

CHAPTER 1

<u>Figure 1-1</u>	26
<u>Figure 1-2</u>	27
<u>Figure 1-3</u>	28
<u>Table 1-1</u>	29
<u>Table 1-2</u>	34
<u>Table 1-3</u>	40

CHAPTER 2

<u>Figure 2-1</u>	93
<u>Figure 2-2</u>	94
<u>Figure 2-3</u>	96
<u>Figure 2-4</u>	98
<u>Figure 2-5</u>	100
<u>Figure 2-6</u>	102
<u>Figure 2-7</u>	104
<u>Table 2-1</u>	105
<u>Table 2-2</u>	106
<u>Table 2-3</u>	108
<u>Table 2-4</u>	109
<u>Table 2-5</u>	110

<u>Table 2-6</u>	110
<u>Table 2-7</u>	112
<u>Table 2-8</u>	113
<u>Table 2-9</u>	114
<u>Table 2-10</u>	115

CHAPTER 3

<u>Figure 3-1</u>	142
<u>Figure 3-2</u>	143
<u>Figure 3-3</u>	145
<u>Figure 3-4</u>	147
<u>Table 3-1</u>	149
<u>Table 3-2</u>	150

CHAPTER 4

<u>Figure 4-1</u>	1811
<u>Figure 4-2</u>	182
<u>Figure 4-3</u>	183
<u>Figure 4-4</u>	185
<u>Figure 4-5</u>	187
<u>Figure 4-6</u>	189
<u>Figure 4-7</u>	191
<u>Figure 4-8</u>	192

<u>Figure 4-9</u>	193
<u>Table 4-1</u>	194
<u>Table 4-2</u>	195
<u>Table 4-3</u>	196

CHAPTER 5

<u>Figure 5-1</u>	221
<u>Figure 5-2</u>	223
<u>Figure 5-3</u>	224
<u>Figure 5-4</u>	226
<u>Figure 5-5</u>	228
<u>Figure 5-6</u>	230
<u>Figure 5-7</u>	231
<u>Figure 5-8</u>	232
<u>Table 5-1</u>	234
<u>Table 5-2</u>	235
<u>Table 5-3</u>	236
<u>Table 5-4</u>	237
<u>Table 5-5</u>	238
<u>Table 5-6</u>	239
<u>Table 5-7</u>	240
<u>Table 5-8</u>	241

CHAPTER 6

<u>Figure 6-1</u>	268
<u>Figure 6-2</u>	270
<u>Figure 6-3</u>	272
<u>Figure 6-4</u>	274
<u>Figure 6-5</u>	275
<u>Figure 6-6</u>	277
<u>Figure 6-7</u>	278
<u>Figure 6-8</u>	279
<u>Figure 6-9</u>	281
<u>Figure 6-10</u>	283
<u>Figure 6-11</u>	284
<u>Figure 6-12</u>	285
<u>Table 6-1</u>	286
<u>Table 6-2</u>	287
<u>Table 6-3</u>	289
<u>Table 6-4</u>	290
<u>Table 6-5</u>	291
<u>Table 6-6</u>	292
<u>Table 6-7</u>	294
<u>Table 6-8</u>	295

<u>Table 6-9</u>	296
<u>Table 6-10</u>	297
<u>Table 6-11</u>	298
 <u>CHAPTER 7</u>	
<u>Figure 7-1</u>	320
<u>Figure 7-2</u>	321
<u>Figure 7-3</u>	323
<u>Figure 7-4</u>	325
<u>Table 7-1</u>	327
<u>Table 7-2</u>	328
<u>Table 7-3</u>	330
<u>Table 7-4</u>	331

LIST OF ACRONYMS AND ABBREVIATIONS

Chapter 1:

Abbreviation	Full Description
ASXL	Additional Sex Combs Like
Asx	Additional Sex Combs
PcG	Polycomb Group
PRC	Polycomb Repressive Complex
PR-DUB	Polycomb Repressive Deubiquitinase
H2AK119	Histone H2A lysine 119 ubiquitination
H3K27	Histone H3 lysine 27 methylation
CBX	Chromobox
PCGF	Polycomb Group Ring Finger
RING1	Ring Finger Protein 1
BCL6	B-Cell Lymphoma 6
BCOR	BCL6 Corepressor
HPH	Polyhomeotic Homolog
KDM2B	Lysine Demethylase 2B
EZH2	Enhancer of Zeste Homolog 2
EZH1	Enhancer of Zeste Homolog 1
EED	Embryonic Ectoderm Development
SUZ12	Suppressor of Zeste 12
BRCA1	Breast Cancer 1
BAP1	BRCA1 Associated Protein 1
FOXK1/2	Forkhead Box K1/2
HCFC1	Host Cell Factor C1

OGT	O-GlcNAc Transferase
MBD5/6	Methyl-CpG Binding Domain Protein 5/6
KDM1B	Lysine Demethylase 1B
Hox	Homeobox
TrxG	Trithorax Group
ETP	Enhancer of Trithorax and Polycomb
IDR	Intrinsically Disordered Regions
ASXH	ASXL Homolog
PHD	Plant Homeodomain
ASXN	ASX N-terminal
KDM1A	Lysine Demethylase 1A
NCOA1	Nuclear Receptor Coactivator 1
HARE-HTH	HBI, ASXL1, restriction endonuclease helix-turn-helix
DEUBAD	Deubiquitinase Adaptor
H3K4me3	Histone H3 lysine 4 trimethylation
PPAR γ	Peroxisome Proliferator-Activated Receptor Gamma
SHPNS	Shashi-Pena Syndrome
CHD	Congenital Heart Disease
BMSC	Bone Marrow Stromal Cell
BMI-1	BMI1 Proto-Oncogene, aka PCGF4
nano LC-MS/MS	Nano Liquid Chromatography Tandem Mass Spectrometry
CDC20	Cell Division Cycle 20
APC/C	Anaphase Promoting Complex/Cyclosome
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats

USP7	Ubiquitin Specific Peptidase 7
MCR	Mutation Cluster Region
NMD	Nonsense-Mediated mRNA Decay
LOF	Loss Of Function
ExAC	The Exome Aggregation Consortium Database
BOS	Bohring-Opitz syndrome
gnomAD	Genome Aggregation Database
BRPS	Bainbridge Ropers Syndrome
KLHL7	Kelch Like Family Member 7
MDS	Myelodysplastic Syndrome
AML	Acute Myeloid Leukemia
CMML	Chronic Myelomonocytic Leukemia
SCLC	Small Cell Lung Cancer
BRD4	Bromodomain Containing 4
PAX9	Paired Box 9
iPSC	Induced Pluripotent Stem Cells
NPC	Neural Progenitor Cells
MPC	Mitochondrial Pyruvate Carrier
CHIP	Clonal Haematopoiesis of Indeterminate Potential
BLM	Bloom Syndrome

Chapter 2:

Abbreviation	Full Description
ASXL1	Additional Sex Combs Like 1
BOS	Bohring-Opitz syndrome
ATAC-seq	Assay for Transposase-Accessible Chromatin using sequencing
DNAm	DNA methylation
ASXL2	Additional Sex Combs Like 2
ASXL3	Additional Sex Combs Like 3
ASXN	ASX N-terminal
HARE-HTH	HBI, ASXL1, restriction endonuclease helix-turn-helix
ASXH	ASXL Homolog
DEUBAD	Deubiquitinase Adaptor
BAP1	BRCA1 Associated Protein 1
PHD	Plant Homeodomain
H3K4me3	Histone H3 lysine 4 trimethylation
PcG	Polycomb Group
PRC	Polycomb Repressive Complex
PR-DUB	Polycomb Repressive Deubiquitinase
H2AK119ub	Histone H2A lysine 119 ubiquitination
H3K27me	Histone H3 lysine 27 monomethylation
H3K27me2	Histone H3 lysine 27 dimethylation
H3K27me3	Histone H3 lysine 27 trimethylation
MDS	Myelodysplastic Syndrome
AML	Acute Myeloid Leukemia
PCP	Planar Cell Polarity
FZD	Frizzled Class Receptor (gene family)

LRP5	LDL Receptor Related Protein 5
LRP6	LDL Receptor Related Protein 6
GSK3B	Glycogen Synthase Kinase 3 Beta
AXIN1	Axis Inhibition Protein 1
AXIN2	Axis Inhibition Protein 2
TCF/LEF	T Cell Factor/Lymphoid Enhancer Factor (gene family)
VANGL2	VANGL Planar Cell Polarity Protein 2
DVL	Disheveled Segment Polarity Protein
DNA	Deoxyribonucleic Acid
RNA	Ribonucleic Acid
PBMC	Peripheral Blood Mononuclear Cells
EDTA	Ethylenediaminetetraacetic Acid
CPT	Cell Preparation Tube
DMEM	Dulbecco's Modified Eagle Medium
FBS	Fetal Bovine Serum
NEAA	Non-Essential Amino Acids
BCA	Bicinchoninic acid assay
TAE	Tris-Acetate-EDTA
HEK293T	Human Embryonic Kidney 293 cells with SV40 T-antigen
CACO2	Colorectal Adenocarcinoma Cells
RT-qPCR	Real-Time Quantitative Polymerase Chain Reaction
RNA-seq	RNA sequencing
RIN	RNA Integrity Number
CUT&RUN	Cleavage Under Targets and Release Using Nuclease
$\Delta\beta$	Delta Beta
PCA	Principal Component Analysis

rRNA	Ribosomal RNA
IGV	Integrative Genomics Viewer
DEG	Differentially Expressed Gene
log2FC	Log2 Fold Change
UTR	Untranslated Region
FDR	False Discovery Rate
GREAT	Genomic Regions Enrichment of Annotations Tool
TPM	Transcripts Per Million
GRIK5	Glutamate Ionotropic Receptor Kainate Type Subunit 5
GREM2	Gremlin 2, BMP Antagonist
UGT3A1	UDP Glucuronosyltransferase Family 3 Member A1
APBA2	Amyloid Beta Precursor Protein Binding Family A Member 2
KCNQ2	Potassium Voltage-Gated Channel Subfamily Q Member 2
OTUD7A	OTU Deubiquitinase 7A
PRXL2A	Peroxiredoxin Like 2A
PELI2	Pellino E3 Ubiquitin Protein Ligase Family Member 2
HOMER	Hypergeometric Optimization of Motif EnRichment
FBXL20	F-box And Leucine-Rich Repeat Protein 20
LINGO1	Leucine Rich Repeat And Ig Domain Containing 1
RUNX3	Runt Related Transcription Factor 3
CTNNB1	Beta Catenin
TSS	Transcription Start Site
RUNX1	Runt Related Transcription Factor 1
PSMA8	Proteasome Subunit Alpha 8
EVC	Ellis Van Creveld Syndrome
WNT1	WNT Family Member 1

WNT7A	WNT Family Member 7A
WNT10B	WNT Family Member 10B
APC	Adenomatous Polyposis Coli
iPSC	Induced Pluripotent Stem Cells

Chapter 3:

Abbreviation	Full Description
BOS	Bohring-Opitz syndrome
ASXL1	Additional Sex Combs Like 1
NPC	Neural Progenitor Cell
iPSC	Induced Pluripotent Stem Cell
NCC	Neural Crest Cell
VANGL	VANGL Planar Cell Polarity Protein
PCP	Planar Cell Polarity
ZIC	Zinc Finger Protein of the Cerebellum
DMEM	Dulbecco's Modified Eagle Medium
FBS	Fetal Bovine Serum
Pen/Strep	Penicillin/Streptomycin
SMAD	SMAD Family Member
BMP4	Bone Morphogenetic Protein 4
WNT	Wingless/Integrated
PFA	Paraformaldehyde
PAX6	Paired Box 6
RT	Reverse Transcription
PCA	Principal Component Analysis
DEG	Differentially Expressed Gene
MEG3	Maternally Expressed 3
NFIB	Nuclear Factor I B
SIK1	Salt-Inducible Kinase 1
TORC	Transducer of Regulated CREB Activity
GREM2	Gremlin 2

HOX	Homeobox
NMD	Nonsense-Mediated Decay
ASXL3	Additional Sex Combs Like 3
NPB	Neural Plate Border
SOX2	SRY-Box Transcription Factor 2
NES	Nestin

Chapter 4:

Abbreviation	Full Description
ASXL1	Additional Sex Combs Like 1
BOS	Bohring-Opitz syndrome
VANGL2	VANGL Planar Cell Polarity Protein 2
mTOR	Mechanistic Target Of Rapamycin
mTORC1	Mechanistic Target Of Rapamycin Complex 1
mTORC2	Mechanistic Target Of Rapamycin Complex 2
pS6K	Phosphorylated S6 Kinase
AML	Acute Myeloid Leukemia
BAP1	BRCA1 Associated Protein 1
AKT	AKT Serine/Threonine Kinase
ATF4	Activating Transcription Factor 4
ISR	Integrated Stress Response
NEAA	Non-Essential Amino Acids
ASNS	Asparagine Synthetase
Asn	Asparagine
SD	Standard Deviation
OCR	Oxygen Consumption Rate
TMRE	Tetramethylrhodamine Ethyl Ester
MTG	Mitotracker Green
FCCP	Carbonyl Cyanide p-(trifluoromethoxy)phenylhydrazone
RNA-seq	RNA Sequencing
CUT&RUN	Cleavage Under Targets and Release Using Nuclease
H3K4me3	Histone H3 Lysine 4 Trimethylation
ATF4	Activating Transcription Factor 4

MPC1	Mitochondrial Pyruvate Carrier 1
MPC2	Mitochondrial Pyruvate Carrier 2
iPSC	Induced Pluripotent Stem Cell
NPC	Neural Progenitor Cell
FGAR	Formylglycinamide-ribonucleotide
ECAR	Extracellular Acidification Rate

Chapter 5:

Abbreviation	Full Description
ASXL1	Additional Sex Combs Like 1
BOS	Bohring-Opitz syndrome
PRC	Polycomb Repressive Complex
BAP1	BRCA1 Associated Protein 1
PHD	Plant Homeodomain
H2AK119	Histone H2A Lysine 119
RNA-seq	RNA Sequencing
NEAA	Non-Essential Amino Acids
KIF22	Kinesin Family Member 22
VUS	Variant of Uncertain Significance
PCA	Principal Component Analysis
DEGs	Differentially Expressed Genes
NMD	Nonsense-Mediated Decay
TRH	Thyrotropin-Releasing Hormone
TSH	Thyroid Stimulating Hormone
RBP4	Retinol Binding Protein 4
ATF3	Activating Transcription Factor 3
REGG	Ras-like, Estrogen-regulated, Growth inhibitor
RRAGD	Ras-related GTP binding D
mTORC1	Mechanistic Target Of Rapamycin Complex 1
SP9	Sp9 Transcription Factor
ROS	Reactive Oxygen Species
TCA	Tricarboxylic Acid Cycle
COSMIC	Catalogue Of Somatic Mutations In Cancer

Chapter 6:

Abbreviation	Full Description
ASXL1	Additional Sex Combs Like 1
BOS	Bohring-Opitz syndrome
AML	Acute Myeloid Leukemia
AML-ASXL1	Acute Myeloid Leukemia with <i>ASXL1</i> mutation
VANGL2	VANGL Planar Cell Polarity Protein 2
GRIK5	Glutamate Ionotropic Receptor Kainate Type Subunit 5
GREM2	Gremlin 2
HOXA	Homeobox A
HOXB4	Homeobox B4
DEU	Differential Exon Usage
RUNX3	Runt-Related Transcription Factor 3
CMML	Chronic Myelomonocytic Leukemia
MDS	Myelodysplastic Syndromes
CHIP	Clonal Hematopoiesis of Indeterminate Potential
TCGA	The Cancer Genome Atlas
GDC	Genomic Data Commons
GTE _x	Genotype-Tissue Expression project
RNA-seq	RNA Sequencing
DNAm	DNA Methylation
p _{adj}	Adjusted P-value
log ₂ FC	Log ₂ Fold Change
PCA	Principal Component Analysis
FDR	False Discovery Rate
DEG	Differentially Expressed Gene

LRP5	Low-Density Lipoprotein Receptor-Related Protein 5
LRP6	Low-Density Lipoprotein Receptor-Related Protein 6
VUS	Variant of Uncertain Significance
TSS	Transcription Start Site
HSC	Hematopoietic Stem Cell
PLLIP	Plasmolipin
ST6GAL1	ST6 Beta-Galactosamide Alpha-2,6-Sialyltransferase 1
SERPINF1	Serpin Family F Member 1
PLEKHG6	Pleckstrin Homology Domain Containing G6
NCBI	National Center for Biotechnology Information
TBL1XR1	Transducin Beta-Like 1 X-Linked Receptor 1
DDX42	DEAD-Box Helicase 42
ROCK1	Rho-Associated Coiled-Coil Containing Protein Kinase 1
NBEAL2	Neurobeachin-Like 2
UGP2	UDP-Glucose Pyrophosphorylase 2
EZH2	Enhancer of Zeste Homolog 2
PRC2	Polycomb Repressive Complex 2
DNMT	DNA Methyltransferase
AML2	Acute Myeloid Leukemia type 2
RD	Runt Domain
CBF β	Core-Binding Factor Beta

Chapter 7:

Abbreviation	Full Description
BSyn	Bloom Syndrome
BLM	Bloom Syndrome Protein
SCE	Sister Chromatid Exchange
CHIP	Clonal Haematopoiesis of Indeterminate Potential
POT1	Protection of Telomeres 1
DNAm	DNA Methylation
GQ	Genotype Quality
VAF	Variant Allele Frequency
DP	Depth of Coverage
RefSeq	Reference Sequence
NCBI	National Center for Biotechnology Information
DNV	De Novo Variant
DNMT3A	DNA Methyltransferase 3 Alpha
ASXL1	Additional Sex Combs Like 1
TET2	Ten-Eleven Translocation 2
NOTCH1	Notch Receptor 1
CUX1	Cut-Like Homeobox 1
ATM	Ataxia Telangiectasia Mutated

ACKNOWLEDGEMENTS

This journey would not have been possible without the support, encouragement, and love from the remarkable group of people in my life.

First and foremost, my heartfelt thanks go to my family—my parents, who instilled in me the values of hard work and perseverance, my brother Andrew, whose phone calls and matcha espresso deliveries always seemed to show up at the most needed times, and to Alex, for being the guinea pig to my research ideas, and for the countless nights and weekends that you have kept me company for those “it will only take a few minutes” experiments. Your belief in me has been my constant source of strength.

My deepest gratitude to my graduate mentor, Dr. Valerie Arboleda, for seeing potential in me and welcoming me into her lab. Your guidance has been instrumental in my growth and accomplishments. I am profoundly grateful for the opportunity to learn and thrive under your guidance.

To my roommates, Jamie, Jen, Shannon, Aisha, and Ollie, thank you for making the home I loved going back to each night. To my labmates, especially Angela and Maneesha, I am incredibly grateful for your solidarity through the highs and lows of every day. Your support has made all the difference. To my mentors that continue to be there for me - Dr. David Raulet, my undergraduate PI at Berkeley, and Dr. Curtis Hodge, my postdoc mentor who taught me how to use pipettes with ease.

A toast to Cliffs of Id, my sanctuary and sanity amidst the chaos of research.

To my committee members, Julian, Harold, and Dinesh, I am forever grateful for your insightful advice and guidance, which have been pivotal in shaping this dissertation.

Finally, I would like to acknowledge the Department of Human Genetics and Pathology, the UCLA-Caltech MSTP programs for their funding and support, and the ASXL Rare Research Endowment (ARRE) for their funding and enthusiastic support of my work.

Chapter 1 is a version of: Lin I, Arboleda VA. The ASXL gene family: regulating development and growth through epigenetic mechanisms. *In preparation*.

Chapter 2 is a version of: Lin I, Wei A, Awamleh Z, Singh M, Ning A, Herrera A, REACH Biobank and Registry, Russell BE, Weksberg R, Arboleda VA. Multiomics of Bohring-Opitz syndrome truncating *ASXL1* mutations identify canonical and noncanonical Wnt signaling dysregulation. *JCI Insight*. 2023 May 22;8(10):e167744. doi: 10.1172/jci.insight.167744. PMID: 37053013; PMCID: PMC10322691.

Chapter 4 is a version of: Lin I, Krall AS, Matulionis N, Ning A, Stiles L, Christofk H, Russell BE, Arboleda VA. *ASXL1* mutations drive metabolic dysregulation and mitochondrial dysfunction in Bohring-Opitz syndrome. *In preparation*.

Chapter 5 is a version of: Lin I, Klein SD, Matulionis N, Krall AS, Corona de la Fuente R, Christofk H, Martinez-Agosto JA, Arboleda VA. Expanding the *ASXL1* mutation spectrum: Transcriptomic and metabolic dysregulation underlying overgrowth phenotype in inherited *ASXL1* missense mutations. *In preparation*.

Chapter 6 is a version of: Lin I, Awamleh Z, Wei A, Russell BE, Weksberg R, Arboleda VA. *ASXL1* mutations that cause Bohring-Opitz syndrome (BOS) or acute myeloid leukemia share epigenomic and transcriptomic signatures. *bioRxiv* 2022.12.15.519823.

Chapter 7 is a version of: Lin I, Wei A, Gebo TA, Boutros PC, Flanagan M, Kucine N, Cunniff C, Arboleda VA, Chang VY. Increased Frequency of Clonal Hematopoiesis of Indeterminate Potential in Bloom Syndrome Probands and Carriers. *medRxiv*. 2024 Feb 6:2024.02.02.24302163. doi: 10.1101/2024.02.02.24302163. PMID: 38370823; PMCID: PMC10871368.

VITA

EDUCATION

INSTITUTION AND LOCATION	DEGREE	START	COMPLETE	FIELD OF STUDY
University of California, Berkeley	BA	08/2014	05/2017	Genetics & Genomics
University of Cambridge		06/2016	08/2016	
University of California, Los Angeles	MD/PhD	08/2018	05/2026	Medicine
			(Expected)	Human Genetics

PUBLICATIONS

Lin I, Wei A, Gebo TA, Boutros PC, Flanagan M, Kucine N, Cunniff C, Arboleda VA, Chang VY. Increased Frequency of Clonal Hematopoiesis of Indeterminate Potential in Bloom Syndrome Probands and Carriers. *medRxiv*. 2024 Feb 6:2024.02.02.24302163. doi: 10.1101/2024.02.02.24302163. PMID: 38370823; PMCID: PMC10871368.

Lin I, Wei A, Awamleh Z, Singh M, Ning A, Herrera A; REACH Biobank and Registry; Russell BE, Weksberg R, Arboleda VA. Multiomics of Bohring-Opitz syndrome truncating ASXL1 mutations identify canonical and noncanonical Wnt signaling dysregulation. *JCI Insight*. 2023 May 22;8(10):e167744. doi: 10.1172/jci.insight.167744. PMID: 37053013.

Lin I, Awamleh Z, Wei A, Russell BE, Weksberg R, Arboleda VA. ASXL1 mutations that cause Bohring-Opitz syndrome (BOS) or acute myeloid leukemia share epigenomic and transcriptomic signatures. *bioRxiv*. 2022.12.15.519823; doi: <https://doi.org/10.1101/2022.12.15.519823>

Lin I, Afshar Y, Goldstein J, Grossman J, Grody WW, Quintero-Rivera F. Central 22q11.2 deletion (LCR22 B-D) in a fetus with severe fetal growth restriction and a mother with severe systemic lupus erythematosus: Further evidence of CRKL haploinsufficiency in the pathogenesis of 22q11.2 deletion syndrome. *Am J Med Genet A*. 2021 Oct;185(10):3042-3047. doi: 10.1002/ajmg.a.62346. Epub 2021 Jul 1. PMID: 34196458.

Bloom JS, Jones EM, Gasperini M, Lubock NB, Sathe L, Munugala C, Boeshaghi AS, Brandenburg OF, Guo L, Boocock J, Simpkins SW, **Lin I**, LaPierre N, Hong D, Zhang Y, Oland G, Choe BJ, Chandrasekaran S, Hilt EE, Butte MJ, Damoiseaux R, Cooper AR, Yin Y, Pachter L, Garner OB, Flint J, Eskin E, Luo C, Kosuri S, Kruglyak L, Arboleda VA. Swab-Seq: A high-throughput platform for massively scaled up SARS-CoV-2 testing. *Nature Biomedical Engineering*, 5 (7). pp. 657-665. ISSN 2157-846X. PMCID: PMC7480060. doi:10.1038/s41551-021-00754-5.

AWARDED FUNDING / GRANTS

2023 *UCLA Leena Peltonen Award for Excellence in Human Genetics*
2023 *ASXL Rare Research Endowment Speaker Travel Award*
2022 *ASHG Stephen Cederbaum Award*

PEER REVIEWING

2024 *Central European Journal of Immunology*

INVITED ORAL PRESENTATIONS (SELECTED)

2023 *ESHG 2023 Invited Speaker*

Lin I, Wei A, Singh M, Awamleh Z, Weksberg R, Russell B, **Arboleda VA**. Truncating ASXL1 mutations in Bohring-Opitz syndrome cause epigenetically driven upregulation of Wnt signaling. Oral presentation at ESHG 2023. 06/13/23. Glasgow, United Kingdom.

2023 *ARRE 2023 Invited Speaker*

Lin I, Wei A, Singh M, Awamleh Z, Weksberg R, Russell B, **Arboleda VA**. Multi-omics of Bohring-Opitz syndrome truncating ASXL1 mutations identify canonical and non-canonical Wnt signaling dysregulation. Oral presentation at ARRE 2023. 07/21/23. Boston, MA.

CHAPTER 1

INTRODUCTION:

The ASXL Gene Family:

Regulating Development and Growth Through Epigenetic Mechanisms

"In every walk with nature, one receives far more than he seeks." - John Muir

Introduction to ASXL genes and their discovery

The Additional Sex Combs-Like (ASXL) gene family, named after the Additional Sex Combs (*Asx*) gene in *Drosophila melanogaster*, encompasses a group of evolutionarily conserved proteins that are instrumental in chromatin remodeling and transcriptional regulation. These proteins are not only crucial in controlling developmental pathways, but also play significant roles in oncogenesis ¹. Initially recognized for their transcriptional regulatory functions, the importance of ASXL genes has expanded dramatically, implicating them in a broad spectrum of human diseases from developmental disorders to various cancers and hematopoietic malignancies. Foundational studies by Jurgens (1985) and Sinclair (1998) laid the groundwork for our current understanding, highlighting the need to delve deeper into the functions of the ASXL gene family and their complex roles in health and disease ^{2,3}.

ASXL homologs function in Polycomb repression and homeotic transformation in model organisms

The initial discovery of the ASXL gene family through genetic screens in *Drosophila melanogaster* revealed their essential role in segment identity determination and sex determination ^{2,3}. Early work demonstrated the critical role of *Asx* in anterior-posterior patterning ⁴, mediated through its interaction with Polycomb Group (PcG) proteins ⁵⁻⁷, which are pivotal in the stable repression of homeotic genes during early developmental stages ⁸⁻¹².

Asx remained fairly elusive in characterization of its role as a PcG protein. The characterization of *Asx* as a PcG protein was further defined with the identification of its interaction with the PcG gene *Calypso* (and its mammalian counterpart *BAP1*) ¹³, forming the Polycomb Repressive Deubiquitinase (PR-DUB) complex ¹⁴. This complex was found to be bound at PcG target genes in *Drosophila*, and is critical for the deubiquitination of histone H2A and the repression of homeobox (*Hox*) genes ¹⁴.

The PcG proteins comprise a large and diverse family that epigenetically suppresses transcription of key developmental genes, organized into three primary polycomb repressive complexes (PRCs): PRC1, PRC2, and PR-DUB. Each complex modifies chromatin through distinct mechanisms, relying on variable compositions of core and accessory subunits. PRC1 complexes catalyze histone 2A lysine 119 (H2AK119ub) ubiquitination, while PRC2 deposits the histone 3 lysine 27 (H3K27) trimethylation mark, stimulating both PRC1 and PRC2 functions. This interaction creates a feedforward loop that likely amplifies polycomb signals across large genomic regions to ensure transcriptional repression ¹⁵. In contrast, PR-DUB reverses the H2AK119 ubiquitination imposed by PRC1, regulating polycomb-targeted repression. The distinctive functions of each group emerge from their structural diversity, where PRC1 acts as an E3 ubiquitin ligase that monoubiquitinates H2AK119 and independently compacts chromatin, potentially interacting directly with transcription machinery. PR-DUB, on the other hand, removes H2AK119 ubiquitination, and PRC2 targets H3K27 for methylation ^{14,16}.

The core of PRC1 is formed by chromobox (CBX) proteins, along with one member from each of the polycomb group ring finger (PCGF; PCGF1–PCGF6), ring finger protein (RING1; RING1a and RING1b), and polyhomeotic homolog (HPH; HPH1–HPH3) families ^{17,18}. Notably, PCGF1 is part of a variant PRC1 complex that includes the transcriptional repressor B cell lymphoma 6 (BCL6), the BCL6 corepressor (BCOR) and the histone lysine demethylase 2B (KDM2B). The essential components of PRC2 include enhancer of zeste 2 (EZH2) or its close homolog EZH1, embryonic ectoderm development (EED), and suppressor of zeste 12 (SUZ12) ¹⁹. The PR-DUB is formed by the ASXL1-3 proteins and deubiquitinase breast cancer type 1 susceptibility protein (BRCA1)-associated protein-1 (BAP1). The core PR-DUB proteins function alongside other subunits such as the transcription factors forkhead box K1 and K2 (FOXK1/2), transcriptional co-regulator host cell factor C1 (HCFC1), glycosyltransferase O-linked N-acetylglucosamine transferase (OGT), methyl-CpG-binding domain 5 and 6 (MBD5/6), and

lysine demethylase 1B (KDM1B), which control specific de-ubiquitination activities vital for gene regulation^{20,21}.

Mutations in PcG genes, including *Asx*, disrupt gene repression, and lead to posterior homeotic transformations²². Typically, mutations in one PcG gene can amplify the effects of mutations in other PcG genes, suggesting overlapping functions or complex interactions^{23,24}. This phenomenon was observed in doubly heterozygous adult *Drosophila* and embryos, where dominant mutant *Asx* alleles produced males with an extra sex combs phenotype and head defects². Sinclair et al. (1998) noted that these *Asx* mutations generally behaved as gain-of-function, with phenotypes more severe than those seen in homozygous deletion mutants^{3,4}.

The mutation spectrum of *Asx* reveals a dichotomy, suggesting roles beyond its involvement in the PcG complex³. Divergent phenotypes in zygotic and germline mutant embryos^{9,12}, as well as mutations driving both anterior and posterior transformations^{22,24,25}, underlie this point. Typically, anterior transformations are characteristic of mutations in the trithorax Group (*trxG*) genes^{22,24,25}, whereas posterior transformations align with a PcG gene function. This dual role highlights the dynamic regulation of gene expression that is critical for the precise control of developmental processes^{3,26}.

PcG and *trxG* proteins maintain an important balance between repression and activation of numerous genes, including the *HOX* genes, which define body segments along the anterior/posterior axis in animals. PcG proteins, including *Asx*, are vital in the stable repression of *Hox* genes during embryogenesis, ensuring accurate segmental identity¹⁴.

The dual function of *Asx*, which contributes to gene repression as part of the PcG complex and gene activation as part of the TrxG complex, as well as its ability to enhance phenotypes of both PcG and TrxG gene mutations when mutated, characterizes *Asx* as an Enhancer of Trithorax and Polycomb (ETP) protein²⁷. The ETP proteins recruit PcG and TrxG

complexes to target chromatin, thus fine-tuning gene expression through both repression and activation pathways ²⁷.

The evolutionary conservation of *ASXL* genes across species highlights their fundamental roles not only in development but also in disease pathology. The elucidation of *Asx* function in *Drosophila* paved the way for the identification and characterization of *ASXL* genes in other organisms, including mammals, and emphasizes the importance of these genes in health and disease ²⁸.

Gene structure of the *ASXL* family

In mammals, the *ASXL* gene family consists of three members: *ASXL1*, *ASXL2*, and *ASXL3* (**Figure 1-1**). The discovery of these genes was driven by comparative genomics and functional studies aimed at understanding the genetic basis of developmental disorders and cancer ^{1,29,30}. *ASXL* genes are homologs of the *Drosophila Asx* gene, and share conserved structural domains and functional motifs essential for their functions of protein-protein interactions, chromatin binding, and enzymatic activities ^{29,30}.

The *ASXL* proteins are characterized by their architectural complexity, including large stretches of intrinsically disordered regions (IDRs) that facilitate flexible interactions with various protein partners. Conserved across *Drosophila Asx* and the vertebrate *ASXL* gene family are the ASX homology (ASXH) and plant homeodomain (PHD) domains ^{29–31}. The ASX N-terminal (ASXN) domain is conserved among vertebrate *ASXL* homologues but not *Drosophila Asx* ³². Both the ASXN and PHD domains act as putative DNA- or histone-recognition modules. Meanwhile, the central regions of the proteins, which encompass the ASXH, ASXM1, and ASXM2 domains, serve as protein–protein interaction modules. These modules facilitate the assembly of critical regulatory complexes with the respective components BAP1, lysine demethyltransferase 1A (KDM1A), nuclear receptor coactivator 1 (NCOA1), as well as nuclear hormone receptors (NHRs), thereby influencing transcriptional outcomes and chromatin states ¹.

These domains are pivotal for DNA or histone recognition, and protein-protein interactions, thereby influencing chromatin states and transcriptional outcomes.

The ASXN domain, positioned at the N-terminus of ASXL proteins, mediates interactions with other chromatin-associated proteins ^{29,30}. The N-terminal HARE-HTH (HBI, ASXL1, restriction endonuclease helix-turn-helix) domain serves as a scaffold for the assembly of protein complexes involved in transcriptional regulation and chromatin remodeling ³². Through these interactions, ASXN facilitates the recruitment of ASXL proteins to specific genomic loci, where they exert their regulatory functions on gene expression.

The ASXH domain, centrally located within ASXL proteins, is essential for protein-protein interactions, particularly with components of the PRC2 and the PR-DUB complex ¹⁴. The ASXH domain contains a characteristic nuclear receptor-coactivator interaction motif LXXLL that mediates protein interactions, and a deubiquitinase adaptor (DEUBAD) domain that activates BAP1. This domain binding function is critical for the recruitment of these complexes to their target genes, facilitating chromatin modifications that influence gene expression patterns crucial for development ³².

The ASXM1 and ASXM2 domains are also conserved among the ASXL family members ¹. The ASXM1 facilitated bindings with NCOA1, influencing gene transcription ³³, while ASXM2 contains the LVxxLL motif, a binding site for nuclear hormone receptors (NHRs) such as RAR α ³³.

The PHD, found at the C-terminus of ASXL proteins, serves as a chromatin-binding module, recognizing and “reading” specific histone modifications such as histone H3 lysine 4 trimethylation (H3K4me3) and facilitating the recruitment of ASXL proteins to target genomic regions, thereby regulating gene expression in a context-dependent manner ^{34–36 37}.

In addition to these conserved domains, ASXL proteins interact with various transcriptional regulators, chromatin-modifying enzymes, and signaling molecules to regulate cellular homeostasis. Within chromatin remodeling complexes such as the PRC2 and the

BAP1-containing complexes, ASXL modulates histone modifications, gene expression, and developmental processes. Furthermore, ASXL proteins exhibit context-dependent activities, acting as both transcriptional activators and repressors ^{3,25,38}, thereby influencing cell fate decisions and tissue homeostasis in a dynamic manner.

Recent structural predictions from AlphaFold3 have provided insights into the structural configuration of ASXL proteins, indicating significant disorder but with regions of lower predicted alignment error in the conserved ASXH and ASXN domains ³⁹. This suggests these domains are structural islands within a largely unstructured sequence, crucial for the proteins' versatile interaction capabilities. (**Figure 1-2A, Figure 1-2B, Figure 1-2C**). The rest of the protein exhibited high predicted alignment error, highlighting extensive IDRs that challenge structural prediction and characterisation, but also likely contribute to the versatile interaction capabilities of the ASXL family.

The interaction with BAP1 appears to induce a transformative effect on the structural configuration of the ASXL proteins. While isolated ASXL proteins exhibited very low template modeling (TM) scores and interface predicted template modeling (ipTM) scores (ASXL1 ipTM = NA, pTM = 0.22; ASXL2 ipTM = NA, pTM = 0.24; ASXL3 ipTM = NA, pTM = 0.22) indicating a high degree of intrinsic disorder, their complexes with BAP1 are predicted to adopt a more structured and organized conformation (ASXL1-BAP1 ipTM = 0.55, pTM = 0.36; ASXL2-BAP1 ipTM = 0.53, pTM = 0.36; ASXL3-BAP1 ipTM = 0.60, pTM = 0.36) (**Figure 1-2D, Figure 1-2E, Figure 1-2F**). This is evident from the increased ipTM scores when bound to BAP1, which rises significantly compared to the isolated forms of ASXL1, ASXL2, and ASXL3. The integration of these AlphaFold3 data suggests that this structural stabilization through BAP1 enhances the functional interaction capacity of the ASXL genes, critical for their role in chromatin modification and gene regulation.

Diverse roles in development and disease

The *ASXL* gene family is pivotal in regulating a wide array of biological processes, from embryonic development to epigenetic regulation. These genes significantly influence gene expression regulation, cellular identity, cell fate decisions, and are intricately linked to various diseases, including cancer and neurodevelopmental disorders ³⁷. The regulatory balance maintained by *ASXL* genes between stem cell pluripotency and differentiated states, particularly through *HOX* gene regulation, underscores their importance. This balance is critical for ensuring precise spatial and temporal gene expression patterns vital for development ^{40–43}.

Studies using knockdown and mutant models have demonstrated that this can cause derepression of posterior *HOXA* genes, primarily due to reduced recruitment of the PRC2 to the *HOXA* locus, and high embryonic lethality ^{44,45} (**Table 1-1**). Furthermore, studies across various animal models have illuminated the roles of *ASXL* genes in crucial processes including hematopoiesis ^{46,47}, neural crest development ⁴⁸, and cardiac development ⁴⁹.

In oncology, disruption to *ASXL1* and *ASXL2* leads to hematopoietic defects and the expansion of myeloid cells ⁵⁰. Loss of *Asx1/1* or *Asx1/2* results in impaired and accelerated development of myeloid neoplasms in vivo ^{44,46,50,51}. Mice heterozygous for a truncating mutations in *Asx1/1*, which mimics pathologic mutations identified in humans ^{42,52–54}, exhibit more subtle effects than those observed in knockouts, knockdowns, or overexpression models of truncated mutant *ASXL1*. Although these mice do not develop hematological malignancies within 18 months, longer studies show that they eventually develop a phenotype resembling myelodysplastic/myeloproliferative neoplasms (MDS/MPN) after 18-24 months, highlighting the long-term impacts of *ASXL1* mutations on hematopoietic diseases ^{42,54}.

Somatic mutations in *ASXL* genes contribute to an aberrant epigenetic landscape characteristic of many tumors, affecting cellular processes like proliferation, survival, and therapy resistance. The dual functionality of the *ASXL* genes, as both a tumor suppressor and an oncogene in myeloid malignancies, illustrate its complex role in disease pathogenesis ^{44,55}.

Diverse biological functions of ASXL genes on development and metabolism

The ASXL gene family plays a pivotal role in metabolic regulation and cellular energy balance, particularly in response to metabolic stress (**Table 1-1**). This function is largely mediated through interactions with peroxisome proliferator-activated receptor gamma (PPAR γ), a nuclear receptor essential for insulin sensitivity and adipocyte differentiation. *ASXL1* typically represses, while *ASXL2* enhances the expression of adipogenic genes, many of which are targets of PPAR γ . Together, these activities regulate adipogenesis and glucose homeostasis^{56,57}.

ASXL gene function is also critical for neural development^{48,58}. Truncating mutations in *ASXL1* disrupt the activation of transcription factors ZIC1 and ZIC4, affecting the migration of neural crest cells, while knockdown of *ASXL3* in *Xenopus* embryos leads to neural tube defects and diminished expression of posterior neural markers^{48,58}. Additionally, reductions in *ASXL3* expression have been associated with compromised neural cell proliferation, cell cycle progression, and apoptosis⁵⁹.

Deficiencies in *ASXL1* and *ASXL2* result in severe cardiac issues. *ASXL1*-deficient mice present with ventricular septal defects, and *ASXL2*-null mice exhibit congenital heart malformations, enlarged hearts, and early mortality. These findings underscore their critical roles in heart formation and function^{49,50,60}. Further studies have shown that the absence of *ASXL2* leads to spontaneous cardiac tissue overgrowth, without cardiomyocyte hypertrophy, suggesting *ASXL2*'s essential role in regulating heart tissue proliferation and differentiation^{49,61}. This role in cardiac function is echoed in human studies, where eight out of 12 Shashi-Pena syndrome (SHPNS) patients, a rare disorder caused by germline truncating mutations in *ASXL2*, had cardiac abnormalities⁶².

Recent in vitro studies have further elaborated the roles of *ASXL3* in cardiac biology. Fu et al. (2021) identified a novel compound heterozygous missense mutation in the *ASXL3* gene in patients with congenital heart disease (CHD), which, when overexpressed in murine

cardiomyocyte HL-1 cells, led to increased apoptosis and reduced cell viability, impacting cardiac structure and fibrosis in mice ⁶³. Liu et al. (2023) further expanded upon these findings and showed that these mutations in *Asx/3* inhibited the proliferation of cardiomyocytes and accelerated cell apoptosis by promoting the expression of long non-coding RNAs, suppressing the formation of FGFR2 transcripts, and inhibiting the Ras/ERK signaling pathway in mouse cardiomyocytes ⁶⁴.

Furthermore, significant findings related to skeletal health have been observed, where the loss of *Asx/1* or *Asx/2* leads to marked changes in bone mass and skeletal homeostasis. A marked decrease in bone mass, reduction in bone marrow stromal cells (BMSCs), and altered differentiation patterns have been noted with the global loss and conditional deletion of *Asx/1* in osteoblasts and their progenitors. Reintroduction of *Asx/1* can restore the self-renewal capabilities of *Asx/1*-deficient BMSCs, demonstrating its critical role in skeletal homeostasis ⁶⁵. Similarly, *Asx/2*-null mice develop osteopetrosis, highlighting dysregulations in skeletal homeostasis driven by the loss of *Asx/2* ⁵⁷.

These studies reveal the ASXL genes' indispensable roles in maintaining cellular and tissue integrity across multiple systems, thereby influencing both development and disease progression.

Binding partners and histone functions of the ASXL gene family

The ASXL gene family profoundly influences chromatin architecture through interactions with critical complexes such as PRC2 and PR-DUB, playing pivotal roles in histone modification and gene repression (**Table 1-1**). These interactions are essential for regulating gene expression patterns that respond to environmental changes, aging, and stress ^{31,66–69}. ASXL proteins notably enhance the enzymatic activity of PRC2, which is responsible for the trimethylation of H3K27me3, a mark typically associated with gene repression (**Figure 1-3**) ^{40,67,68}. In studies using mouse embryonic stem cells (mESCs), ASXL1, ASXL2, and ASXL3 have

all been identified as binding partners for PR-DUB⁶⁶. Through its association with PR-DUB, ASXL proteins counteract the ubiquitination effects mediated by PRC1, regulating key gene expression patterns essential for cellular differentiation and development^{31,69}.

The interaction of truncated mutant ASXL1 with BAP1 forms a hyperactive complex, promoting spontaneous differentiation to the mast cell lineage and aberrant myeloid differentiation in hematopoietic progenitor cells. This hyperactive complex significantly reduces levels of H2AK119ub and H3K27me3, with a selective upregulation of genes through the presence of H3K4me3 at gene promoters, thus affecting the regulation of posterior *HOXA* genes by removing H2AK119 ubiquitination^{55,70}. Additionally, knockdown or knockout of ASXL1 has been shown to substantially decrease H3K27me3 levels due to loss of interactions with the PRC2, along with decreased global levels of H3K4me3^{40,44,46,71}. These changes highlight the crucial role of ASXL1 in maintaining proper chromatin remodeling dynamics.

Further research has explored how wildtype ASXL1 interacts with other molecular partners, such as PCGF4, a component of the PRC1 complex and also known as BMI-1, and OGT, which stabilizes ASXL1 through glycosylation (**Figure 1-3**). These interactions are crucial for the functional integrity and stability of ASXL1^{54,71}. However, truncated forms of ASXL1 lack the ability to bind to PCGF4, leading to disrupted chromatin modulation and gene expression⁵⁴, providing insights into the pathophysiology of truncating mutations in ASXL1.

Signaling pathways regulated by the ASXL gene family

ASXL genes are integral to various signaling pathways critical for cellular function and development. Recent research has revealed significant interactions between ASXL1 and the Wnt signaling pathway, where truncating mutations lead to the aberrant activation of both the canonical Wnt pathway and the non-canonical planar cell polarity pathway. This dysregulation has profound implications, as the Wnt pathway is critical for numerous developmental processes and its disruption can lead to various developmental disorders⁷². The canonical Wnt

pathway typically promotes cell proliferation and differentiation, whereas the planar cell polarity pathway influences the orientation of cells within the plane of a tissue.

ASXL1 mutations have also been shown to impact the mTOR signaling pathway, leading to overactive Akt/mTOR signaling, and aberrant cellular proliferation⁷³. This finding is consistent with the study by Youn et al. (2017), which showed that wild type *ASXL1* interacts with phosphorylated AKT, influencing the AKT-E2F regulatory axis and contributing to cellular senescence. Together, these interactions illustrate how *ASXL1* can act to both promote and regulate cellular functions, impacting growth and aging⁴⁵.

Furthermore, *ASXL2* modulates retinoic-acid induced transcription, which is essential for gene expression during embryonic development. Khan et al. (2014) reported that while the C-terminus of *ASXL2* enhances retinoic-acid induced transcription, the full-length protein does not⁷⁴. This suggests that *ASXL2* may have regulatory domains that are contextually activated or inhibited.

ASXL3 has also been implicated in influencing cell proliferation and apoptosis via interactions with the PI3K-Akt pathways⁷⁵, as well as the Ras/ERK signaling pathway through the modulation of lncRNAs and FGFR2 transcripts, leading to altered cell proliferation and apoptosis⁶⁴.

Overall, the *ASXL* family of genes emerges as crucial regulators within multiple signaling pathways, impacting cellular behavior and development. Through animal and cell models, these studies provide a detailed map of the *ASXL* genes' functions, especially in chromatin remodeling, transcriptional governance, and developmental pathways.

Regulatory mechanisms governing *ASXL* proteins

The regulatory mechanisms controlling *ASXL* gene expression are complex and appear to be influenced by various developmental and environmental factors. Recent studies in proteomics using methods such as nano liquid chromatography tandem mass spectrometry

(nano LC-MS/MS), have provided insights into ASXL1's protein interactions and post-translational modifications. For example, the protein degradation of ASXL1 is influenced by its interaction with cell division cycle 20 (CDC20), a coactivator of the anaphase-promoting complex (APC/C), an E3 ubiquitin ligase complex that regulates cell cycle progression. Overexpression of CDC20 promotes ASXL1 degradation, whereas its depletion through the clustered regularly interspaced short palindromic repeats (CRISPR)-Cas9 system leads to increased ASXL1 expression ⁷⁶. Through immunoprecipitation of FLAG-tagged ASXL1 mutant plasmid overexpression and co-transfection assays of ASXL1 coupled to luciferase reporter, ASXL1 was also identified to interact with other molecular partners such as PCGF4/BMI-1 and OGT. These interactions are essential for maintaining the functional integrity and stability of ASXL1 ^{54,71}. Furthermore, ASXL1 is subject to ubiquitination at lysine 351 (K351) which leads to its degradation, playing an important role in ASXL1 protein turnover. On the other hand, interaction with ubiquitin-specific peptidase 7 (USP7) stabilizes the ASXL1 protein ⁷⁷.

Furthermore, the interaction of ASXL proteins with chromatin-modifying complexes such as PRC2 and PR-DUB suggests an epigenetic regulatory layer that might include feedback mechanisms to control their own expression. Preliminary data from our studies suggests that *ASXL1* regulates gene expression of *ASXL3* in neural tissues. Overall, while substantial progress has been made in understanding the interactions and modifications of ASXL proteins, much remains to be discovered about the comprehensive regulatory networks that govern their function.

Mechanisms of truncating mutation-driven dysfunction

Truncating mutations in *ASXL* genes result in prematurely terminated proteins and are critically linked to various diseases, including myeloid malignancies and rare neurodevelopmental disorders. These mutations disrupt normal ASXL protein functions, leading to aberrant gene regulation and cellular dysfunction. The proposed mechanisms through which

these truncating mutations exert their effects involve alterations in chromatin remodeling, disruption of epigenetic regulation, and interference with normal protein-protein interactions^{40,55,72}.

ASXL proteins, integral to chromatin remodeling complexes like PRC2 and PR-DUB, are crucial for regulating gene expression through histone modifications⁴⁰. Truncating mutations in ASXL genes can lead to the production of proteins that lack functional domains necessary for interactions with these chromatin remodeling complexes, impairing the recruitment of these complexes to their target genes (**Figure 1-3**). For instance, the loss of the ASXH domain or the PHD finger in ASXL proteins can prevent the recruitment of PRC2 to target genes, leading to reduced H3K27me3 and aberrant gene activation that would normally be repressed⁴⁰. Similarly, disruption of the PR-DUB complex due to truncating mutations of ASXL proteins may affect the deubiquitination of histone H2A, altering gene expression patterns critical for cell differentiation and proliferation^{31,78}.

Additionally, truncating mutations can create aberrant ASXL protein products that lose the ability to interact with key regulatory partners or gain new, potentially pathogenic interactions. This can disrupt normal signaling pathways and gene regulatory networks, contributing to the development and progression of disease. For instance, truncated forms of ASXL1 observed in myeloid malignancies have been proposed to aberrantly recruit or stabilize repressive complexes at target genes, leading to the silencing of tumor suppressor genes or the activation of oncogenes, thereby promoting leukemogenesis^{1,55,79}.

Some proposed mechanisms of action for truncating mutations in ASXL genes include a gain of function effect, dominant negative effect—where the truncated protein interferes with the function of the wild-type protein—and haploinsufficiency—where a single functional copy of the gene is insufficient to maintain normal cellular function^{52,77,80}.

Early research in ASXL3 posited that pathogenic truncating variants might primarily exert a dominant negative effect. This hypothesis stemmed from the observation that these mutations,

and other truncating mutations found across the other *ASXL* gene family, often cluster in the penultimate exon. This effectively separates the N-terminal protein scaffolding functions from the C-terminal chromatin/DNA targeting activities, thereby disrupting both domains' functional integrity^{81,82}. However, Srivastava et al. (2015) found that mRNA transcripts from the c.1448dupT variant of *ASXL3* were prone to nonsense-mediated decay (NMD), leading to reduced *ASXL3* expression and suggesting that haploinsufficiency might also play a significant role in the disease mechanism⁶⁹. Yet, in analysis of cells from 3 unrelated patients with SHAPNS, Shashi et al. (2016) showed that the mutant transcripts were expressed and not subject to NMD, supporting a dominant-negative effect⁸³.

Thus, truncating mutations in *ASXL* genes disrupt normal cellular function through various mechanisms, including the impairment of chromatin remodeling activities, interference with protein-protein interactions, and the induction of dominant negative effects or haploinsufficiency.

Pathogenic *ASXL* germline mutations: inheritance and mutation types

Truncating germline mutations in *ASXL* genes manifest in distinct syndromes, notably BOS (#MIM 605039, *ASXL1*, MIM# 612990)^{84–86}, SHAPNS (MIM#617190, *ASXL2*, MIM #612991)⁸³, and Bainbridge-Ropers syndrome (BRPS, MIM#615485, *ASXL3*, MIM# 617768)⁸¹. These syndromes are characterized by developmental delays, intellectual disabilities, and unique morphological traits. Typically occurring *de novo*, these mutations often appear in the penultimate exon of the genes, affecting the protein's structure and function dramatically (**Figure 1-1, Table 1-2**)^{81,83,87}. Notably, some mutations such as p.Arg404* in *ASXL1* (n=5 *de novo*) and p.Arg1036* in *ASXL3* (n=4 *de novo*, n=1 unknown inheritance, n=2 likely inherited), have been recurrently identified in multiple, unrelated patients.

Interestingly, to date, only one truncating mutation causing BOS, SHPNS, or BRPS has been identified within the PHD domain of these genes. All other pathogenic truncating mutations

identified have been found 5' to the PHD. The mutation, a heterozygous frameshift mutation (p.Ser2238Thrfs*3) in ASXL3, was identified in a 47-year-old patient presenting with atypical features of BRPS, including severe behavioral problems from early adolescence, autism spectrum disorder, and adult-onset tonic clonic epilepsy. This case was further complicated by a family history of intellectual disability, affecting the patient's parents, one sibling sharing the same parents, and three maternal half-siblings. The full sibling also suffered from severe behavioural problems, autism spectrum disorder, and childhood-onset tonic clonic epilepsy. Unfortunately, the parents and full sibling were deceased at the time of testing for the proband, making it impossible to verify the inheritance pattern of the mutation in the proband ⁸⁸.

Despite significant impacts on disease pathogenesis, the relationship between specific ASXL genotypes and clinical phenotypes remains poorly understood. Extensive studies, such as those on BOS patients, have yet to clarify this relationship, highlighting the multifactorial influences on the manifestation of these disorders ⁸⁹. Similarly, a study by Yu et al. (2021) in BRPS patients did not demonstrate any genotype-phenotype differences, nor significant differences in brain malformation or presence of seizure based on genotype ⁸⁰.

Recent studies have broadened our understanding of ASXL mutations, identifying atypical and novel variants that challenge our comprehension of these genes' roles in development and disease. While mutations causing BOS, SHPNS, and BRPS are typically found *de novo*, inherited mutations have also been identified in both BOS and BRPS (**Table 1-2**), often inherited from asymptomatic carriers, indicating variable penetrance and expressivity ⁹⁰. The absence of mosaicism in blood samples does not rule out its presence in other tissues, suggesting more complex genetic inheritance patterns that may include gonadal mosaicism, as evidenced by *de novo* pathogenic variants in non-twin siblings with the absence of mutations identified in the parents ^{91,92}.

Missense and mosaic variants add further complexity to the genetic landscape. Some missense variants identified cause entirely distinct phenotypes, while other missense variants

drive phenotypes similar to those of truncating variants, but with other additional atypical features ⁹³. Studies have also identified mosaic variants, found in varying percentages in blood and saliva samples, indicate that even a 30% mosaicism rate can significantly influence phenotypes ⁹². An example includes a patient with a de novo pathogenic *ASXL3* variant, present in 30-35% of blood and saliva samples, who exhibited considerable developmental challenges ⁹².

Duplications in the 20q11.2 band, which includes the *ASXL1* gene, have been implicated in a newly characterized microduplication syndrome with developmental delays and distinct craniofacial features ⁹⁴. Additionally, mutations in *KLHL7* (kelch-like family number 7) have been identified in patients with BOS-like phenotypes, demonstrating that similar clinical presentations can originate from mutations in different genes ⁹⁵.

Clinical consequences of somatic *ASXL* mutations in malignancies

Somatic mutations in *ASXL* genes are key drivers in a variety of cancers, with *ASXL1* mutations particularly associated with poor prognosis in myeloid malignancies such as myelodysplastic syndromes (MDS), acute myeloid leukemia (AML), and chronic myelomonocytic leukemia (CMML) ⁹⁶⁻⁹⁹. These mutations are predominantly frameshift or nonsense mutations that cause truncation of the protein before the PHD domain, and result in a stable C-terminal truncated product ^{41,98}. Found in approximately 10-15% of patients with AML and up to 40% of patients with MDS, these mutations are linked to poor prognosis, rapid disease progression, treatment resistance, higher rates of relapse, and decreased overall survival.

Conversely, germline mutations in *ASXL1* which cause BOS do not appear to increase the risk of myeloid malignancies in these individuals. However, several other cancers have been reported in patients with BOS, including medulloblastoma, medulloblastoma ¹⁰⁰, Wilms tumor and hepatoblastoma ^{89,101,102}. Notably, the 7% incidence of renal neoplastic processes suggests a link between BOS and bilateral Wilms tumor, prompting established screening protocols for

early detection ¹⁰¹. Additionally, hepatoblastoma, a rare but aggressive liver cancer in children, has been documented in three BOS cases. Given its severity and the potential for metastasis, updated screening guidelines recommend vigilant monitoring for hepatoblastoma in BOS patients, despite its less than 1% occurrence rate.

Mutations in *ASXL2* are implicated in both SHPNS and a range of malignancies, including AML, as well as bladder, prostate, breast, and pancreatic cancers ^{1,83}. In pediatric populations, somatic *ASXL2* mutations are primarily observed in hematologic malignancies, particularly AML cases involving mutations in core binding factors ¹⁰³. Notably, about 23% of AML cases with t(8;21)(q22;q22); RUNX1-RUNX1T1 gene fusions harbor *ASXL2* mutations, making it the second most common mutation after the *KIT* gene ^{47,104}. However, *ASXL2* mutations are rarely found in other core-binding factor AML types ¹⁰⁵. Intriguingly, unlike *ASXL1* mutations which are predominantly found in older patients, *ASXL2* mutations occur at comparable rates in both pediatric and adult cases. Although these mutations do not significantly impact overall survival, they are associated with a higher rate of disease relapse, demonstrating its influence on the clinical course of these diseases ^{104,105}.

ASXL3 mutations are also linked to a variety of solid tumors, including melanomas and lung adenocarcinomas, despite its rare occurrence in myeloid neoplasms ^{1,106,107}. Only a few cases of *ASXL3* mutations in myeloid neoplasms have been documented in The Cancer Genome Atlas (TCGA) program study ¹⁰⁸. One particularly noteworthy case involved a pediatric AML patient with an *ASXL3* mutation that expanded significantly from less than 1% at diagnosis to 30% at relapse, illustrating its potential impact on disease progression and therapy resistance ¹⁰⁹.

While *ASXL2* and *ASXL3* play less definitive roles in cancers compared to *ASXL1*, emerging research suggests their potential oncogenic or tumor-suppressive effects vary by cancer type ^{37,110,111}. For instance, elevated expression of *ASXL2* is associated with poor prognosis in breast ¹¹², colorectal ¹¹³ and pancreatic cancers ¹¹⁴. Conversely, overexpression of

ASXL3 was correlated with increased genomic instability in small cell lung cancer (SCLC) lines¹⁰⁷, potentially through protein-protein interactions with BAP1 and bromodomain-containing protein 4 (BRD4)¹⁰⁶. These interactions were highlighted in the study by Tsuboyama et al. (2022), which demonstrated that dysregulations in the BAP1 histone H2A deubiquitinase complex, involving the BAP1/ASXL3/BRD4 axis, contribute to the pathogenesis of human SCLC. Pharmacologic targeting of this axis disrupts ASXL3 function¹¹⁵, thereby affecting SCLC cell fate by driving overexpression of the transcription factor paired box protein 9 (PAX9)¹¹⁶.

Furthermore, Shukla et al. (2017) provided insights into epigenetic mechanisms contributing to pulmonary carcinogenesis by demonstrating that ASXL3 was markedly upregulated in induced pluripotent stem cells from normal human small airway epithelial cells and in SCLC lines, correlating with increased genomic copy number and tumorigenic properties¹⁰⁷.

This evolving understanding highlights the intricate relationship between normal development, epigenetic regulation, and oncogenesis driven by ASXL genes, highlighting their significant roles across a spectrum of cancers.

Challenges in ASXL research and therapeutic development

Research into the ASXL gene family, crucial for both developmental disorders and malignancies, faces significant challenges due to their broad biological functions and roles in epigenetic regulation. The redundancy and potential compensatory mechanisms within the ASXL family add layers of complexity to therapeutic development, often making it difficult to pinpoint the specific contributions of ASXL proteins to disease pathogenesis. The presence of three ASXL family members suggests overlapping functions that may compensate for the dysfunction of one gene, complicating the targeting strategies. Preliminary data from our studies suggests that ASXL1 regulates gene expression of ASXL3 in neural progenitor tissues, but not

in stem cells or haematopoietic cells, suggesting that one *ASXL* gene may regulate the expression of another, but only in certain tissues and developmental stages.

The interaction of *ASXL* proteins with multiple co-factors and other epigenetic modifiers, such as the deubiquitinating enzyme BAP1 and the PRC2 complex, also make it challenging to dissect the specific contributions of *ASXL* proteins to gene expression regulation. This complexity not only offers potential therapeutic targets but also poses significant challenges in developing specific inhibitors or modulators.

Moreover, the challenges of modeling diseases associated with *ASXL* mutations reflect the complexity of these conditions, which span developmental disorders to various cancers. Traditional models like CRISPR-edited cell lines or knockout mice may not fully capture the tissue-specific and development-specific dysregulations caused by *ASXL* mutations, leading to discrepancies within laboratory findings and between research and clinical realities. For example, the study of germline mutations and their implications across different tissues demands patient-derived samples, given that *ASXL* genes have such important roles in development. Furthermore, recent literature suggests that metabolic rewiring may occur and thus lead to differences between a patient's cell with the mutation and a wild type cell with a mutation induced.

Adding to these challenges, a critical gap in the research toolkit is the lack of specific antibodies for *ASXL* proteins. This gap hinders the ability to precisely identify, quantify, and understand the interactions and localization of these proteins, impacting both basic research and the development of targeted therapies. The IDRs of *ASXL* proteins further complicate conventional approaches to studying their structure and function, restricting the use of common techniques like N-terminus tagging, a common technique for tracking protein location and interactions, which is a key tool for tracking protein interactions.

There are also many scientific and clinical hurdles in studying and identifying targeted therapies against *ASXL* gene mutations. One of the primary challenges lies in achieving high

specificity in targeting these genes. *ASXL1*, for example, plays a significant role in myeloid malignancies and is a notoriously difficult target to modulate without affecting non-pathogenic pathways. This difficulty is compounded by *ASXL1*'s role in the epigenetic regulation of gene expression ⁷², where precision targeting is crucial to avoid unintended effects on healthy cells. Therapeutic strategies that inhibit *ASXL1*'s interaction with PRC2, intended to counteract its oncogenic effects in AML, may have off-target effects compromising the essential epigenetic regulation in healthy, non-cancerous cells.

Additionally, the heterogeneity observed within populations with germline *ASXL* mutations highlights the variable impacts of *ASXL* mutations on gene function and disease progression, demanding a nuanced and stratified approach to treatment ⁸⁹. The presence of nonsense mutations in *ASXL* genes in phenotypically normal individuals poses additional questions. Ropers and Wienker (2015) identified 56 predicted loss-of-function (LOF) variants in *ASXL1* within The Exome Aggregation Consortium (ExAC) database ¹¹⁷, which includes exomes from 60,706 unrelated individuals selected to exclude severe genetic disorders such as Bohring-Opitz syndrome (BOS). Similarly, nine putative LOF mutations were identified in *ASXL3* in this cohort ¹¹⁷.

In the updated Genome Aggregation Database (gnomAD) v4.1.0, spanning 730,947 exome sequences and 76,215 whole genome sequences from unrelated individuals, we identified 64 predicted LOF mutations in *ASXL1*, 18 in *ASXL2*, and 24 in *ASXL3* ¹¹⁸. These findings provoke several hypotheses, including that LOF variants might result in subtle manifestations that escape clinical detection, or that disease modifiers could attenuate clinical phenotypes. Additionally, effects of alternative splicing and escape from RNA degradation processes might allow some variants to exert a dominant negative effect. These observations suggest that other factors may influence the penetrance and expressivity of *ASXL* mutations.

Key gaps in research

The research on ASXL genes is still evolving, with significant gaps that need to be addressed to better understand their functions and implications in diseases. Advancements in genomic sequencing technologies and model systems are critical for overcoming these challenges. An integrated approach that combines genetic, biochemical, and clinical insights is crucial for developing effective therapies that can improve patient care and outcomes.

Rare genetic conditions such as BOS, SHAPNS, BRPS, often lack targeted treatments. These disorders, while individually rare, collectively represent a substantial challenge in pediatric care, with significant financial implications ¹¹⁹. Addressing these research gaps not only promises to improve the quality of life for individuals with these conditions but also provides a deeper understanding of complex disease mechanisms that could inform broader therapeutic strategies.

The ASXL gene family has profound implications across both developmental biology and oncology. This review has underscored the dual functionality of ASXL genes, highlighting their critical roles in normal development and their contributions to the pathogenesis of a range of disorders, from neurodevelopmental syndromes to diverse malignancies. The intricate molecular mechanisms, involving both germline and somatic mutations, highlight the complex nature of these genes and their protein products in cellular regulation.

Understanding the multifaceted roles of ASXL genes opens avenues for innovative therapeutic approaches. However, the path towards effective treatments is still fraught with challenges, primarily driven by the complexities of ASXL protein interactions and gene regulatory networks. Overcoming these challenges requires a nuanced understanding of ASXL function across different tissues and developmental stages.

Project goals and methodologies

This thesis aims to elucidate the complex interplay between genetics, epigenetics, and metabolism in BOS and related disorders, leveraging multi-omics and metabolomic approaches to uncover underlying disease mechanisms and identify potential therapeutic targets. The specific studies included are designed to bridge significant gaps in our current understanding of BOS, and *ASXL* genes, and establish a framework for studying other genetic disorders.

The primary objective of this research project is to elucidate the functional impacts of *ASXL1* mutations, explore how these genetic alterations affect cellular and molecular pathways with a particular focus on developmental and growth regulation. In **Chapter 2**, we assayed patient-derived fibroblasts and blood samples from BOS patients using a comprehensive, integrative approach towards epigenetic, transcriptomic, and proteomic data. Our analyses uncovered significant dysregulations within both canonical and non-canonical Wnt signaling pathways.

A secondary goal is to characterize the multi-organ dysregulation caused by *ASXL1* mutations, aiming to understand the systemic effects across different tissues and organ systems. In **Chapter 3**, we extended our studies to BOS patient-derived induced pluripotent stem cells (iPSCs) and neural progenitor cells (NPCs), and utilized a 2D cell model developed by Dr. Jianping Fu's lab to model neural crest cells from BOS iPSCs. We focused specifically on early neural derivatives given the neurodevelopmental defects that are characteristic of BOS. This approach identified significant induction defects in neural crest cells, providing a scalable method to study neurodevelopmental impairments and craniofacial anomalies in BOS.

Another crucial aim of this project is to identify therapeutic targets through comprehensive multi-omics and metabolomics studies. In **Chapter 4**, we explored the metabolic landscape of BOS across patient-derived fibroblasts, iPSCs, and NPCs. We revealed alterations in mitochondrial respiration, amino acid dependency, and stress signaling pathways that are dependent upon the state of differentiation. Our findings suggested that these metabolic

changes are driven by MPC (mitochondrial pyruvate carrier) deficiency, a direct consequence of truncating *ASXL1* mutations. These represent potential metabolic vulnerabilities that could be therapeutically targeted.

Our research extended to examining the broader spectrum of phenotypes associated with different types of *ASXL1* mutations, including missense mutations. We explored how these mutations influence a wide range of systemic effects beyond the classical presentations seen in BOS. In **Chapter 5**, we presented a case of a patient with an overgrowth syndrome, macrocephaly, and normal neurodevelopment. This patient harbored a paternally inherited *ASXL1* missense mutation in the PHD, and had a distinct phenotypic, transcriptomic, and metabolomic phenotype from BOS patients.

Additionally, this thesis seeks to compare the disease pathways by investigating the shared and distinct molecular features of diseases linked to *ASXL1* mutations, such as BOS and acute myeloid leukemia (AML). In **Chapter 6**, we leveraged RNA sequencing and DNA methylation studies to compare the epigenetic and transcriptomic alterations in BOS and AML due to *ASXL1* mutations. This comparison identified specific molecular pathways shared between these conditions and potential therapeutic targets.

In **Chapter 7**, we examined cancer predisposition in genetic disorders, specifically the role of somatic mutations in clonal hematopoiesis of indeterminate potential (CHIP) genes among Bloom Syndrome patients and *BLM* variant carriers. Bloom Syndrome, caused by mutations in the *BLM* gene, is characterized by defects in DNA repair mechanisms that lead to genomic instability and a heightened cancer risk^{90,91}. We identified clonal hematopoiesis as a potential mechanism for increased cancer risk in these both probands and carriers.

Through integrating genetic and omics findings with metabolic studies, we highlight specific molecular disruptions and biomarkers driven by the *ASXL1* genes, and other genes, and their implications for disease pathogenesis. The overarching goal is to identify therapeutic

targets for patients with rare diseases such as BOS, and to lay a framework with which to examine other genetic diseases.

Figures and Tables

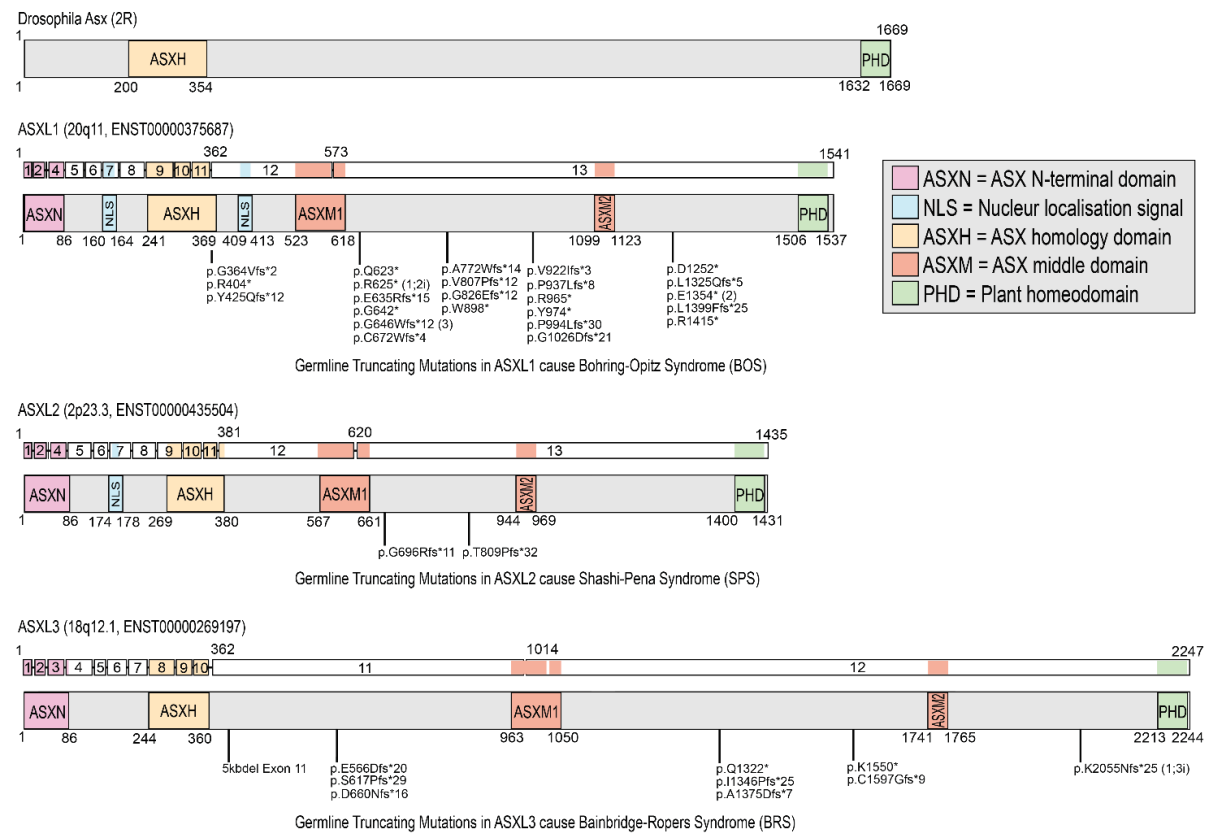


Figure 1-1: ASXL family gene structures and mutations associated with germline neurodevelopmental disorders

Scaled visualizations of the exon and protein structures of *Drosophila Asx* gene and mammalian ASXL gene family. This figure highlights the domain structures and binding partners of the proteins, and the truncating mutations that cause the germline neurodevelopmental disorders Bohring-Opitz syndrome (BOS), Shashi-Pena syndrome (SHAPNS), and Bainbridge-Ropers syndrome (BRPS).

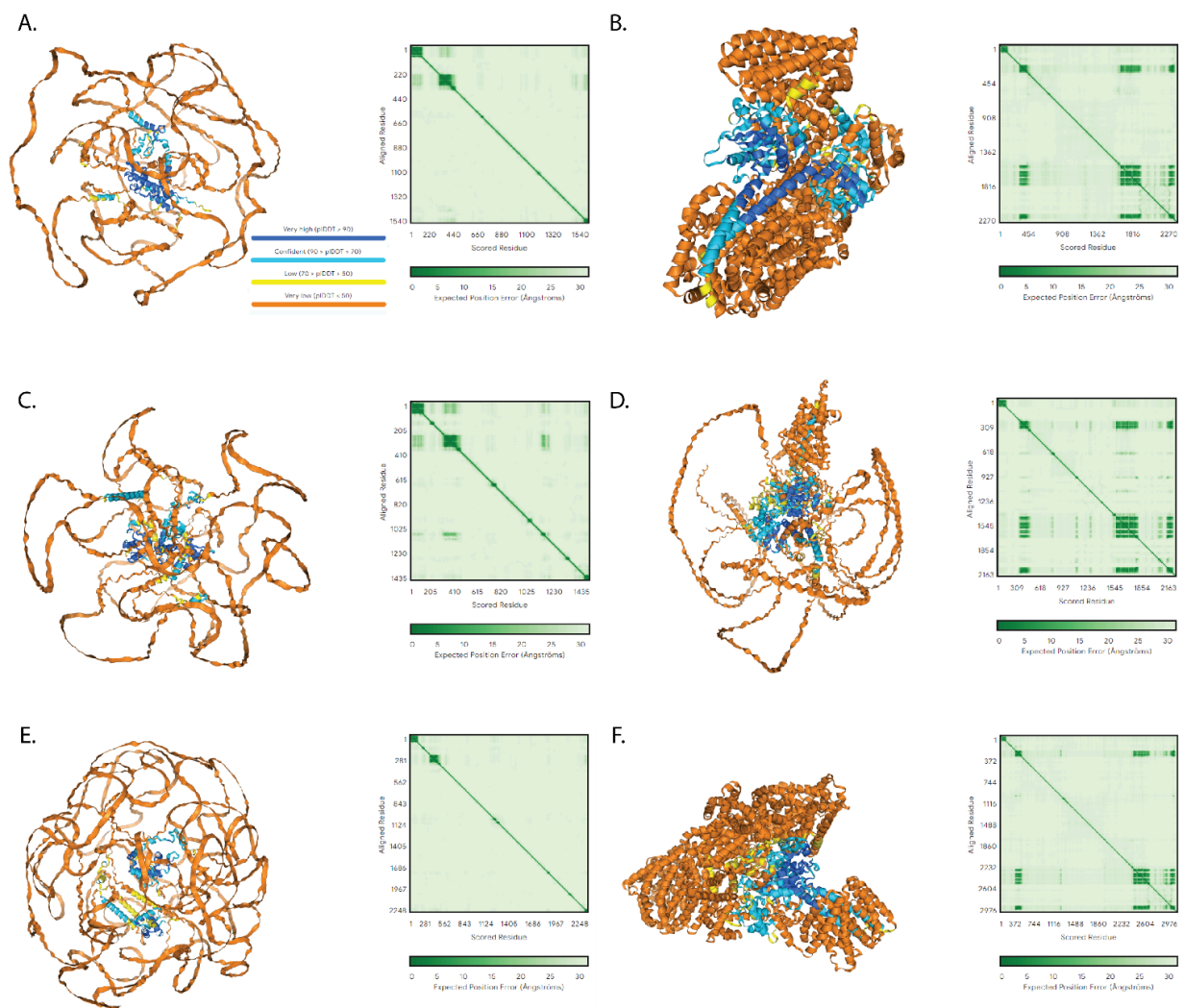


Figure 1-2: Structural models and error predictions for ASXL proteins with and without BAP1 binding

AlphaFold3-predicted three-dimensional structure with predicted local-distance difference test (pLDDT) scores, and predicted aligned error plots for:

(A) ASXL1, **(B)** ASXL1 and BAP1, **(C)** ASXL2, **(D)** ASXL2 and BAP1, **(E)** ASXL3, and **(F)** ASXL3 and BAP1.

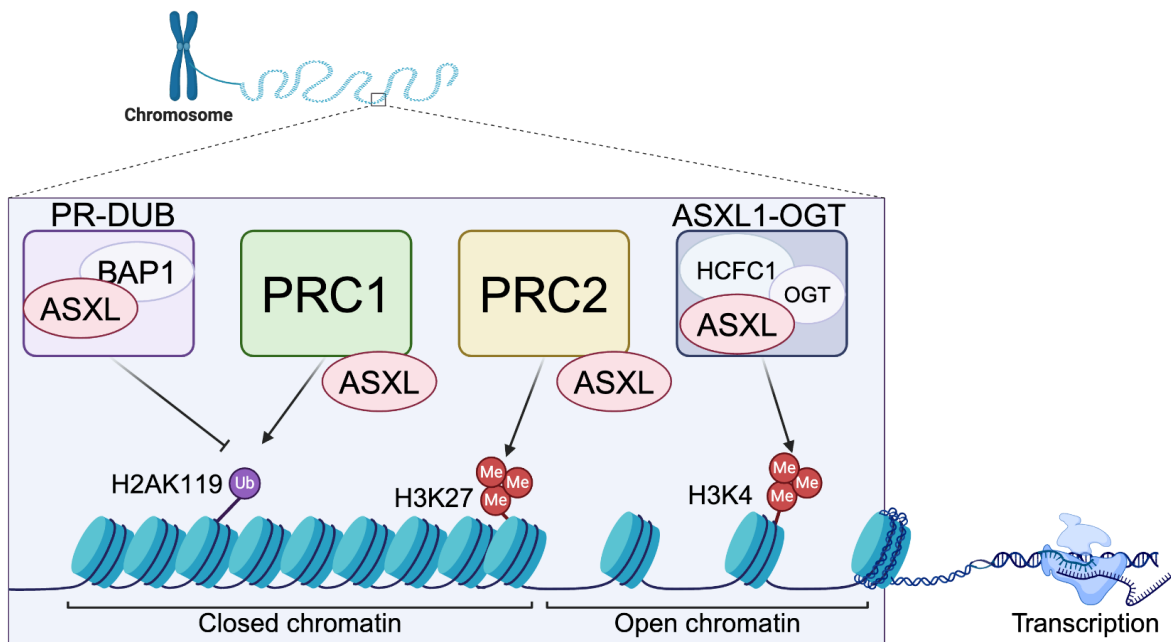


Figure 1-3: Interaction network of ASXL proteins with binding partners and subsequent chromatin modifications

The binding partners of the ASXL gene family, specifically the complexes PR-DUB, PRC1, PRC2, and ASXL1-OGT. Key interactions are shown, along with the specific histone modifications mediated by each complex. For PR-DUB, the histone deubiquitination activity and its effects on chromatin condensation are depicted. The diagram also highlights how PRC1 and PRC2 contribute to histone ubiquitination and methylation, respectively, impacting gene expression through chromatin remodeling. The role of ASXL1-OGT in catalyzing O-GlcNAcylation of histones and its implications for chromatin structure are also presented. Arrows indicate the direction of interaction and the flow of effects leading to chromatin remodeling, influencing gene expression patterns crucial for cellular functions and identity.

Gene/s	Key Findings	Cell/Animal Model
Development		
ASXL1/2/3	Cardiac Development - Loss of Asxl1 leads to ventricular septal defects. ⁴⁷ - Loss of Asxl2 results in congenital heart malformations, including thickened compact myocardium, AV valve stenosis, and patent ductus arteriosus. ^{47,65} - ASXL3 mutations inhibited the proliferation of cardiomyocytes and accelerated cell apoptosis ⁶², impacting cardiac structure and fibrosis in mice ⁶¹	Asxl1(-/-) mice ⁴⁷ Asxl2(-/-) mice ^{47,65} Asxl3 compound heterozygous missense mutation in HL-1 mouse cardiomyocytes ^{61,62}
ASXL1/2	Metabolic and Adipogenic Regulation - ASXL1 represses, while ASXL2 increases the expression of adipogenic genes, most of which are PPARγ targets. ^{54,55} - ASXL2 regulates glucose and lipid homeostasis via interaction with PPARγ. ⁵⁵ - ASXL2 deletion is unresponsive to high-fat diet ^{55,118}	Asxl2(-/-) mice ⁵⁵ Asxl2 deletion in myeloid cells (Asxl2ΔLysM) mice ¹¹⁸ 3T3-L1 preadipocyte cells ⁵⁴
ASXL1/3	Neural Development - ASXL1 and ASXL3 are crucial for proper neural development, including neural folding and neural crest emigration ^{46,56}. - Truncating ASXL1 mutations reduce activation of transcription factors ZIC1 and ZIC4, leading to neural crest emigration defects. ⁴⁶ - Loss of ASXL3 perturbs neural crest development ^{56,57}. - Inhibition of ASXL3 differentially modifies Noggin neural inducing activity, impacting development ⁵⁶.	ASXL1(+/-tm) and ASXL1(tm/tm) iPSCs ⁴⁶ Xenopus embryos injected with ASXL3-MO ⁵⁶ ASXL3 compound het. mutant mice ⁵⁷
ASXL1/2	Hematopoiesis and Myeloid Neoplasms Impaired hematopoiesis and accelerated myeloid neoplasms are observed in various mouse and zebrafish models with knockout, heterozygous loss, or truncating mutations of ASXL1 ^{39,40,42,44,50,51,53,68,71,119,120},	Asxl1(-/-) mice ⁴² Asxl1(+/-) mice ⁴² Asxl1(+/-tm) mice ^{39,40,50,51,71,119} Asxl1fl/fl Mx1 Cre or Vav Cre conditional knockout

	and ASXL2 ^{35,48,49}. - Myeloid neoplasms driven by ASXL1 mutations can be prevented by loss of Bap1 ⁴¹. - Depletion of ASXL1 impairs erythroid differentiation and maturation ¹²¹.	mice ³⁸ <i>Bap1</i>Δ/+;<i>Asx1</i>(+/tm) mice ⁴¹ <i>Asx1</i>(tm/tm) zebrafish ¹²⁰ ASXL1-tm overexpression in HEK293T, 32Dcl3 cells ⁷⁵ and LK, EML cells ⁵³ <i>Asx12</i>(-/-) mice ^{48,49} <i>Asx12</i>+/fl Mx1 Cre and <i>Asx12</i>fl/fl Mx1 Cre ³⁵ Murine c-Kit+ bone marrow cells ⁶⁸
ASXL1	Regulation of Posterior HOXA genes - Knockdown, heterozygous mutant KI mouse, and transient overexpression of mutant <i>Asx1</i> result in derepression and upregulation of posterior HOXA genes via loss of PRC2 recruitment to the HOXA locus ^{38-41,75}. - ASXL1 expression in ASXL1-null cells suppresses <i>HoxA</i> gene expression ³⁸. - ASXL1 mutation alters H2AK119ub at promoters of <i>Hoxa</i> genes, which can be partially restored by reducing Bap1 ⁴¹.	Cancer cell lines ³⁸ <i>Asx1</i>(+/tm) mice ³⁹⁻⁴¹ ASXL1-tm overexpression in HEK293T, 32Dcl3 cells ⁷⁵
ASXL1	Embryonic Development ASXL1 knockout results in high embryonic lethality (80% ⁴²; 100% ⁴³)	<i>Asx1</i>(-/-) mice ^{42,43}
ASXL1	Body Weight ASXL1 knockout results in low body weights in homozygous C-terminus truncated zebrafish ¹²⁰.	<i>Asx1</i>(tm/tm) zebrafish ¹²⁰
ASXL1/2	Skeletal Development - Loss of ASXL1 leads to dramatic bone loss and impaired BMSC self-renewal and differentiation ⁶³ - Loss of <i>Asx12</i> in leads to osteopetrosis in mice ⁵⁵.	<i>Asx1</i>(-/-) BMSCs ⁶³ <i>Asx12</i>(-/-) mice ⁵⁵
Binding Partners, Complexes, and Histone Modifications		
ASXL1/2	Interaction with WTIP (Wilms tumor 1-interacting protein) WTIP is a binding partner of ASXL1 ¹²²	<i>Asx1</i>(-/-) mice ¹²² HeLa cells ⁷²

	- ASXL1 deletion upregulated WTIP target genes ¹²² - The C-terminus of ASXL2 interacts with WTIP. ⁷²	
ASXL1/2/3	PR-DUB mediated deubiquitination of H2AK119ub - ASXL1, ASXL2, and ASXL3 can all bind to BAP1 ^{38,53,64,67,123} via the ASXH domain ⁷⁶ to form the Polycomb repressive deubiquitinase (PR-DUB) complex - Truncated ASXL1 also binds BAP1 and form a stable, hyperactive, complex ^{68,71} - PR-DUB catalyzes the removal of monoubiquitination on lysine 119 of histone H2A - Monoubiquitination of wild-type ASXL2 is required for activation of BAP1 catalytic activity ²⁹	<i>Bap1</i>(-/-), <i>Asxl1</i>(-/-), <i>Asxl2</i>(-/-), <i>Asxl3</i>(-/-), and <i>Asxl1/2/3</i>(-/-) mESCs ⁶⁴ FLAG-tagged BAP1 mice ¹²³ <i>Asxl1</i>(+/tm) mice ⁷¹ BRPS patient (<i>ASXL3</i>(+/tm)) derived fibroblasts ⁶⁷ <i>ASXL1</i>-tm overexpression in LK, EML cells ⁵³ Cancer cell lines ³⁸ HEK293T and murine myeloid 32Dcl3 cells ⁶⁸ HeLa and HEK293T cells ²⁹ Proteomics ⁷⁶
ASXL1/2/3	Polycomb Repressive Complex 1 mediated H2AK119ub - Wild type ASXL1 interacts with PRC1 ⁴⁰ - Truncated ASXL1 cannot interact with PRC1, leading to loss of PRC1-mediated H2AK119ub and subsequent upregulation of PRC1 target genes. This can be attenuated by reducing BAP1 levels ^{40,124}. - Loss of <i>Asxl2</i> led to significant reduction of global H2AK119ub ⁴⁸. - Truncating ASXL3 mutation leads to increased H2AK119Ub1 ⁶⁷	<i>Asxl1</i>(+/tm) mice ^{40,51} <i>ASXL1</i>-tm overexpression in LK, EML cells ⁵³ <i>Asxl1</i>(tm/tm) zebrafish ¹²⁰ HSPCs overexpressing truncated <i>ASXL1</i>(tm) ¹²⁴ <i>Asxl2</i>(-/-) mice ⁴⁸ BRPS patient (<i>ASXL3</i>(+/tm)) derived fibroblasts ⁶⁷
ASXL1/2	Polycomb Repressive Complex 2 mediated H3K27me3 - Loss of ASXL1 in various models leads to a reduction of PRC2-mediated H3K27me3 repression . ^{38,42,53,121} - Loss of <i>Asxl2</i> led to significant reduction of global H3K27me3, but at lesser extent than <i>Asxl1</i> ^{35,48,125} <i>ASXL1</i>-null cells show a global increase in H3K27me3 expression, while heterozygous truncating mutations lead to no global change but an altered	<i>Asxl1</i>(-/-) mice ⁴² <i>Asxl1</i>fl/fl Mx1 Cre or Vav Cre ⁴⁴ <i>Asxl1</i>(+/tm) mice ^{51,52,126} <i>Asxl1</i>(-/-) MEFs ⁴³ <i>Asxl2</i>(-/-) mice ¹²⁵ <i>Asxl2</i>(-/-) murine bone marrow ⁴⁸ <i>ASXL1</i>-tm overexpression in LK, EML cells ⁵³ <i>ASXL1</i> knockdown in cord blood (CB) CD34+ cells ¹²¹

	H3K27me3 profile in murine bone marrow^{51,52} ASXL1 interacts with PRC2 complex components such as EZH2^{38,43}	ASXL1 and ASXL2 shRNAs in SKNO-1 cells³⁵ Cancer cell lines³⁸
ASXL1	ASXL1-OGT Complex mediated H3K4me3 - ASXL1/OGT complex is required for H3K4me3⁶⁹ and knockout of ASXL1 or the presence of heterozygous truncating mutations leads to loss of H3K4me3^{42,51,69,71,121} - ASXL1 interacts with OGT and HCFC1⁶⁹ - OGT directly stabilises ASXL⁶⁹	Asxl1(-/-) mice⁴² Asxl1(+/-) mice^{51,71} ASXL1 knockdown in cord blood (CB) CD34+ cells¹²¹ HEK293T cells⁶⁹
ASXL1	RNA Polymerase II - ASXL1 interacts with RNA polymerase II¹²⁷ - ASXL1 deletion results in global increase in transcription¹²⁸	Asxl1(-/-) murine BMSCs^{127,128}
ASXL1	Bromodomain-containing 4 (BRD4) ASXL1 interacts with Bromodomain-containing 4 (BRD4), a myeloid malignancy oncoprotein⁵⁰	Asxl1(+/-) mice⁵⁰
ASXL1/2	EZH2 (PRC2 complex component) - ASXL1 cooperates with EZH2 to repress p16Ink4a, a mediator of cellular senescence. - Mutant ASXL1 disrupts EZH2 bindings and leads to elevated p16Ink4a expression, resulting in cellular senescence⁴³ - Asxl2 co-localises with Ezh2 to directly repress β-MHC¹²⁵	Asxl1(-/-) MEFs⁴³ Asxl2(-/-) mice¹²⁵
ASXL1	Paraspeckle Formation - Paraspeckles are crucial for hematopoietic stem and progenitor cell repopulating potential¹²⁹. - ASXL1 forms phase-separated organelles known as paraspeckles, while mutant ASXL1 cannot¹²⁹.	Asxl1(+/-) murine BMSCs¹²⁹ HEK293T and HL-60 cells¹²⁹
Signaling Pathways		
ASXL1	Wnt Pathway Truncating ASXL1 mutations lead to aberrant activation of the canonical Wnt	BOS patient (ASXL1(+/-)) derived blood and fibroblasts⁷⁰

	and non-canonical planar cell polarity pathway ⁷⁰ .	
ASXL1/3	mTOR Pathway - ASXL1 mutation leads to overactive Akt/mTOR signaling and aberrant proliferation ⁷¹. - Wild type ASXL1 interacts with phosphorylated AKT and affects the AKT-E2F axis leading to cellular senescence ⁴³. - ASXL3 knockdown leads to dysregulation of genes enriched in PI3K-AKT processes ⁷³	Asxl1(-/-) MEFs ⁴³ Asxl1(+/-tm) mice ⁷¹ ASXL3 knockdown in P19 cells ⁷³
ASXL2	Retinoic-Acid Induced Transcription C-terminus of ASXL2 enhances retinoic-acid induced transcription while full length ASXL2 does not ⁷²	HeLa cells ⁷²
ASXL3	Ras/ERK signaling pathway ASXL3 interacts with the Ras/ERK signaling pathway through the modulation of lncRNAs and FGFR2 transcripts, leading to altered cell proliferation and apoptosis ⁶².	Asxl3 compound heterozygous missense mutation in HL-1 mouse cardiomyocytes ⁶²

Table 1-1: Overview of experimental studies on ASXL gene family: functions, interactions, and regulatory mechanisms

Comprehensive overview of experimental studies on the ASXL gene family. These studies detail the physiological and pathophysiological functions of the ASXL gene family across development, interactions with various binding partners and complexes, effects on histone modifications, and regulation of signaling pathways.

ASXL1 NM_015338	ASXL1 NP_056153	# De novo Patients* Except *i = inherited	Reference
N/A	Del of Exons 5 - 13	1	Previously unpublished
c.1049G>A	p.Trp350*	1	130
c.1091delG	p.Gly364Valfs*2	1	70
c.1210C>T	p.Arg404*	5	87,89,131–133
c.1272_1273delGT	p.Tyr425Glnfs*12	1	89,101
c.1406dupC	p.Ser470Leufs*15	1	89
c.1446dupA	p.Val483Serfs*2	1	89
c.1517_1518delGA	p.Arg506Asnfs*3	1	93
c.1544_1545delTG	p.Val515Glyfs*13	1	89
c.1719+5G>C	N/A	1	89
c.1720-2A>G	p.Ile574Valfs*22	1	134
c.1867C>T	p.Gln623*	2	89,93
c.1873C>T	p.Arg625*	1, 2 maternally inherited	Previously unpublished
c.1910_1922del	p.Glu635Argfs*15	1	89
c.1924G>T	p.Gly642*	1	89,101
c.1927_1928insG	p.Gly644Trpfs*9	1	89
c.1934dup	p.Gly646Trpfs*12	3	89,135,136
c.1981A>T	p.Arg661*	1	89
c.2013_2014del	p.Cys672Trpfs*4	1	89,101
c.2024_2025del	p.Pro675Argfs*42	1	89
c.2033dupG	p.Gly680Argfs*38	1	137
c.2036dup	p.Gly680Argfs*38	2	89,138
c.2100dupT	p.Pro701Serfs*16	1	139
c.2197C>T	p.Gln733*	1	87
c.2324T>G	p.Leu775*	2	89,101
c.2332C>T	p.Gln778*	2	87,89
c.2407_2411del	p.Gln803Thrfs*17	1	140
c.2414_2415delinsA	p.Pro805Glnfs*13	1	89
c.2416_2417dupAC	p.Val807Profs*12	1	89

c.2468T>G	p.Leu823*	1	87
c.2477delG	p.Gly826GluFs*12	1	70
c.2535dupC	p.Ser846GlnFs*5	1	87
c.2689delC	p.His897IleFs*11	1	130
c.2694G>A	p.Trp898*	1	Previously unpublished
c.2759_2762dupCATC	p.Val922LeuFs*3	1	89,101
c.2773C>T	p.Gln925*	1	87
c.2810delC	p.Pro937LeuFs*8	1	70
c.2893C>T	p.Arg965*	2, 3 inherited from mosaic parent	70,93,140,141
c.2922C>A	p.Tyr974*	1	89
c.2966_2967delCT	p.Ser989*	1	89
c.2981delC	p.Pro994LeuFs*30	1	89
c.3077delG	p.Gly1026AspFs*21	1	89,101
c.3083C>A	p.Ser1028*	1	87
c.3596_3606del	p.Ala1199GluFs*9	1	89
c.3601delC	p.Gln1201LysFs*16	1	89
c.3754_3758del	p.Asp1252*	1	89
c.3762delT	p.Ser1255GlnFs*25	1	142
c.3894_3897dupAGGC	p.Gln1300ArgFs*30	1	89
c.3974_3977delTTCCins AGCAGGA	p.Leu1325GlnFs*5	1	Previously unpublished
c.4060G>T	p.Glu1354*	3	70,93
c.4116_4117delTT	p.Phe1373CysFs*6	1	143
c.4127dup	p.Pro1377SerFs*3	1	89
c.4196dupT	p.Leu1399PheFs*25	2	144, Previously unpublished
c.4243C>T	p.Arg1415*	2	89
ASXL2 NM_018263.6	ASXL2 NP_0607433.4	# De novo Patients* Except *i = inherited	Reference
c.1288G>T	p.Glu430*	1	81
c.1225_1228delCCAA	p.Pro409AsnFs*13	1	81

c.1792C>T	p.Gln598*	1	60
c.2081dupG	p.Gly696Argfs*11	2	81,145
c.2142_2152del	p.Ser714Argfs*5	1	146
c.2424delC	p.Thr809Profs*32	2	81, Previously unpublished
c.2468_2477del	p.Gly823Valfs*15	1	147
c.2472delC	p.Ser825Valfs*16	1	81
c.2485C>T	p.Gln829*	1	148
c.2955_2967delins	p.Glu987Argfs*15	1	149
c.2971_2974delGGAG	p.Gly991Argfs*3	1	81
ASXL3 NM_030632.2	ASXL3 NP_085135.1	# De novo Patients* Except *i = inherited	Reference
N/A	5kb del of Exon 11	1	145
c.55_879del	p.Ala19_Gln293del	1	90
c.187C>T	p.Arg63*	1	90
c.1063G>T	p.Glu355*	1	150
c.1074T>A	p.Tyr358*	1	151
c.1082dupA	p.Leu362Alafs*23	1	151
c.1095_1096delAA	p.Glu367Glyfs*17	1	90
c.1201delG	p.Ala401Glnfs*8	1	151
c.1210C>T	p.Gln404*	2	79,152
c.1217dupA	p.Ser407Lysfs*25	1	153
c.1219delA	p.Ser407Alafs*2	1	80
c.1314_1316delinsA	p.Ser439Argfs*7	1	154
c.1318dupG	p.Glu440Glyfs*7	1	155
c.1369G>T	p.Glu457*	1	80
c.1377_1378delGA	p.Glu459Aspfs*2	1	78
c.1396C>T	p.Gln466*	1	79
c.1406_1409dup	p.His470Glnfs*3	1	90
c.1421_1422ins	p.Leu483*	1	156
c.1422dupT	p.Glu475fs*1	1	79
c.1448dupT	p.Thr484Asnfs*5	2	67,157

c.1451_1454delCCAA	p.Thr484Ilefs*8	1	90
c.1484insTGAA	p.Asp497*	1	151
c.1491dupT	p.Asn498*	1	151
c.1500delA	p.Glu500Aspfs*5	1	90
c.1505_1508dupGTGT	p.Met504Cysfs*5	1	90
c.1534_1535delTT	p.Leu512Asnfs*4	2	90
c.1535T>A	p.Leu512*	1	90
c.1579dupC	p.Gln527Profs*15	1	90
c.1579delC	p.Gln527Argfs*4	1	90
c.1606C>T	p.Gln536*	1	90
c.1632_1637delins31	p.Pro545Leufs*10	1	92
c.1698_1699delGA	p.Glu566Aspfs*20	3	90,144, Previously unpublished
c.1783C>T	p.Gln595*	1	151 90
c.1795G>T	p.Glu599*	1	158
c.1849_1850delAG	p.Ser617Profs*29	1	Previously unpublished
c.1884_1885delAG	p.Gly629Argfs*17	1	90
c.1895dupC	p.Gln633Thrfs*14	1	93
c.1897_1898delCA	p.Gln633Valfs*13	2	145,159
c.1921_1924dupACTC	p.Pro642Hisfs*6	1	90
c.1978_1981delACAG	p.Asp660Asnfs*16	2	79, Previously unpublished
c.1990C>T	p.Gln664*	2	90,93
c.2791_2792delCA	p.Gln931Glufs*19	1, 1 paternally inherited	90,160
c.2992_2995delGAGA	p.Glu998Lysfs*26	1	161
c.3006delT	p.Arg1004Glufs*21	1	162
c.3028delC	p.Pro1010Leufs*14	1	163
c.3039+1G>A	N/A	2	164,165
c.3043C>T	p.Gln1015*	2	166,167
c.3106C>T	p.Arg1036*	4, 1 unknown, 2 likely inherited	80,90,91,164,168,169
c.3127_3128dupGT	p.Gly1045Valfs*99	1	151
c.3178dupC	p.Arg1060Profs*50	2	90,151

c.3284_3288del	p.Thr1096Asnfs*12	1	92
c.3307A>T	p.Lys1103*	1	78
c.3313_3316delCAGA	p.Thr1106Argfs*36	1	164
c.3349C>T	p.Arg1117*	2	93,170
c.3355dup	p.His1119Profs*7	1	151
c.3364C>T	p.Gln1122*	1	67
c.3382delC	p.Arg1128Glyfs*15	1	90
c.3419delC	p.Pro1140Glnfs*3	1	90
c.3464C>A	p.Ser1155*	2	90,171
c.3494_3495delGT	p.Cys1165*	2	80,172
c.3613G>T	p.Glu1205*	1	80
c.3635T>G	p.Leu1212*	1	151
c.3811_3814dupAACA	p.Thr1272Lysfs*8	1	90
c.3964C>T	p.Gln1322*	2	144, Previously unpublished
c.4034_4035dupCC	p.Ile1346Profs*15	2	144, Previously unpublished
c.4060_4061delTC	p.Ser1354Hisfs*2	1	144
c.4072_4073delGT	p.Val1358Leufs*8	1	80
c.4090G>T	p.Gly1364*	1	173
c.4120_4123dupATAG	p.Ala1375Aspfs*7	2	90, Previously unpublished
c.4144C>T	p.Gln1382*	1	151
c.4219_4220delCT	p.Leu1407Glyfs*20	1	90
c.4322C>G	p.Ser1441*	1	93
c.4330C>T	p.Arg1444*	3	67,151,174
c.4336_4337delAG	p.Arg1446glyfs*3	1	90
c.4360C>T	p.Gln1454*	2	90
c.4399C>T	p.Arg1467*	3, 1 unknown	78,90
c.4441dupC	p.Leu1481Profs*12	3, 1 maternally inherited	90,160
c.4462_4465del	p.Thr1488Serfs*17	1	90
c.4479_4483del	p.Ser1493Argfs*29	1	90
c.4509_4513dup	p.Val1505Aspfs*3	1	92
c.4534C>T	p.Gln1512*	1, 2 maternally inherited	90,160

c.4648A>T	p.Lys1550*	2	144, Previously unpublished
c.4788_4816delinsT	p.Cys1597Glyfs*9	2	144, Previously unpublished
c.4871_4874delA	p.His1624Argfs*10	1	90
c.4894_4895delGA	p.Glu1632Ilefs*13	1	90
c.4906C>T	p.Gln1636*	1	145
c.5455C>T	p.Gln1819*	1	78
c.5659A>T	p.Arg1887*	1	90
c.6165delA	p.Lys2055Asnfs*25	1, 3 maternally inherited	Previously unpublished
c.6697_6710dup	p.Ser2238Thrfs*3	1	88

Table 1-2: Clinical reports of mutations in *ASXL1*, *ASXL2*, and *ASXL3* genes associated with BOS, SHAPNS, and BRPS

Clinical reports of mutations in *ASXL1*, *ASXL2*, and *ASXL3* genes associated with Bohring-Opitz syndrome (BOS), Shashi-Pena syndrome (SHAPNS), and Bainbridge-Ropers syndrome (BRPS) as published up to March 2024.

Gene	Germline Truncation Mutation	Somatic Truncation Mutation	Amplification and Overexpression
ASXL1	Bohring-Opitz Syndrome (BOS)	Myeloid malignancies	Cervical cancer
		Colorectal cancer	
		Head and neck squamous cell carcinoma	
		Liver cancer	
		Prostate cancer	
		Breast cancer	
ASXL2	Shashi-Pena Syndrome (SHAPNS)	Myeloid malignancies	Adult T-cell leukemia/lymphoma
		Prostate cancer	Colorectal cancer
		Pancreatic cancer	Pancreatic cancer
		Breast cancer	Breast cancer
ASXL3	Bainbridge-Ropers Syndrome (BRPS)	Melanoma	Small cell lung cancer

Table 1-3: Comparative overview of disorders associated with germline and somatic mutations in *ASXL* genes

This table presents a comprehensive comparison of clinical disorders linked to mutations in *ASXL1*, *ASXL2*, and *ASXL3* genes, distinguishing between germline and somatic mutations.

References

1. Katoh, M. Functional and cancer genomics of ASXL family members. *Br. J. Cancer* **109**, 299–306 (2013).
2. Jürgens, G. A group of genes controlling the spatial expression of the bithorax complex in *Drosophila*. *Nature* **316**, 153–155 (1985).
3. Sinclair, D. A. *et al.* The Additional sex combs gene of *Drosophila* encodes a chromatin protein that binds to shared and unique Polycomb group sites on polytene chromosomes. *Development* **125**, 1207–1216 (1998).
4. Sinclair, D. A., Campbell, R. B., Nicholls, F., Slade, E. & Brock, H. W. Genetic analysis of the additional sex combs locus of *Drosophila melanogaster*. *Genetics* **130**, 817–825 (1992).
5. Harding, K. & Levine, M. Gap genes define the limits of antennapedia and bithorax gene expression during early development in *Drosophila*. *EMBO J.* **7**, 205–214 (1988).
6. Irish, V. F., Martinez-Arias, A. & Akam, M. Spatial regulation of the Antennapedia and Ultrabithorax homeotic genes during *Drosophila* early development. *EMBO J.* **8**, 1527–1537 (1989).
7. White, R. A. & Lehmann, R. A gap gene, hunchback, regulates the spatial expression of Ultrabithorax. *Cell* **47**, 311–321 (1986).
8. Wedeen, C., Harding, K. & Levine, M. Spatial regulation of Antennapedia and bithorax gene expression by the Polycomb locus in *Drosophila*. *Cell* **44**, 739–748 (1986).
9. McKeon, J. & Brock, H. W. Interactions of the Polycomb group of genes with homeotic loci of *Drosophila*. *Roux's Arch. Dev. Biol.* **199**, 387–396 (1991).
10. Struhl, G. & Akam, M. Altered distributions of Ultrabithorax transcripts in extra sex combs mutant embryos of *Drosophila*. *EMBO J.* **4**, 3259–3264 (1985).
11. Jones, R. S. & Gelbart, W. M. The *Drosophila* Polycomb-group gene Enhancer of zeste contains a region with sequence similarity to trithorax. *Mol. Cell. Biol.* **13**, 6357–6366

- (1993).
12. Simon, J., Chiang, A. & Bender, W. Ten different Polycomb group genes are required for spatial control of the *abdA* and *AbdB* homeotic products. *Development* **114**, 493–505 (1992).
 13. Gaytán de Ayala Alonso, A. *et al.* A genetic screen identifies novel polycomb group genes in *Drosophila*. *Genetics* **176**, 2099–2108 (2007).
 14. Scheuermann, J. C. *et al.* Histone H2A deubiquitinase activity of the Polycomb repressive complex PR-DUB. *Nature* **465**, 243–247 (2010).
 15. Levine, S. S. *et al.* The core of the polycomb repressive complex is compositionally and functionally conserved in flies and humans. *Mol. Cell. Biol.* **22**, 6070–6078 (2002).
 16. Aranda, S., Mas, G. & Di Croce, L. Regulation of gene transcription by Polycomb proteins. *Sci Adv* **1**, e1500737 (2015).
 17. Wang, L. *et al.* Hierarchical recruitment of polycomb group silencing complexes. *Mol. Cell* **14**, 637–646 (2004).
 18. Cao, R., Tsukada, Y.-I. & Zhang, Y. Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing. *Mol. Cell* **20**, 845–854 (2005).
 19. Margueron, R. & Reinberg, D. The Polycomb complex PRC2 and its mark in life. *Nature* **469**, 343–349 (2011).
 20. Hauri, S. *et al.* A High-Density Map for Navigating the Human Polycomb Complexome. *Cell Rep.* **17**, 583–595 (2016).
 21. Kloet, S. L. *et al.* The dynamic interactome and genomic targets of Polycomb complexes during stem-cell differentiation. *Nat. Struct. Mol. Biol.* **23**, 682–690 (2016).
 22. Milne, T. A., Sinclair, D. A. & Brock, H. W. The Additional sex combs gene of *Drosophila* is required for activation and repression of homeotic loci, and interacts specifically with Polycomb and super sex combs. *Mol. Gen. Genet.* **261**, 753–761 (1999).
 23. Locke, J., Kotarski, M. A. & Tartof, K. D. Dosage-dependent modifiers of position effect

- variegation in *Drosophila* and a mass action model that explains their effect. *Genetics* **120**, 181–198 (1988).
24. Kennison, J. A. The Polycomb and trithorax group proteins of *Drosophila*: trans-regulators of homeotic gene function. *Annu. Rev. Genet.* **29**, 289–303 (1995).
 25. Brock, H. W. & Fisher, C. L. Maintenance of gene expression patterns. *Dev. Dyn.* **232**, 633–655 (2005).
 26. Schuettengruber, B., Bourbon, H.-M., Di Croce, L. & Cavalli, G. Genome Regulation by Polycomb and Trithorax: 70 Years and Counting. *Cell* **171**, 34–57 (2017).
 27. Fisher, C. L. *et al.* Additional sex combs-like 1 belongs to the enhancer of trithorax and polycomb group and genetically interacts with Cbx2 in mice. *Dev. Biol.* **337**, 9–15 (2010).
 28. Fisher, C. L., Berger, J., Randazzo, F. & Brock, H. W. A human homolog of Additional sex combs, ADDITIONAL SEX COMBS-LIKE 1, maps to chromosome 20q11. *Gene* **306**, 115–126 (2003).
 29. Katoh, M. & Katoh, M. Identification and characterization of ASXL3 gene in silico. *Int. J. Oncol.* **24**, 1617–1622 (2004).
 30. Katoh, M. & Katoh, M. Identification and characterization of ASXL2 gene in silico. *Int. J. Oncol.* **23**, 845–850 (2003).
 31. Daou, S. *et al.* The BAP1/ASXL2 histone H2A deubiquitinase complex regulates cell proliferation and is disrupted in cancer. *J. Biol. Chem.* **290**, 28643–28663 (2015).
 32. Aravind, L. & Iyer, L. M. The HARE-HTH and associated domains: novel modules in the coordination of epigenetic DNA and protein modifications. *Cell Cycle* **11**, 119–131 (2012).
 33. Cho, Y.-S., Kim, E.-J., Park, U.-H., Sin, H.-S. & Um, S.-J. Additional sex comb-like 1 (ASXL1), in cooperation with SRC-1, acts as a ligand-dependent coactivator for retinoic acid receptor. *J. Biol. Chem.* **281**, 17588–17598 (2006).
 34. Katoh, M. Functional proteomics of the epigenetic regulators ASXL1, ASXL2 and ASXL3: a convergence of proteomics and epigenetics for translational medicine. *Expert Rev.*

- Proteomics* **12**, 317–328 (2015).
35. Jain, K. *et al.* Characterization of the plant homeodomain (PHD) reader family for their histone tail interactions. *Epigenetics Chromatin* **13**, 3 (2020).
 36. Sanchez, R. & Zhou, M.-M. The PHD finger: a versatile epigenome reader. *Trends Biochem. Sci.* **36**, 364–372 (2011).
 37. Micol, J.-B. *et al.* ASXL2 is essential for haematopoiesis and acts as a haploinsufficient tumour suppressor in leukemia. *Nat. Commun.* **8**, 15429 (2017).
 38. Gildea, J. J., Lopez, R. & Shearn, A. A screen for new trithorax group genes identified little imaginal discs, the *Drosophila melanogaster* homologue of human retinoblastoma binding protein 2. *Genetics* **156**, 645–663 (2000).
 39. Abramson, J. *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* (2024) doi:10.1038/s41586-024-07487-w.
 40. Abdel-Wahab, O. *et al.* ASXL1 mutations promote myeloid transformation through loss of PRC2-mediated gene repression. *Cancer Cell* **22**, 180–193 (2012).
 41. Inoue, D. *et al.* Myelodysplastic syndromes are induced by histone methylation–altering ASXL1 mutations. *J. Clin. Invest.* **123**, 4627–4640 (2013).
 42. Uni, M. *et al.* Modeling ASXL1 mutation revealed impaired hematopoiesis caused by derepression of p16Ink4a through aberrant PRC1-mediated histone modification. *Leukemia* **33**, 191–204 (2019).
 43. Guo, Y. *et al.* Reduced BAP1 activity prevents ASXL1 truncation-driven myeloid malignancy in vivo. *Leukemia* **32**, 1834–1837 (2018).
 44. Wang, J. *et al.* Loss of Asxl1 leads to myelodysplastic syndrome-like disease in mice. *Blood* **123**, 541–553 (2014).
 45. Youn, H. S. *et al.* Asxl1 deficiency in embryonic fibroblasts leads to cellular senescence via impairment of the AKT-E2F pathway and Ezh2 inactivation. *Sci. Rep.* **7**, 5198 (2017).
 46. Abdel-Wahab, O. *et al.* Deletion of Asxl1 results in myelodysplasia and severe

- developmental defects in vivo. *J. Exp. Med.* **210**, 2641–2659 (2013).
47. Yamato, G. *et al.* ASXL2 mutations are frequently found in pediatric AML patients with t(8;21)/RUNX1-RUNX1T1 and associated with a better prognosis. *Genes Chromosomes Cancer* **56**, 382–393 (2017).
 48. Matheus, F. *et al.* Pathological ASXL1 Mutations and Protein Variants Impair Neural Crest Development. *Stem Cell Reports* **12**, 861–868 (2019).
 49. McGinley, A. L., Li, Y., Deliu, Z. & Wang, Q. T. Additional sex combs-like family genes are required for normal cardiovascular development. *Genesis* **52**, 671–686 (2014).
 50. Madan, V. *et al.* ASXL2 regulates hematopoiesis in mice and its deficiency promotes myeloid expansion. *Haematologica* **103**, 1980–1990 (2018).
 51. Li, J. *et al.* Loss of Asxl2 leads to myeloid malignancies in mice. *Nat. Commun.* **8**, 15456 (2017).
 52. Yang, H. *et al.* Gain of function of ASXL1 truncating protein in the pathogenesis of myeloid malignancies. *Blood* **131**, 328–341 (2018).
 53. Nagase, R. *et al.* Expression of mutant Asxl1 perturbs hematopoiesis and promotes susceptibility to leukemic transformation. *J. Exp. Med.* **215**, 1729–1747 (2018).
 54. Hsu, Y.-C. *et al.* The distinct biological implications of Asxl1 mutation and its roles in leukemogenesis revealed by a knock-in mouse model. *J. Hematol. Oncol.* **10**, 139 (2017).
 55. Balasubramani, A. *et al.* Cancer-associated ASXL1 mutations may act as gain-of-function mutations of the ASXL1–BAP1 complex. *Nat. Commun.* **6**, 1–15 (2015).
 56. Park, U.-H., Yoon, S. K., Park, T., Kim, E.-J. & Um, S.-J. Additional sex comb-like (ASXL) proteins 1 and 2 play opposite roles in adipogenesis via reciprocal regulation of peroxisome proliferator-activated receptor {gamma}. *J. Biol. Chem.* **286**, 1354–1363 (2011).
 57. Izawa, T. *et al.* ASXL2 Regulates Glucose, Lipid, and Skeletal Homeostasis. *Cell Rep.* **11**, 1625–1637 (2015).
 58. Lichtig, H. *et al.* Modeling Bainbridge-Ropers Syndrome in *Xenopus laevis* Embryos. *Front.*

Physiol. **11**, 75 (2020).

59. Zhang, S., Fu, F., Zhen, L., Li, R. & Liao, C. Alteration of long non-coding RNAs and mRNAs expression profiles by compound heterozygous ASXL3 mutations in the mouse brain. *Bioengineered* **12**, 6935–6951 (2021).
60. Baskind, H. A. *et al.* Functional conservation of Asxl2, a murine homolog for the Drosophila enhancer of trithorax and polycomb group gene Asx. *PLoS One* **4**, e4750 (2009).
61. Brunner, R. *et al.* Asxl2^{-/-} Mice Exhibit De Novo Cardiomyocyte Production during Adulthood. *J Dev Biol* **4**, (2016).
62. Yuan, M. *et al.* A newborn with a pathogenic variant in ASXL2 expanding the phenotype of SHAPNS: a case report and literature review. *Transl Pediatr* **12**, 86–96 (2023).
63. Fu, F. *et al.* Compound heterozygous mutation of the ASXL3 gene causes autosomal recessive congenital heart disease. *Hum. Genet.* **140**, 333–348 (2021).
64. Liu, Z. *et al.* ASXL3 gene mutations inhibit cell proliferation and promote cell apoptosis in mouse cardiomyocytes by upregulating lncRNA NONMMUT063967.2. *Biochem Biophys Rep* **35**, 101505 (2023).
65. Zhang, P. *et al.* Loss of Asxl1 Alters Self-Renewal and Cell Fate of Bone Marrow Stromal Cell, Leading to Bohring-Opitz-like Syndrome in Mice. *Stem Cell Reports* **6**, 914–925 (2016).
66. Kolovos, P. *et al.* PR-DUB maintains the expression of critical genes through FOXK1/2- and ASXL1/2/3-dependent recruitment to chromatin and H2AK119ub1 deubiquitination. *Genome Res.* **30**, 1119–1130 (2020).
67. Lai, H.-L. & Wang, Q. T. Additional sex combs-like 2 is required for polycomb repressive complex 2 binding at select targets. *PLoS One* **8**, e73983 (2013).
68. Cao, R. *et al.* Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science* **298**, 1039–1043 (2002).
69. Srivastava, A. *et al.* De novo dominant ASXL3 mutations alter H2A deubiquitination and

- transcription in Bainbridge–Ropers syndrome. *Hum. Mol. Genet.* **25**, 597–608 (2015).
70. Asada, S. *et al.* Mutant ASXL1 cooperates with BAP1 to promote myeloid leukaemogenesis. *Nature Communications* vol. 9 Preprint at <https://doi.org/10.1038/s41467-018-05085-9> (2018).
 71. Inoue, D. *et al.* A novel ASXL1-OGT axis plays roles in H3K4 methylation and tumor suppression in myeloid malignancies. *Leukemia* **32**, 1327–1337 (2018).
 72. Lin, I. *et al.* Multi-omics of Bohring-Opitz syndrome truncating ASXL1 mutations identify canonical and non-canonical Wnt signaling dysregulation. *JCI Insight* (2023) doi:10.1172/jci.insight.167744.
 73. Fujino, T. *et al.* Mutant ASXL1 induces age-related expansion of phenotypic hematopoietic stem cells through activation of Akt/mTOR pathway. *Nat. Commun.* **12**, 1826 (2021).
 74. Khan, F. F., Li, Y., Balyan, A. & Wang, Q. T. WTIP interacts with ASXL2 and blocks ASXL2-mediated activation of retinoic acid signaling. *Biochem. Biophys. Res. Commun.* **451**, 101–106 (2014).
 75. Liu, Z.-Q. *et al.* Identification of differential microRNAs and messenger RNAs resulting from ASXL transcriptional regulator 3 knockdown during heart development. *Bioengineered* **13**, 9948–9961 (2022).
 76. Asada, S. *et al.* Leukemogenic Functions of Mutant ASXL1 Are Regulated By CDK-Mediated Phosphorylation. *Blood* **134**, 731 (2019).
 77. Inoue, D., Nishimura, K., Kozuka-Hata, H., Oyama, M. & Kitamura, T. The stability of epigenetic factor ASXL1 is regulated through ubiquitination and USP7-mediated deubiquitination. *Leukemia* **29**, 2257–2260 (2015).
 78. Sahtoe, D. D., van Dijk, W. J., Ekkebus, R., Ovaa, H. & Sixma, T. K. BAP1/ASXL1 recruitment and activation for H2A deubiquitination. *Nat. Commun.* **7**, 10292 (2016).
 79. Xia, Y.-K. *et al.* Tumor-derived neomorphic mutations in ASXL1 impairs the BAP1-ASXL1-FOXK1/K2 transcription network. *Protein Cell* **12**, 557–577 (2021).

80. Yu, K. P.-T., Luk, H.-M., Fung, J. L. F., Chung, B. H.-Y. & Lo, I. F.-M. Further expanding the clinical phenotype in Bainbridge-Ropers syndrome and dissecting genotype-phenotype correlation in the ASXL3 mutational cluster regions. *Eur. J. Med. Genet.* **64**, 104107 (2021).
81. Bainbridge, M. N. *et al.* De novo truncating mutations in ASXL3 are associated with a novel clinical phenotype with similarities to Bohring-Opitz syndrome. *Genome Med.* **5**, 11 (2013).
82. Kuechler, A. *et al.* Bainbridge-Ropers syndrome caused by loss-of-function variants in ASXL3: a recognizable condition. *Eur. J. Hum. Genet.* **25**, 183–191 (2017).
83. Shashi, V. *et al.* De Novo Truncating Variants in ASXL2 Are Associated with a Unique and Recognizable Clinical Phenotype. *Am. J. Hum. Genet.* **99**, 991–999 (2016).
84. Russell, B., Tan, W.-H. & Graham, J. M., Jr. Bohring-Opitz Syndrome. in *GeneReviews®* (eds. Adam, M. P. *et al.*) (University of Washington, Seattle, Seattle (WA), 2018).
85. Bohring, A. *et al.* Severe end of Opitz trigonocephaly (C) syndrome or new syndrome? *Am. J. Med. Genet.* **85**, 438–446 (1999).
86. Bohring, A., Oudesluijs, G. G., Grange, D. K., Zampino, G. & Thierry, P. New cases of Bohring-Opitz syndrome, update, and critical review of the literature. *Am. J. Med. Genet. A* **140**, 1257–1263 (2006).
87. Hoischen, A. *et al.* De novo nonsense mutations in ASXL1 cause Bohring-Opitz syndrome. *Nat. Genet.* **43**, 729–731 (2011).
88. Verhoeven, W. *et al.* Phenotypic characterization of an older adult male with late-onset epilepsy and a novel mutation in ASXL3 shows overlap with the associated Bainbridge-Ropers syndrome. *Neuropsychiatr. Dis. Treat.* **14**, 867–870 (2018).
89. Russell, B. E. *et al.* Clinical findings in 39 individuals with Bohring-Opitz syndrome from a global patient-driven registry with implications for tumor surveillance and recurrence risk. *Am. J. Med. Genet. A* **191**, 1050–1058 (2023).
90. Schirwani, S. *et al.* Expanding the phenotype of ASXL3-related syndrome: A comprehensive description of 45 unpublished individuals with inherited and de novo

- pathogenic variants in ASXL3. *Am. J. Med. Genet. A* **185**, 3446–3458 (2021).
91. Koboldt, D. C. *et al.* A de novo nonsense mutation in ASXL3 shared by siblings with Bainbridge-Ropers syndrome. *Cold Spring Harb Mol Case Stud* **4**, (2018).
 92. Schirwani, S. *et al.* Mosaicism in ASXL3-related syndrome: Description of five patients from three families. *Eur. J. Med. Genet.* **63**, 103925 (2020).
 93. Cuddapah, V. A. *et al.* Understanding the phenotypic spectrum of ASXL-related disease: Ten cases and a review of the literature. *Am. J. Med. Genet. A* **185**, 1700–1711 (2021).
 94. Avila, M. *et al.* Delineation of a new chromosome 20q11.2 duplication syndrome including the ASXL1 gene. *Am. J. Med. Genet. A* **161A**, 1594–1598 (2013).
 95. Cheraghi, S., Moghbelinejad, S., Najmabadi, H., Kahrizi, K. & Najafipour, R. A novel PTC mutation in the BTB domain of KLHL7 gene in two patients with Bohring-Opitz syndrome-like features. *Eur. J. Med. Genet.* **63**, 103849 (2020).
 96. Pratcorona, M. *et al.* Acquired mutations in ASXL1 in acute myeloid leukemia: prevalence and prognostic value. *Haematologica* **97**, 388–392 (2012).
 97. Genovese, G. *et al.* Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N. Engl. J. Med.* **371**, 2477–2487 (2014).
 98. Gelsi-Boyer, V. *et al.* Mutations in ASXL1 are associated with poor prognosis across the spectrum of malignant myeloid diseases. *J. Hematol. Oncol.* **5**, 12 (2012).
 99. Micol, J.-B. & Abdel-Wahab, O. The Role of Additional Sex Combs-Like Proteins in Cancer. *Cold Spring Harb. Perspect. Med.* **6**, (2016).
 100. Hastings, R. *et al.* Bohring–Opitz (Oberklaid–Danks) syndrome: clinical study, review of the literature, and discussion of possible pathogenesis. *Eur. J. Hum. Genet.* **19**, 513–519 (2011).
 101. Russell, B. *et al.* Clinical management of patients with ASXL1 mutations and Bohring-Opitz syndrome, emphasizing the need for Wilms tumor surveillance. *Am. J. Med. Genet. A* **167A**, 2122–2131 (2015).

102. Patel, K., McQuaid, S., Ketterl, T., Benedetti, D. J. & Sokol, E. Two cases of hepatoblastoma in Bohring-Opitz syndrome: An emerging association. *Pediatr. Blood Cancer* e31010 (2024).
103. Shiba, N. *et al.* Whole-exome sequencing reveals the spectrum of gene mutations and the clonal evolution patterns in paediatric acute myeloid leukaemia. *Br. J. Haematol.* **175**, 476–489 (2016).
104. Micol, J.-B. *et al.* Frequent ASXL2 mutations in acute myeloid leukemia patients with t(8;21)/RUNX1-RUNX1T1 chromosomal translocations. *Blood* **124**, 1445–1449 (2014).
105. Jahn, N. *et al.* Incidence and Prognostic Relevance of ASXL2 Mutations in Adult CBF-AML with t(8;21)(q22;q22): A Study of the German-Austrian AML Study Group (AMLSG). *Blood* **126**, 3818 (2015).
106. Szczepanski, A. P. *et al.* ASXL3 bridges BRD4 to BAP1 complex and governs enhancer activity in small cell lung cancer. *Genome Med.* **12**, 63 (2020).
107. Shukla, V. *et al.* ASXL3 Is a Novel Pluripotency Factor in Human Respiratory Epithelial Cells and a Potential Therapeutic Target in Small Cell Lung Cancer. *Cancer Res.* **77**, 6267–6281 (2017).
108. Cancer Genome Atlas Research Network *et al.* Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N. Engl. J. Med.* **368**, 2059–2074 (2013).
109. Masetti, R. *et al.* Genomic complexity and dynamics of clonal evolution in childhood acute myeloid leukemia studied with whole-exome sequencing. *Oncotarget* **7**, 56746–56757 (2016).
110. Oak, J. S. & Ohgami, R. S. Focusing on frequent ASXL1 mutations in myeloid neoplasms, and considering rarer ASXL2 and ASXL3 mutations. *Current medical research and opinion* vol. 33 781–782 (2017).
111. Duployez, N. *et al.* Unlike ASXL1 and ASXL2 mutations, ASXL3 mutations are rare events in acute myeloid leukemia with t(8;21). *Leuk. Lymphoma* **57**, 199–200 (2016).

112. Park, U.-H. *et al.* ASXL2 promotes proliferation of breast cancer cells by linking ER α to histone methylation. *Oncogene* **35**, 3742–3752 (2016).
113. Cui, R. *et al.* Elevated Expression of ASXL2 is Associated with Poor Prognosis in Colorectal Cancer by Enhancing Tumorigenesis and Inducing Cell Proliferation. *Cancer Manag. Res.* **12**, 10221–10228 (2020).
114. Wang, G. *et al.* Identification of Candidate Biomarker ASXL2 and Its Predictive Value in Pancreatic Carcinoma. *Front. Oncol.* **11**, 736694 (2021).
115. Tsuboyama, N. *et al.* Therapeutic targeting of BAP1/ASXL3 sub-complex in ASCL1-dependent small cell lung cancer. *Oncogene* **41**, 2152–2162 (2022).
116. Zhao, Z. *et al.* PAX9 Determines Epigenetic State Transition and Cell Fate in Cancer. *Cancer Res.* **81**, 4696–4708 (2021).
117. Ropers, H. H. & Wienker, T. Penetrance of pathogenic mutations in haploinsufficient genes for intellectual disability and related disorders. *Eur. J. Med. Genet.* **58**, 715–718 (2015).
118. Chen, S. *et al.* Author Correction: A genomic mutational constraint map using variation in 76,156 human genomes. *Nature* **626**, E1 (2024).
119. Gonzaludo, N., Belmont, J. W., Gainullin, V. G. & Taft, R. J. Estimating the burden and economic impact of pediatric genetic disease. *Genet. Med.* **21**, 1781–1789 (2019).
120. Zou, W. *et al.* Myeloid-specific Asxl2 deletion limits diet-induced obesity by regulating energy expenditure. *J. Clin. Invest.* **130**, 2644–2656 (2020).
121. Guo, Y. *et al.* ASXL1 alteration cooperates with JAK2V617F to accelerate myelofibrosis. *Leukemia* **33**, 1287–1291 (2019).
122. Fang, X. *et al.* Asxl1 C-terminal mutation perturbs neutrophil differentiation in zebrafish. *Leukemia* vol. 35 2299–2310 Preprint at <https://doi.org/10.1038/s41375-021-01121-8> (2021).
123. Shi, H. *et al.* ASXL1 plays an important role in erythropoiesis. *Sci. Rep.* **6**, 28789 (2016).
124. Moon, S., Um, S.-J. & Kim, E.-J. Role of Asxl1 in kidney podocyte development via its

- interaction with Wtip. *Biochem. Biophys. Res. Commun.* **466**, 560–566 (2015).
125. Dey, A. *et al.* Loss of the tumor suppressor BAP1 causes myeloid transformation. *Science* **337**, 1541–1546 (2012).
 126. Bai, J. *et al.* Reducing hyperactivated BAP1 attenuates mutant ASXL1-driven myeloid malignancies in human haematopoietic cells. *Cancer Lett.* **519**, 78–90 (2021).
 127. Lai, H.-L. *et al.* Maintenance of adult cardiac function requires the chromatin factor Asxl2. *J. Mol. Cell. Cardiol.* **53**, 734–741 (2012).
 128. Inoue, D. & Kitamura, T. C-terminal-truncating ASXL1 mutations induce MDS via repression of PRC2 function. *Experimental Hematology* vol. 41 S18 Preprint at <https://doi.org/10.1016/j.exphem.2013.05.067> (2013).
 129. Zhang, P. *et al.* Loss of ASXL1 in the bone marrow niche dysregulates hematopoietic stem and progenitor cell fates. *Cell Discov* **4**, 4 (2018).
 130. Braun, T. P. *et al.* Asxl1 deletion disrupts MYC and RNA polymerase II function in granulocyte progenitors. *Leukemia* **37**, 478–487 (2023).
 131. Yamamoto, K. *et al.* A histone modifier, ASXL1, interacts with NONO and is involved in paraspeckle formation in hematopoietic cells. *Cell Rep.* **36**, 109576 (2021).
 132. Zhao, J. *et al.* Novel truncating mutations in ASXL1 identified in two boys with Bohring-Opitz syndrome. *Eur. J. Med. Genet.* **64**, 104155 (2021).
 133. Zhu, X. *et al.* Whole-exome sequencing in undiagnosed genetic diseases: interpreting 119 trios. *Genet. Med.* **17**, 774–781 (2015).
 134. Carlston, C. M. *et al.* Pathogenic ASXL1 somatic variants in reference databases complicate germline variant interpretation for Bohring-Opitz Syndrome. *Hum. Mutat.* **38**, 517–523 (2017).
 135. Zhao, W. *et al.* A de novo Variant of ASXL1 Is Associated With an Atypical Phenotype of Bohring-Opitz Syndrome: Case Report and Literature Review. *Front Pediatr* **9**, 678615 (2021).

136. Leon, E. *et al.* Extending the phenotypic spectrum of Bohring-Opitz syndrome: Mild case confirmed by functional studies. *Am. J. Med. Genet. A* **182**, 201–204 (2020).
137. Urreizti, R. *et al.* The ASXL1 mutation p.Gly646Trpfs*12 found in a Turkish boy with Bohring-Opitz Syndrome. *Clin Case Rep* **6**, 1452–1456 (2018).
138. Kibe, M. *et al.* Lethal persistent pulmonary hypertension of the newborn in Bohring-Opitz syndrome. *Am. J. Med. Genet. A* **176**, 1245–1248 (2018).
139. Efthymiou, S. *et al.* A de novo truncating mutation in ASXL1 associated with segmental overgrowth. *J. Genet.* **98**, (2019).
140. Arunachal, G., Danda, S., Omprakash, S. & Kumar, S. A novel de-novo frameshift mutation of the ASXL1 gene in a classic case of Bohring-Opitz syndrome. *Clin. Dysmorphol.* **25**, 101–105 (2016).
141. Urreizti, R. *et al.* Screening of CD96 and ASXL1 in 11 patients with Opitz C or Bohring-Opitz syndromes. *Am. J. Med. Genet. A* **170A**, 24–31 (2016).
142. Magini, P. *et al.* Two novel patients with Bohring-Opitz syndrome caused by de novo ASXL1 mutations. *Am. J. Med. Genet. A* **158A**, 917–921 (2012).
143. Bedoukian, E., Copenheaver, D., Bale, S. & Deardorff, M. Bohring-Opitz syndrome caused by an ASXL1 mutation inherited from a germline mosaic mother. *Am. J. Med. Genet. A* **176**, 1249–1252 (2018).
144. Wang, D. *et al.* Bohring-Opitz syndrome caused by a novel ASXL1 mutation (c.3762delT) in an IVF baby: A case report. *Medicine* **101**, e28759 (2022).
145. Dangiolo, S. B., Wilson, A., Jobanputra, V. & Anyane-Yeboah, K. Bohring-Opitz syndrome (BOS) with a new ASXL1 pathogenic variant: Review of the most prevalent molecular and phenotypic features of the syndrome. *Am. J. Med. Genet. A* **167A**, 3161–3166 (2015).
146. Ayoub, M. C., Anderson, J. T., Russell, B. E. & Wilson, R. B. Examining the neurodevelopmental and motor phenotypes of Bohring-Opitz syndrome (ASXL1) and Bainbridge-Ropers syndrome (ASXL3). *Front. Neurosci.* **17**, 1244176 (2023).

147. Awamleh, Z. *et al.* DNA methylation signature associated with Bohring-Opitz syndrome: a new tool for functional classification of variants in ASXL genes. *Eur. J. Hum. Genet.* (2022) doi:10.1038/s41431-022-01083-0.
148. Zheng, Y. *et al.* Identification of a de novo variant in the ASXL2 gene related to Shashi-Pena syndrome. *Mol Genet Genomic Med* **11**, e2251 (2023).
149. Ho, S. K. L. *et al.* Extending the phenotype of Shashi-Pena syndrome: a case report and review of literature. *Clin. Dysmorphol.* **32**, 139–146 (2023).
150. Jiao, Z. *et al.* A de novo and novel nonsense variants in ASXL2 gene is associated with Shashi-Pena syndrome. *Eur. J. Med. Genet.* **65**, 104454 (2022).
151. Murphy, S. A., O'Rourke, N. E. & O'Grady, M. J. Prepubertal onset of type 2 diabetes in Shashi-Pena syndrome due to ASXL2 mutation. *Am. J. Med. Genet. A* **188**, 2803–2807 (2022).
152. Wang, Q. *et al.* De novo nonsense variant in ASXL3 in a Chinese girl causing Bainbridge-Ropers syndrome: A case report and review of literature. *Mol Genet Genomic Med* **10**, e1924 (2022).
153. Balasubramanian, M. *et al.* Delineating the phenotypic spectrum of Bainbridge-Ropers syndrome: 12 new patients with de novo, heterozygous, loss-of-function mutations in ASXL3 and review of published literature. *J. Med. Genet.* **54**, 537–543 (2017).
154. Švantnerová, J. *et al.* ASXL3 De Novo Variant-Related Neurodevelopmental Disorder Presenting as Dystonic Cerebral Palsy. *Neuropediatrics* **53**, 361–365 (2022).
155. Alqaisi, D. 'a & Hassona, Y. Oral findings and healthcare management in Shashi-Pena syndrome. *Spec. Care Dentist.* **42**, 432–436 (2022).
156. Dad, R. *et al.* Hyperventilation-athetosis in ASXL3 deficiency (Bainbridge-Ropers) syndrome. *Neurol Genet* **3**, e189 (2017).
157. Bacrot, S. *et al.* Whole exome sequencing diagnoses the first fetal case of Bainbridge-Ropers syndrome presenting as pontocerebellar hypoplasia type 1. *Birth*

- Defects Res* **110**, 538–542 (2018).
158. Zheng S., Chen H. & Mo M. [A case of Bainbridge-Ropers syndrome with autism in conjunct with ASXL3 gene variant and its clinical analysis]. *Zhonghua Yi Xue Yi Chuan Xue Za Zhi* **38**, 671–673 (2021).
159. Li J. *et al.* [Genetic analysis and prenatal diagnosis for a Chinese pedigree affected with Bainbridge-Ropers syndrome]. *Zhonghua Yi Xue Yi Chuan Xue Za Zhi* **39**, 1228–1232 (2022).
160. Li, J.-R. *et al.* Novel mutation in the ASXL3 gene in a Chinese boy with microcephaly and speech impairment: A case report. *World J Clin Cases* **8**, 6465–6472 (2020).
161. Dinwiddie, D. L. *et al.* De novo frameshift mutation in ASXL3 in a patient with global developmental delay, microcephaly, and craniofacial anomalies. *BMC Med. Genomics* **6**, 32 (2013).
162. Schirwani, S. *et al.* Familial Bainbridge-Ropers syndrome: Report of familial ASXL3 inheritance and a milder phenotype. *Am. J. Med. Genet. A* **191**, 29–36 (2023).
163. Contreras-Capetillo, S. N., Vilchis-Zapata, Z. H., Ribbón-Conde, J. & Pinto-Escalante, D. Global developmental delay and postnatal microcephaly: Bainbridge-Ropers syndrome with a new mutation in ASXL3. *Neurol. (Engl. Ed.)* (2018) doi:10.1016/j.nrleng.2017.01.005.
164. Wayhelova, M. *et al.* Novel de novo frameshift variant in the ASXL3 gene in a child with microcephaly and global developmental delay. *Mol. Med. Rep.* **20**, 505–512 (2019).
165. Chinen, Y. *et al.* Mild prominence of the Sylvian fissure in a Bainbridge-Ropers syndrome patient with a novel frameshift variant in ASXL3. *Clin Case Rep* **6**, 330–336 (2018).
166. Myers, K. A. *et al.* Childhood-onset generalized epilepsy in Bainbridge-Ropers syndrome. *Epilepsy Res.* **140**, 166–170 (2018).
167. Hori, I. *et al.* Novel splicing mutation in the ASXL3 gene causing Bainbridge-Ropers syndrome. *Am. J. Med. Genet. A* **170**, 1863–1867 (2016).
168. Khan, T. R., Dolce, A. & Goodspeed, K. A case of Bainbridge-Ropers syndrome with breath

- holding spells and intractable epilepsy: challenges in diagnosis and management. *BMC Neurol.* **22**, 60 (2022).
169. Wu, K. & Cong, Y. Case report : a novel ASXL3 gene variant in a Sudanese boy. *BMC Pediatr.* **21**, 557 (2021).
170. Guo, H. *et al.* Inherited and multiple de novo mutations in autism/developmental delay risk genes suggest a multifactorial model. *Mol. Autism* **9**, 64 (2018).
171. Duan F., Zhai Y. & Kong X. [Analysis of ASXL3 gene variant in a child with Bainbridge-Ropers syndrome]. *Zhonghua Yi Xue Yi Chuan Xue Za Zhi* **38**, 275–277 (2021).
172. Zhang R., He X. H., Lin H. Y. & Yang X. H. [Bainbridge-Ropers syndrome with ASXL3 gene variation in a child and literature review]. *Zhonghua Er Ke Za Zhi* **56**, 138–141 (2018).
173. Qiao, L., Liu, Y., Ge, J. & Li, T. Novel Nonsense Mutation in ASXL3 causing Bainbridge-Ropers Syndrome. *Indian Pediatr.* **56**, 792–794 (2019).
174. Yang, L. *et al.* Bainbridge-ropers syndrome caused by loss-of-function variants in ASXL3: Clinical abnormalities, medical imaging features, and gene variation in infancy of case report. *BMC Pediatr.* **20**, 287 (2020).
175. Lyu Y. *et al.* [Diagnosis of Bainbridge-Ropers syndrome due to de novo ASXL3 variant by high throughput sequencing]. *Zhonghua Yi Xue Yi Chuan Xue Za Zhi* **37**, 452–454 (2020).
176. Fu, C. *et al.* Chromosomal microarray and whole exome sequencing identify genetic causes of congenital hypothyroidism with extra-thyroidal congenital malformations. *Clin. Chim. Acta* **489**, 103–108 (2019).

CHAPTER 2

Multi-omics of Bohring-Opitz syndrome: Identifying Wnt Signaling Dysregulation

Abstract

ASXL1 (additional sex combs–like 1) plays a crucial role in the epigenetic regulation of early developmental gene expression. *De novo* protein-truncating mutations in *ASXL1* lead to Bohring-Opitz syndrome (BOS; OMIM #605039), a rare neurodevelopmental condition characterized by severe intellectual disabilities, distinctive facial features, hypertrichosis, an increased risk of Wilms tumor, and variable congenital anomalies, including heart defects and severe skeletal abnormalities that give rise to a typical BOS posture. Notably, these *ASXL1* variants responsible for BOS are also prevalent somatic driver mutations in acute myeloid leukemia.

In our study, we utilized primary cells from individuals with BOS (n = 18) and controls (n = 49) to dissect the gene regulatory changes induced by *ASXL1* mutations. We employed comprehensive multi-omics assays, including chromatin accessibility, DNA methylation, histone methylation binding, and transcriptome analysis in peripheral blood and skin fibroblasts. Our findings reveal that, irrespective of the cell type, *ASXL1* mutations exert strong cross-tissue effects, disrupting multiple layers of the epigenome.

The data indicate a broad activation of canonical Wnt signaling at both the transcriptional and protein levels. Additionally, we observed an upregulation of *VANGL2*, a gene encoding a planar cell polarity pathway protein that functions through noncanonical Wnt signaling to guide tissue patterning and cell migration.

This multiomics approach not only identifies the fundamental impact of *ASXL1* mutations but also reveals potential therapeutic targets for addressing both BOS and myeloid leukemias.

Introduction

Role of ASXL1 in chromatin modification and gene regulation.

ASXL1 (additional sex combs–like 1) is a crucial protein in embryonic development ¹. Eutherian mammals possess three ASX homologs, namely ASXL1, ASXL2, and ASXL3, sharing conserved domains: the ASX N-terminal (ASXN) domain with a HARE-HTH domain involved in putative DNA binding ², the ASX homology (ASXH) domain featuring a characteristic LXXLL motif mediating protein interactions, and a deubiquitinase adaptor (DEUBAD) domain activating BRCA1-associated protein 1 (BAP1), along with the C-terminal plant homeodomain (PHD) important for reading post translational histone modifications like histone H3 Lys4 trimethylation (H3K4me3) ^{3–5}. ASXL1 plays a pivotal role in modulating the epigenetic landscape, regulating downstream transcription. The ASXH and PHD domains are highly conserved and integral to large epigenetic complexes such as the polycomb repressive complexes (PRCs) and the polycomb-repressive deubiquitinase (PR-DUB) complex. These complexes direct genome-wide transcriptional regulation ^{3,6,7}, influencing activation or repression of essential developmental genes.

In mammals, *ASXL1* is essential in early neural crest ⁸, cardiac ⁹, and hematopoietic development ^{10,11}. *ASXL1* directs histone modifications through PRC1, PRC2 and the PR-DUB complex. PRC1 and -2 control developmental stemness and differentiation, impacting neural progenitor cell multipotency and cortical neuron production during corticogenesis ¹². These complexes achieve this through histone modifications like monoubiquitination at histone H2A lysine 119 (H2AK119ub) via PRC1 and mono-, di-, and trimethylation at histone H3K27 (H3K27me/me2/me3) via PRC2 ^{3,13,14}. These complexes also regulate key signaling pathways, such as Wnt signaling, in tissue-specific and developmental contexts, with PRC2 accessory proteins influencing Wnt signaling during erythropoiesis ¹⁵ and PRC1 participating in the Wnt/ β -catenin pathway through a positive feedback loop in hepatocellular carcinoma ¹⁶.

The PR-DUB complex, including ASXL1 binding to BAP1^{17–19}, plays critical roles in brain development^{20–22} and myeloid differentiation regulation through H2AK119 deubiquitination^{18,23}. Despite recognizing that *ASXL1* protein-truncating mutations disrupt core developmental processes in multiple organ systems, its role as an essential chromatin modifier in human development remains incompletely elucidated.

ASXL1 mutations are implicated in Bohring-Opitz syndrome and hematologic malignancies.

De novo protein-truncating mutations in *ASXL1* cause Bohring-Opitz syndrome (BOS; OMIM #605039), a rare genetic disorder characterized by profound intellectual disability, developmental delay, seizures, and variable anomalies, including heart defects, an increased risk of Wilms tumor, and the distinctive BOS posture^{24,25}. As of 2018, only 46 clinically diagnosed individuals have been reported, with less than half (20/46) molecularly confirmed²⁴. Pathogenic truncating mutations in *ASXL1* concentrate in exons 12 and 13, resulting in premature truncation of the highly conserved C-terminal PHD domain³. Notably, truncation mutations in *ASXL1* homologs *ASXL2* (Shashi-Pena syndrome, OMIM #6171901)²⁶ and *ASXL3* (Bainbridge-Ropers syndrome, OMIM #615485)^{22,27} do not exhibit enrichment or clustering in the homologous C-terminal region despite sharing multiple conserved domains³.

The same protein-truncating mutations causing BOS are observed as somatic driver mutations in myeloid malignancies such as chronic myelomonocytic leukemia (~45%), myelodysplastic syndromes (MDS, 16%), myeloproliferative neoplasms (~10%), and acute myeloid leukemia (AML, secondary 30%, *de novo* 6.5%)^{28–31}, as well as in a premalignancy state called clonal hematopoiesis of indeterminate potential^{32,33}. *ASXL1* regulates the delicate balance between proliferation and differentiation of stem progenitor cell populations^{8,34}. Thus, both germline and somatic *ASXL1* mutations disrupt the proliferation-differentiation balance, promoting stem-cell identity over differentiation in BOS and myeloid leukemias.

Recent ASXL1 functional studies involve transgenic overexpression of mutant and wild-type ASXL1 proteins in cell-line systems. However, interpreting these studies is challenging as they do not reflect endogenous levels or cell-type-specific functions of ASXL1, and are often conducted in nonhuman model systems ³⁵. Given the crucial functions of ASXL1 as part of PRC1 and PRC2 in neural progenitor multipotency ¹² and the key role of the PR-DUB complex in brain development ^{20–22}, we hypothesized that protein-truncating mutations of *ASXL1* dysregulate global transcription, cellular homeostasis, and downstream signaling pathways.

We also explored whether BOS cells harbored decreased mRNA or protein levels since protein-truncating ASXL1 mutations in exons 12 and 13 are predicted to escape nonsense-mediated decay ³⁶. These truncated proteins may exert a gain-of-function effect ³⁷. The broad phenotypic effects of *ASXL1* mutations in BOS patients and myeloid malignancies suggest that *ASXL1* drives essential and core gene regulatory features across early development and disease. Our hypothesis posited that *ASXL1* mutations drive common cellular effects, detectable as broad and high-effect epigenomic and transcriptomic signatures across cells and tissues.

To investigate BOS molecular pathogenesis, we collected samples from 18 patients with BOS, comprising one of the largest *ASXL1* rare disease cohorts published, and conducted a multi-omics analysis. Our approach circumvents confounding variables associated with cancer-derived cell models that harbor additional karyotypic and genomic mutations ^{30,38}, focusing the study on the singular effect of pathogenic *ASXL1* mutations. This integrated multi-omics study characterizes the impact of *ASXL1* truncating mutations on the epigenome and transcriptome, revealing for the first time, to our knowledge, that *ASXL1* mutations aberrantly activate the Wnt signaling and noncanonical Wnt planar cell polarity (PCP) pathways.

Canonical Wnt signaling pathway.

Canonical Wnt signaling is an evolutionarily conserved pathway essential for development ³⁹ and plays critical roles in stem cell biology and hematopoiesis regulation ⁴⁰. In

the Wnt-active state, WNTs (translated products of the *WNT* gene) bind to the transmembrane receptor Frizzled (FZD), stimulating the coreceptor low density receptor-related protein 5 and 6 (LRP5 and -6), critical for initiating Wnt signaling transduction ⁴¹. Activation of FZD and LRP5/6 recruits and inactivates the destruction complex targeting phosphorylation, ubiquitination, and degradation of β -catenin through the proteasome. Inactivation of the destruction complex occurs through the dissociation of glycogen synthase kinase 3 β (GSK3 β) from the inhibition protein AXIN, concurrent with inactivation of the default quiescent-state β -catenin destruction complex ^{42–44}. This prevents β -catenin phosphorylation and subsequent degradation ^{45,46}. Activation of the canonical Wnt pathway leads to cytoplasmic accumulation and subsequent nuclear translocation of β -catenin, which, along with T cell factor/lymphoid enhancer factor (TCF/LEF) transcription factors, transcriptionally coactivates Wnt target genes ⁴⁷.

To delve deeper into cross-omics dysregulation, we explored van Gogh–like 2 (VANGL2), a noncanonical Wnt and PCP protein highly dysregulated across all the -omics assays conducted on blood and fibroblast BOS samples. VANGL2 regulates polarized cellular migration, differentiation, and tissue morphology during development. In the nervous system, the PCP pathway regulates neuronal maturation, playing a functional role in neural complex formation ⁴⁸. VANGL2 intersects with the canonical Wnt pathway through the activation of Dishevelled (DVL), initiating noncanonical Wnt pathways and driving c-JUN–mediated expression. Our hypothesis suggests that upregulation of *VANGL2* disrupts canonical and noncanonical signaling pathways, underpinning the clinical phenotypes observed in patients with BOS and other mutations in *ASXL1*. To date, no studies have established a link between *ASXL1* disorders and dysregulation of Wnt signaling.

In this study, we identify and investigate the mechanisms by which truncating *ASXL1* mutations in BOS may dysregulate the canonical and noncanonical Wnt signaling pathways. These findings reveal what we believe are novel regulatory mechanisms underlying BOS pathogenesis.

Methods

Patient Recruitment

This project was approved by the UCLA IRB# 11-001087. In conjunction with Dr. Bianca Russell and the ASXL Biobank at the University of California, Los Angeles, written informed consent from participants and guardians were obtained for skin punch biopsies and blood samples. Deep phenotyping and review of medical records was also collected.

Sample Collection

All individual level-data was de-identified prior to analysis and samples were collected for follow-up experiments. We collected blood from 12 BOS individuals and 38 controls, and skin punch biopsies from 6 individuals and 7 controls (**Table 2-1 and Table 2-2**). We used two independent samples (blood and patient-derived fibroblast) from 4/18 BOS individuals to filter out noise of genetic background. Whole blood was processed to isolate DNA, RNA and PBMCs and skin-punch biopsy samples were processed to create patient-derived skin fibroblasts in the Translational Pathology Core Laboratory at UCLA.

Blood Sample Collection and Processing.

Peripheral blood samples were collected into EDTA tubes for DNA extraction and PBMC extraction, CPT tubes for PBMC extraction and PAXgene™ Blood RNA Tube (BDBiosciences, 762165) for RNA extraction. Whole blood DNA was extracted using the Qiagen Blood and Cell Culture DNA Midi Kit (Qiagen #13343), and whole blood RNA was extracted using the MagMAX™ for Stabilized Blood Tubes RNA Isolation kit (Thermo Fisher Scientific, #4451893) with DNase treatment. All samples were quantified on the nanodrop, Qubit 4.0 and TapeStation Bioanalyzer for sample quality control. DNA and RNA were stored at -80C.

Cell Culture

Patient-derived fibroblast cell lines were grown in Dulbecco's medium (DMEM, Gibco™, #11-995-073), 10% fetal bovine serum (FBS), 1% Non-essential Amino Acid (NEAA, 100X,

Gibco™, 11140-050) and 1% PenStrep at 37°C in 5% CO₂ incubators. All cell lines were tested for mycoplasma (MycoAlert PLUS Mycoplasma Detection Kit, Lonza #LT07-318) on a monthly basis.

Extractions

Whole Cell Lysis Extraction

Cells were harvested in ice-cold 1X cell lysis buffer (10X, Cell Signaling #9803S) containing 1X Halt™ Protease and Phosphatase Inhibitor Cocktail (100X, Thermo Scientific™ #PI78442). Soluble lysate fractions were isolated by constant agitation at 4°C for 30 minutes at 1000 rpm and then centrifugation at 20,000 × g, for 20 min at 4 °C.

Cytoplasmic and Nuclear Extraction

1 million cells were harvested into a 1.5ml eppendorf tube and centrifuged at 500 ×g for 5 min to form a pellet. Cells were washed with ice cold PBS with 0.5mM Sodium Butyrate and centrifuged at 500 × g for 2 min to form a pellet. After centrifugation, the supernatant was discarded. Ice-cold CERI with 1X Halt™ Protease and Phosphatase Inhibitor Cocktail (100X, Thermo Scientific™ #PI78442) was added to each cell pellet. Cytoplasmic and nuclear extraction was conducted using the Thermo Scientific CE/NER Kit (#78833) according to protocol.

Histone Extraction

Histones were extracted as per established protocol ⁴⁹.

Western Blot

Protein was quantified using the Pierce BCA Protein Assay Kit (Thermo Scientific #23225). For Western blots, we used 15ug or 30µg of whole cell lysate, 10µg of cytoplasmic extract, 3ug of nuclear extract or 1ug of acid-extracted histone per well for 30 minutes at 130V. Proteins were then transferred to 0.2uM nitrocellulose membranes (BioRad #1704271) using

the semi-dry TransBlot for 7 min at 25 mA. The membranes were blocked with 5% milk in 1x TBST and then probed with primary antibodies (1:1000) overnight at 4°C. The antibodies for whole cell lysate, cytoplasmic and nuclear extracts are shown in **Table 2-3**.

Membranes were then probed with secondary antibodies (1:10,000), at room temperature for one hour according to manufacturer dilutions. These secondary antibodies were Goat anti-Rabbit IgG Highly Cross-Adsorbed, Alexa Fluor™ 680 (Invitrogen #A-32734) and Goat anti-Mouse IgG Highly Cross-Adsorbed, Alexa Fluor™ 488 (Invitrogen #A-11001). Blots were imaged on a Biorad ChemiDoc.

DNA Gel Electrophoresis and Gel Extraction

DNA fragments were analyzed by electrophoresis on a 2% agarose gel in 1xTAE buffer at 90V for 30-60 minutes. The gel was stained with SYBR™ Safe DNA Gel Stain (Invitrogen, S33102) for visualization under UV light. Target bands were excised from the gel using a clean scalpel and extracted using the Zymoclean Gel DNA Recovery Kit (Zymo, #D4001) following the manufacturer's instructions. The DNA was eluted in 30 µl of Elution Buffer and quantified using a NanoDrop™ spectrophotometer (Thermo Fisher Scientific).

Sanger sequencing by Genewiz

Gel-extracted PCR products and plasmids were sequenced through GeneWiz and visualized with SnapGene.

Lipofectamine 3000 transfections

For transfection experiments, fibroblast cells were grown in DMEM (Gibco™, #11-995-073), 10% FBS, 1% NEAA, and 1% PenStrep, and HEK293T cells and CACO2 cells were grown in DMEM (Gibco™, #11-995-073), 10% FBS, and 1% PenStrep at 37°C in 5% CO2 incubators. CACO2 cells were a gift from the Tontono lab at the University of California, Los Angeles.

RT-qPCR

RT-qPCR was conducted using TaqPath™ 1-Step RT-qPCR Master Mix, CG (Applied Biosystems, A15300) according to protocol (Document #100020171, Rev 1.00) on QuantStudio™ 3.

10ng of RNA was used per well in 20ul total well volume, and each sample and primer combination was conducted in triplicate. Machine settings were for Standard Curve and Fast run mode. RNA isolated from patient blood and patient-derived fibroblast cell lines were subject to quantitative reverse transcriptase polymerase chain reaction (RT-qPCR) to validate RNA-seq findings. RNA isolated from transfection experiments were subject to RT-qPCR to identify effects of the ASXL1 truncated plasmids or ASXL1 siRNAs on cell transcriptome.

Next Generation Sequencing

RNA-sequencing

Fibroblasts were grown to 80-90% confluency in T75 flasks before total RNA extraction via RNeasy Plus Mini Kit (Qiagen, #74136). Blood samples were first treated with QIAseq FastSelect –Globin (Qiagen, #334376). All samples were then prepared with QIAseq FastSelect –rRNA HMR (Qiagen, #334386) and Illumina Truseq® Stranded Total RNA Library Prep Gold (Illumina, 20020599), before sequencing to a minimum of 30 million paired end 150bp reads per sample.

Fibroblast and blood samples were quantified on the Qubit 4.0 and Tapestation Bioanalyzer 4150 using High Sensitivity RNA Tapestation (Agilent, #5067-5579) to ensure high-quality RNA (RIN $\geq$ 9.0 for fibroblasts, RIN $\geq$ 5.3 for blood). All libraries were quantified on the Qubit 4.0 and library quality was assessed using Agilent D1000 DNA Tapestation (Agilent, #5067-5582) before multiplexing and sequencing on a NovaSeq flow cell for a minimum of 30 million paired-end 150bp reads per sample.

ATAC-sequencing

Fibroblast lines were cultured to 80-90% confluency in T75 flasks and 50,000 freshly isolated cells per line were treated with Tn5 transposase as per established protocol⁵⁰. Libraries were sequenced to a minimum of 40 million paired-end reads per sample with 75bp length. After purification and PCR amplification for library generation, the libraries were double-sided bead purified using AMPure XP beads (Beckman, A63881) to remove contaminating primer dimers. All libraries were quantified on the Qubit 4.0 and library quality was assessed using Agilent High Sensitivity DNA TapeStation (Agilent, 5067-5584) before multiplexing and sequencing on a Nextseq 550 for a minimum of 40 million paired-end reads per sample with 75bp length.

Cleavage Under Targets & Release Using Nuclease (CUT&RUN)

CUT&RUN libraries were prepared with the established CUT&RUN Assay Kit (Cell Signaling, #86652S). CUT&RUN libraries were prepared using the SimpleChIP ChIP-seq DNA Library Prep Kit for Illumina (Cell Signaling, #56795S) with SimpleChIP ChIP-seq Multiplex Oligos for Illumina (Cell Signaling, #47538S) according to manufacturer guidelines, with modifications for CUT&RUN as noted in the CUT&RUN Assay Kit (Cell Signaling, #86652S). The following antibodies were used for CUT&RUN libraries: IgG isotype (DA1E) (Cell Signaling, #66362, negative control, 1:20), H3K4me3 (C42D8) (Cell Signaling, #9751S, 1:50) and H3K27me3 (C36B11) (Cell Signaling, #9733S, 1:50) according to protocol. For DNA purification, the DNA Purification Buffers and Spin Columns for ChIP and CUT&RUN (Cell Signaling, #14209), with elution in 50 μ l. All libraries were quantified on the Qubit 4.0 and library quality assessed using Agilent High Sensitivity DNA TapeStation (Agilent, 5067-5584) before multiplexing and sequencing on a Nextseq 550 Mid Output Kit (150 cycles) with 1% PhiX for a minimum of 5 million paired-end reads per sample with 50bp read length.

Samples were analyzed using the established NF-Core CUT&RUN analysis pipeline⁵¹.

DNA methylation

DNA methylation (DNAm) data underwent processing through our established pipeline⁵². For DNA methylation (DNAm), the absolute difference between the means of β values for BOS patients versus controls was calculated for each CpG, obtaining the delta beta ($\Delta\beta$) value through linear regression modeling. Highly differentially methylated sites ($|\Delta\beta| > 5\%$) were filtered, and significant CpG sites were identified with FDR < 0.05. To discern dysregulated biological mechanisms, queried CpG sites were analyzed using GREAT⁵³, filtering for gene ontologies with $p_{\text{adj}} < 0.05$. Principal component analysis (PCA) and hierarchical clustering utilized the BOS DNAm episignature sites (413 CpG sites)⁵².

Data Analysis

RNA-seq Bioinformatic Analysis.

Samples were demultiplexed after sequencing, and fastq sequencing files were processed using established pipelines in the Arboleda lab (**Figure 2-1A**). Raw read quality, adaptor content, and duplication rates were assessed with *FastQC v0.11.8*.⁵⁴ Raw reads were then aligned against the Gencode human genome version hg38 (GR38) version 31 using STAR 2.7.0e⁵⁵ with default parameters. Gene counts from raw reads were generated using featureCounts 1.6.5⁵⁶ from the Subread package. Only reads that uniquely mapped to exons of a gene were counted per gene. Differential expression analysis was completed using DESeq2 v1.24.0⁵⁷, adjusting for sample gender. Multidimensional scaling based on the top 500 genes with the highest variance was performed and verified that BOS samples clustered together and away from controls. We confirmed effective ribosomal RNA (rRNA) removal.

Differential expression analysis was performed, adjusting for sample sex. Genes with a p-adjusted (Benjamini-Hochberg) less than 0.05 were classified as significantly differentially expressed (Wald's test), and fold changes⁵⁷ were shrunk using approximate posterior estimation for GLM coefficients. We further filtered our differentially expressed genes (DEGs) for

$\text{abs}(\log_2\text{FC}) \geq 0.58$, corresponding to an absolute log fold change ≥ 1.5 . Gene ontology ^{58,59} over-enrichment tests were completed using clusterProfiler v3.12.0 ⁶⁰ by submitting differentially expressed genes against all genes from the Gencode hg38 annotation, version 31. Gene ontologies were classified as significantly enriched when p-adjusted (Benjamini-Hochberg) was less than 0.05 (hypergeometric test). *HOMER v4.9 findMotifs.pl* ⁶¹ was used to identify enrichment of motifs within differentially expressed genes' promoters. Differentially expressed genes were compared against HOMER's default human, RefSeq-based promoter set, which yielded de novo and known motif enrichments for motif lengths of 8, 10, and 12. Sequencing tracks were visualized using Integrated Genomic Viewer (IGV 2.9.4) and GVIZ.

ATACseq bioinformatic analysis.

Fastq files were processed through our best practise bioinformatic in-house pipeline (**Figure 2-1B**). Quality of reads were assessed using FastQC. Raw reads were then aligned to GENCODE Human genome version hg38 (GR38) version 31 using BWA-MEM ⁶². BAM files were then sorted, indexed, filtered against chrX, chrY, and MT reads using *SAMtools*. *Picard* tools were then used to generate insert size histograms and remove duplicates from BAM files. Narrow peaks from each sample were called using MACS2 callpeak ⁶³; any peak that overlapped by at least one base was then merged using BEDtools merge ⁶⁴ *merge*. Reads overlapping merged peaks were counted using featureCounts ⁵⁶. Technical replicates clustered closely on PCA and were collapsed during analysis. Peaks were annotated using *annotatePeaks.pl* from HOMER to the nearest gene or regulatory element, including exon, intron, promoter, 5' untranslated region (UTR), 3' UTR, and other genomic features ⁶¹.

DESeq2 was used to identify differentially open peaks between disease and control samples, adjusting for sample gender. Peaks with p-adjusted value (Benjamini-Hochberg) less than 0.05 were classified as significantly differentially open (Wald's test), and fold changes were shrunk using approximate posterior estimation for GLM coefficients. Significant peaks were

identified as promoter peaks if their distance from their respective closest gene was less than 1kb upstream or 2kb downstream relative to the gene transcription start site. Gene ontology over-enrichment tests were completed using clusterProfiler by submitting the closest genes to significantly differentially open peaks against all genes from the Gencode hg38 annotation, version 31. Gene ontologies were classified as significantly enriched when p-adjusted (Benjamini-Hochberg) was less than 0.05 (hypergeometric test). Enriched de novo and known motifs were identified in significant peaks using findMotifsGenome.pl from HOMER ⁶¹. Significant peaks were identified as promoter peaks if their distance from their respective closest gene was less than 1kb upstream or 2kb downstream relative to the gene transcription start site. Sequencing tracks were visualized using Integrated Genomic Viewer (IGV 2.9.4) and GVIZ.

DNA methylation bioinformatic analysis.

We used DNA methylation generated from BOS and control blood and fibroblasts and processed as described in Awamleh *et al.* ⁵². We calculated the absolute difference between the means of the β value for BOS patients versus controls for each CpG to obtain the delta beta ($\Delta\beta$) value using linear regression modeling. We filtered for highly differentially methylated sites ($|\Delta\beta| > 5\%$). Significant CpG sites were identified if they were below a false discovery rate (FDR) < 0.05 . To identify biological mechanisms that were dysregulated, we queried these CpG sites using GREAT ⁵³. We filtered for gene ontologies with padj < 0.05 for significance.

Statistics

RNA-seq, ATAC-seq, and CUT&RUN reads were aligned to the human genome (hg38) and featureCounts (v1.6.5) was used to generate count matrices for genes or chromatin regions. DESeq2 (v1.24.0), which internally normalizes sample library size and utilizes the negative binomial distribution for testing, was used to identify differentially expressed genes (RNA-seq) or differentially accessible chromatin regions (ATAC-seq, CUT&RUN) after adjusting

for sample sex, which was identified as a covariate to adjust for through PCAs generated based on the top 500 most variably expressed genes. Genes or regions were identified as significant if Benjamini-Hochberg adjusted p-values (p_{adj}) < 0.05 . We further filtered for highly dysregulated significant genes or regions using an $abs(log2FoldChange) \geq 0.58$, corresponding to an $abs(FoldChange) \geq 1.5$ with reference to control samples. Gene ontology over-enrichment tests were completed using clusterProfiler v3.12.0 by submitting differentially expressed genes against all genes from the Gencode hg38 annotation, version 31. Gene ontologies were classified as significantly enriched when p_{adj} (Benjamini-Hochberg) < 0.05 (hypergeometric test).

DNA β (β) values were calculated using minfi⁶⁵, and differentially methylated sites were identified by running Limma regression modeling, while adjusting for the covariates of age, sex, and ethnicity. Our significance cut-offs were $FDR < 0.05$ (Benjamini-Hochberg). We filtered for significantly methylated CpG sites with $abs(\Delta\beta) > 0.10$, where $\Delta\beta$ represents the difference in average DNA β (β) between BOS and controls. GREAT⁵³ was run on significant CpG sites to identify biological mechanisms that were dysregulated, filtering for gene ontologies with $p_{adj} < 0.05$ for significance.

Results

Study design.

We recruited 18 individuals with BOS (**Table 2-1**), 25 sex-matched and genetically related controls, and 30 sex- and age-matched controls (**Table 2-2**), comprising the largest cohort of BOS patients studied at the molecular level. All BOS individuals had clinically identified *ASXL1* mutations in the last 2 exons of the gene predicted to cause protein-truncating variants (NM_015338.6), between amino acids 364 to 1415 (**Figure 2-2A**, **Table 2-1**). A summary of patient demographics is provided in **Table 2-4**. This cohort includes novel patients first reported

in 2023 by Russell et al ⁶⁶, and a subset of patients were previously reported by Awamleh et al. ⁵². Clinical data are available through the ASXL registry.

Consent was obtained from individuals, who donated peripheral blood (n=14) , a skin punch biopsy (n=8), or both. Multi-omics analysis was conducted across tissues, examining chromatin accessibility (ATAC-seq)⁶⁷, histone modifications H3K4me3 and H3K27me3 enrichment using cleavage under targets and release using nuclease (CUT&RUN)⁶⁸, DNA methylation (DNAm, Illumina Epic Arrays)⁵², and global transcriptome (RNA-seq) (**Figure 2-2B**, **Figure 2-2C**). The study aimed to understand tissue-dependent and -independent dysregulation caused by truncating *ASXL1* mutations. We derived disease-specific DNAm signatures ⁵² and transcriptional signatures from RNA-seq using blood and fibroblast cells (**Figure 2-2B**), and integrated our findings to examine the trans-omics dysregulation (**Figure 2-2D**).

ASXL1 transcript and protein expression levels were not significantly different in BOS patient samples.

ASXL1 exhibits a broad expression profile across various tissues and cell types, consistent with the clinical manifestation affecting multiple organ systems. Thus, we hypothesized that *ASXL1* mutations may have high-level epigenomic and transcriptomic signatures across affected cells and tissues.

To examine *ASXL1* expression between BOS and control samples, we performed RT-qPCR using *ASXL1* primers (**Table 2-5**), revealing no significant differences in *ASXL1* expression levels between BOS and control samples in both blood (n=4 BOS, n=4 control) and fibroblast (n=6 BOS, n=6 control) samples.

We also analyzed RNA-seq data to evaluate *ASXL1* transcript levels. In blood samples, the absolute fold change ($|\log_2\text{FoldChange}|$) was 0.11, falling below the threshold for a strong fold change (0.58), with an adjusted p value (padj) of 0.29, failing to meet our significance

criterion of $\text{padj} < 0.05$. Similarly, for fibroblasts, the $|\log_2\text{FoldChange}|$ was -0.01 , and padj was 0.95 .

To explore whether the truncating mutation influenced wild-type ASXL1 protein levels, we meticulously selected a high-quality antibody through testing multiple commercially available options (**Table 2-3**). This selection process involved CACO2 cells, characterized by elevated ASXL1 RNA expression, and HEK293T cells transfected with FLAG-tagged truncated ASXL1 (54). Two antibodies (ab228009, MABE1933) were identified, with ab228009 exhibiting higher ASXL1 expression in CACO2 cells and demonstrating overlap with the ASXL1 construct probed with a high-affinity FLAG antibody.

Western blotting identified no bulk differences in total ASXL1 protein levels between BOS ($n=5$) and control ($n=5$) fibroblasts (**Figure 2-2E**). Western blot bands for ASXL1 were consistently detected at the same molecular weight between BOS and control samples, and truncated proteins were not detected at a reduced molecular weight for any BOS sample tested. This is consistent with other literature examining endogenous, truncated human ASXL1 expression ⁶⁹.

Finally, we extracted histones from fibroblast cells to quantify global changes for histone modification. The identified mutations in BOS lead to the premature termination of ASXL1 before the PHD domain—a critical histone- and DNA-binding domain known for recognizing specific histone modifications, including H3K4me and H3K27me3 (5). It is also implicated in influencing levels of H3K27me3 and H2AK119ub (19). We performed histone immunoblotting, focusing on the specific modifications mentioned, using representative BOS ($n=5$) and control ($n=5$) fibroblast samples. Our analysis revealed no significant changes in the levels of H3K4me3, H3K27me3, or H2AK119Ub between BOS and control groups (**Figure 2-2F**). These findings suggest that the observed ASXL1 mutations in BOS do not result in overt alterations to these key histone modifications in the examined fibroblast cells.

The study aimed to understand if truncating mutations affected *ASXL1* transcript or protein levels, but no significant differences were observed.

Transcriptomic profiling in BOS patient samples.

One of the fundamental roles of *ASXL1* is to regulate the epigenome, activating and repressing the expression of specific genes and pathways during developmental time points. The intricate interplay between epigenetic modifications and transcriptional regulation prompted us to conduct comprehensive RNA-seq analyses on blood and dermal fibroblasts from both BOS patients (n=8 blood, n=7 fibroblasts) and controls (n=11 blood, n=7 fibroblasts). A total of 33 RNA-seq libraries were generated, ensuring a minimum of 30 million reads per sample across different groups.

In line with our qRT-PCR data, normalized *ASXL1* gene expression exhibited no significant differences between BOS (n=8) and control (n=11) blood samples, as well as between BOS (n=7) and control (n=7) fibroblast samples. Examining the relationship between *ASXL1* reads and total sequencing reads revealed a correlation in both blood and fibroblast samples, with fibroblasts exhibiting higher normalized *ASXL1* read counts per sequencing coverage, consistent with the tissue-specific expression pattern.

Further exploration into the correlation between *ASXL1* reads at mutation sites and total *ASXL1* reads demonstrated no significant correlation ($R^2 < 0.01$ for blood samples and $R^2 = 0.10$ for fibroblast samples). Notably, individual variations were observed, such as Pt4 fibroblasts with only 2 reads at the mutation site compared to a total of 3968 *ASXL1* reads, and Pt8 blood with 100 reads at the mutation site compared to 1314 total *ASXL1* reads.

Allelic Distribution in *ASXL1* in BOS patient samples.

To evaluate allelic representation, we assessed the *ASXL1* reference and mutant alleles in patient blood (**Figure 2-3A**, n=8) and fibroblast (**Figure 2-3B**, n=7) RNA-seq data. Calculating the variant allele frequency (VAF) at the mutation site across each BOS sample, most samples

conformed to the expected germline range, between 30% to 70% of the reads at the site (**Figures 2-3A and 2-3B**). This suggests that these late-truncating alleles escape nonsense-mediated decay.

An exception to the expected germline allele distribution was identified in Pt6, who underwent treatment for Wilms tumor before sample collection, displaying 36.9% pathogenic allele in fibroblast RNA-seq, but intriguingly, 100% of reads (27/27) in blood RNA-seq over the *ASXL1* mutation site contained the pathogenic mutation and no reference allele was identified, suggesting loss of heterozygosity. This suggests a potential selective growth advantage during tumorigenesis, treatment, or new clonal hematopoiesis. Pt6, with truncating mutation at amino acid 672, underwent treatment with actinomycin D and vincristine for Wilms tumor years prior to sample collection.

Transcriptomic profiling of Bohring-Opitz syndrome reveals tissue-specific transcriptional dysregulation

To assess the pathogenesis of truncating *ASXL1* mutations, we examined the impact on global transcription across diverse tissues. *ASXL1*, expressed ubiquitously at low levels throughout the body, exhibits higher expression in the skin (32.55 transcripts per million (TPM)) than in whole blood (8.03 TPM). Anticipating a greater number of differentially expressed genes (DEGs) in blood RNA-seq, given its bulk assessment of heterogeneous cell types, we observed substantial gene expression dysregulation in blood, identifying 2118 significant DEGs (**Figure 2-3C**), in contrast to 177 in fibroblast RNA-seq (**Figure 2-3D**). The top 30 DEGs for BOS fibroblast are shown in **Table 2-6**, and the top 20 DEGs for BOS blood are shown in **Table 2-7**.

We identified an outlier in our blood RNA-seq analysis—Pt7 (**Figure 2-3A**), aged 29 at sample collection (**Table 2-1**). Pt7 defies the typical age distribution of BOS, associated with high infant mortality (27%), with most patients not surviving into early adulthood (56). The rarity of BOS and the age profile of available samples constrained our ability to independently assess

Pt7 for age-related effects. Corrective measures included incorporating an age-matched control, Ctrl19 (age 28, **Table 2-2**), integrating analyses across diverse -omics assays with varying control group ages, and assessing significant genes for age-specific bias. Moreover, we examined if Pt7's transcript expression resembled other BOS cases before downstream analysis.

After filtering for fold change ($|\log_2FC| \geq 0.58$), 1097 significant DEGs were identified in blood and 155 in fibroblasts. Both blood and fibroblast RNA-seq analyses revealed a predominance of upregulated DEGs in BOS patients compared to controls—590/1097 DEGs (53.8%) in blood (**Figure 2-3C**) and 125/155 DEGs (80.6%) in fibroblast cells (**Figure 2-3D**). Top DEGs in blood included *GRIK5* (Glutamate receptor, Kainate 5, $\log_2FC=3.8$), *VANGL2* (Vang-like protein 2, $\log_2FC=3.8$), and *GREM2* (Gremlin 2, BMP antagonist, $\log_2FC=-2.5$), while fibroblast DEGs featured *UGT3A2* (UDP Glycosyltransferase Family 3 Member A2, $\log_2FC=4.8$), *VANGL2* ($\log_2FC=2.5$), *GRIK5* ($\log_2FC=2.5$) and *GREM2* ($\log_2FC=-2.1$). Gene ontology analyses unveiled an unexpected commonality in BOS tissues, despite their distinct origins, identifying enrichment in the biological process of regulating neuron projection development ($padj=9.82 \times 10^{-5}$ in blood, $padj=0.02$ in fibroblasts) (**Figures 2-3E and 2-3F**), driven by genes that play key roles in early morphogenesis, particularly neurodevelopmental processes.

Our exploration of gene ontology (GO) in blood and fibroblast RNA-seq not only revealed distinct gene expression alterations but also highlighted tissue-specific changes. Blood DEGs were enriched for hematological processes, encompassing T-cell activation ($padj=3.23 \times 10^{-8}$), neutrophil activation ($padj=1.90 \times 10^{-5}$), axogenesis ($padj=1.90 \times 10^{-5}$), and leukocyte cell-cell adhesion ($padj=3.62 \times 10^{-5}$) (**Figure 2-3E**). In contrast, fibroblast DEGs showed enrichment in structural cell processes, particularly dysregulation of potassium ion transport ($padj=0.004$) and regulation of membrane potential ($padj=0.02$) (**Figure 2-3F**).

Transcriptomic profiling of Bohring-Opitz syndrome reveals tissue-independent transcriptional dysregulation

To examine potential common, cross-tissue effects of truncating *ASXL1* mutations, we performed a correlation analysis on BOS blood RNA-seq DEGs and BOS fibroblast RNA-seq DEGs to unveil shared dysregulated transcripts (**Figure 2-3G**). We identified a core subset of 25 genes dysregulated across both tissue types, with 84% (21/25) of DEGs showing dysregulation in the same direction. This compelling evidence suggests a robust tissue-independent effect of *ASXL1* mutations transcending tissue types. Noteworthy genes within this core subset include *VANGL2*, *GRIK5*, and *GREM2*, emphasizing their pivotal role in the broader impact of *ASXL1* mutations.

Chromatin accessibility dysregulation unveils aberrant activation of developmental pathways

ATAC-seq analysis of BOS (n=6) and control (n=6) fibroblast cells was conducted to examine chromatin accessibility. Chromatin accessibility, identified by increased reads over a genomic region in ATAC-seq, positively correlated with gene expression BOS gene expression profiles, establishing a crucial link between epigenetic modifications and transcriptional dysregulation (**Figure 2-3D**)^{67,70}.

All ATAC-seq libraries displayed the expected distribution of fragment lengths, with an average of 52% of fragments being small (<200 bp length) representing open chromatin regions, and progressively fewer fragments of larger size spanning nucleosomes (>300 bp length). Distinct chromatin landscapes were identified for BOS patient fibroblasts compared to controls. Notably, BOS fibroblasts exhibited a mean insert size of 184bp, indicative of more open chromatin regions, while control fibroblasts had a median insert size of 242bp, reflecting less open chromatin regions. PC1 (47% variance) on PCA and heatmap analysis underscored the clear separation of BOS and control samples, emphasizing the impact of *ASXL1* mutations on chromatin accessibility (**Figure 2-4A**).

A comprehensive examination of differentially accessible regions revealed 4336 significant peaks, with 70.02% (3036/4336) more open and 29.98% (1300/4336) more closed regions in BOS patients compared to controls. These peaks corresponded to 3054 unique genes, with 25.0% (763/3054) represented by more than one significant differentially accessible peak. Key genes, such as *APBA2* (*Amyloid beta precursor protein binding family A member 2*), *KCNQ5* (Potassium voltage-gated channel subfamily Q member 5), *OTUD7A* (OTU deubiquitinase 7A), *PRXL2A* (Peroxisredoxin like 2A) and *PELI2* (Pellino E3 ubiquitin protein ligase family member 2), were among those affected, each represented by multiple peaks. HOMER annotation^{61,71} unveiled top differentially accessible peaks mapping to *FBXL20* ($\log_2FC=-8.3$), *LINGO1* ($\log_2FC=-7.2$), *RUNX3* ($\log_2FC=4.4$) and *CTNNB1* (β -catenin, $\log_2FC=3.8$). The top 20 significantly dysregulated transcriptional start site (TSS) and 5'UTR peaks are listed in **Table 2-8**.

Gene set enrichment analysis identified dysregulation in crucial developmental systems, encompassing muscle development and differentiation ($p_{adj}=9.11 \times 10^{-13}$), skeletal system development ($p_{adj}=1.0 \times 10^{-9}$), regulation of neurogenesis ($p_{adj}=1.35 \times 10^{-9}$), limb morphogenesis ($p_{adj}=1.38 \times 10^{-8}$), renal system development ($p_{adj}=2.14 \times 10^{-8}$), and cardiac septum morphogenesis ($p_{adj}=2.67 \times 10^{-8}$) (**Figure 2-4B**). Motif enrichment analysis using both known motif and *de novo* methods identified significant enrichment of transcription factor binding sites, including JunB ($p_{adj}=1 \times 10^{-208}$, BOS 18.43%, background 5.23%), RUNX1 ($p_{adj}=1 \times 10^{-203}$, BOS 26.45%, background 10.09%) and Fra1 ($p_{adj}=1 \times 10^{-196}$, BOS 17.44%, background 4.95%), emphasizing the core factors driving dysregulation of differentially expressed genes in BOS.

Chromatin accessibility is correlated with the dysregulation of gene expression in BOS.

To unravel the connection between transcriptional dysregulation in BOS and chromatin accessibility, we integrated DEGs identified from fibroblast RNA-seq (BOS n=6, control n=6) with differentially accessible chromatin regions pinpointed through ATAC-seq (BOS n=6, control

n=6). This revealed a positive correlation ($R^2=0.405$) between increased chromatin accessibility and increased gene expression for the same gene (**Figure 2-4C**).

Seventy-one differentially accessible chromatin regions corresponded to 20.9% (37/117) unique DEGs in the fibroblast RNA-seq dataset. For DEGs represented by multiple significantly differentially accessible peaks, we validated that these peaks were concordant in the direction of change. Notably, a subset of these peaks mapped to the promoter-transcriptional start site (TSS) region, typically indicative of more pronounced effects on transcription. Approximately 19.7% (14/71) of the peaks mapped to promoter-TSS regions, including the differentially accessible peak for *VANGL2*. This association accounted for 35.1% (13/37) of unique DEGs in the fibroblast RNA-seq data, underscoring a clear link between increased promoter accessibility and increased transcription.

Further extending our findings across tissue types, a subset of DEGs identified through ATAC-seq exhibited dysregulation in RNA-seq analyses of both BOS fibroblast and blood samples (**Table 2-9**). Seven genes emerged as significant and differentially expressed across multiple omics levels (ATAC-seq and RNA-seq) and tissues (blood and fibroblasts): *VANGL2*, *PRXL2A*, *APBA2*, *OTUD7A*, *KCNQ5*, *PELI2*, and *GREM2*. These genes, pivotal in body patterning, neuron function, neural plate development, and ubiquitination, underscore their multifaceted involvement in the pathogenesis of BOS.

Truncating *ASXL1* mutations dysregulate DNA methylation, and contribute to changes in gene expression.

The negative correlation between endogenous CpG methylation levels at promoter regions and gene expression is well-established (**Figure 2-2D**)⁷⁰. Increased DNA CpG methylation or more closed chromatin correlates with diminished gene expression. To generate a BOS-specific DNA methylation (DNAm) signature, we conducted genome-wide DNAm

profiling in blood from BOS individuals (n=13) compared to sex- and age-matched control subjects (n=26) (**Table 2-1** and **Table 2-2**)⁵².

DNA_m analysis unveiled 8596 differentially methylated CpG sites (FDR<0.05), associated with ENSEMBL genes, with 5773 CpG sites overlapping genes and mapping to 3803 unique genes. Employing linear regression modeling, 763 differentially methylated CpG sites surpassed a threshold of $|\Delta\beta|>0.10$ (10% DNA_m difference). These sites corresponded to 163 unique genes, 71 of which (43.6%) were represented by more than one differentially methylated CpG site. After filtering for effect size ($\text{abs}(\log_2\text{FC})>0.58$), 24 unique genes remained, encompassing 50 differentially methylated CpG sites, including Proteasome 20S subunit alpha 8 (*PSMA8*) represented by 8 sites, Ellis van Creveld (*EVC*) represented by 6 sites, and *LRP5* represented by 3 sites.

Examining the potential downstream effects of statistically significant differential methylation on transcription, we correlated BOS blood DNA_m with RNA-seq expression profiles. This identified 672 differential CpG sites that mapped to 341 unique DEGs in the blood RNA-seq dataset, and we observed 143 genes represented by more than one differential CpG site. Genes represented by multiple CpG sites were analyzed for concordance across the same gene and analyzed in correlation with gene expression values for each sample. We filtered for highly differentially methylated sites ($|\Delta\beta|>5\%$) and highly differentially expressed transcripts ($|\log_2\text{FoldChange}(\log_2\text{FC})|>1.5$). After filtering, we retained 50/672 CpG sites (7.44%), corresponding to 24/341 unique genes (7.04%) (**Figure 2-5A**). 11 genes were represented by >1 CpG site, including *PSMA8* with 8 CpG sites and *GRIK5* with 2 CpG sites. *VANGL2* was represented by 1 CpG site. The correlation between these DNA methylation differences and transcriptomic changes indicates a steady-state effect of dysregulated DNA methylation on gene expression.

Three differentially methylated CpG sites within the transcriptional start site (TSS) were identified for *EVC* and *EVC2*. Both genes were hypomethylated at the TSS (-5% and -3.5%,

respectively) and transcriptionally upregulated ($\log_2FC=1.50$ and 1.54 , respectively) in BOS patient blood.

We also generated BOS DNAm data using BOS fibroblasts ($n=8$) and control fibroblasts ($n=8$). We identified 444 CpG sites with a nominal p-value of 0.005 and $|\Delta\beta|>0.10$. Despite not meeting $FDR<0.05$ criteria for significance, the fibroblast DNAm analysis displayed lower significance and effect levels compared to blood, this aligned with our RNA-seq findings when comparing across tissues. The inherent homogeneity in fibroblast cell culture, as opposed to the heterogeneity of primary tissue like blood, likely contributed to this observation. Furthermore, the reduced number of DEGs identified in fibroblast RNA-seq compared to blood RNA-seq may have resulted in no significant correlations between fibroblast DNAm and RNA-seq.

Gene set enrichment analysis of DNA methylation and RNA sequencing reveals activation of canonical Wnt pathway signaling in BOS samples.

We observed concordant findings in the DNAm and RNA-seq datasets which prompted us to conduct gene ontology analysis for common gene targets using two complementary methods, GREAT⁵³ and clusterProfiler v3.12.0⁶⁰. We applied a significance threshold of $p_{adj}<0.05$. Both analyses unveiled significant gene set enrichments in the Wnt signaling pathway (GO#0060070, 17 genes, $p_{adj}=0.045$), anterior-posterior pattern specification process (GO#0007389, 19 genes, $p_{adj}=0.045$), and regulation of neuron projection development (GO#0010975, 19 genes, $p_{adj}=0.045$), along with other biologically relevant pathways (**Figure 2-5B, Table 2-10**). Key genes enriched in these pathways include *PSMA8*, *WNT7A*, *FZD3*, *LRP5* and *LRP6* in canonical Wnt signaling and body pattern specification. *VANGL2* strongly contributed to the enrichment in the latter pathway. The identification of overlapping gene targets and pathways between the methylome and transcriptome in BOS patients suggests robust, coordinated, epigenetically-driven effects of truncating *ASXL1* mutations.

PSMA8 emerged as a pivotal driver gene in the canonical Wnt-signaling enrichment identified in blood DNA methylation and RNA-seq data. We observed significant hypermethylation of *PSMA8* across all 8 CpG sites within the promoter and TSS region, with $\Delta\beta$ ranging from 6.1% to 18.9% (**Figure 2-5C**). Transcriptionally, *PSMA8* was downregulated in BOS patient blood ($\log_2FC=-2.92$, **Figure 2-5D**). Interestingly, *PSMA8* showed no expression in BOS (n=7) or control (n=7) fibroblasts, with normalized read counts of 0 in all samples and a read count of 1 in one BOS fibroblast sample.

Truncating *ASXL1* mutations disrupt canonical and non-canonical Wnt signaling.

The integration of various datasets unveiled dysregulation of the canonical Wnt pathway (**Figure 2-6A**), as evidenced by gene set enrichment analysis of RNA-seq (**Figure 2-3E**) and DNA methylation studies (**Figure 2-5B**). This dysregulation was characterized by a consistent pattern of upregulation in the RNA-seq differentially expressed genes (DEGs) associated with the canonical Wnt/ β -catenin pathway. Specifically, we observed the upregulation of Wnt ligands: *WNT1* ($\log_2FC=0.87$), *WNT7A* ($\log_2FC=0.97$), and *WNT10B* ($\log_2FC=0.53$) and upregulation of Wnt ligand receptor, *Frizzled* receptor *FZD3* ($\log_2FC=0.87$).

Western blot quantification of representative BOS fibroblasts (n=5) and control (n=5) fibroblasts did not reveal significant differences in two key proteins regulated through spatial distribution, VANGL2 and β -Catenin, using bulk whole-cell lysate (**Figure 2-6B and 2-6C**). However, the analysis using a commercial Wnt antibody kit demonstrated a fold increase of 1.5x for AXIN1, 2.8x for AXIN2, 3.5x for DVL2, and 1.5x for DVL3 averaged across BOS patient samples compared to controls (**Figure 2-6C**). Notably, while the levels of ASXL1, histone marks, VANGL2, β -Catenin, and β -actin were consistent between BOS samples, Pt3 and Pt4 exhibited control-like levels of AXIN2, DVL2, and, to some extent, DVL3.

RNA-seq and DNA methylation studies further identified aberrant upregulation and hypomethylation of Wnt signaling co-receptors *LRP5* (**Figure 2-6D to 2-6F**) and *LRP6* (**Figure**

2-6G to 2-6I). We observed significant transcriptional upregulation of *LRP5* ($\log_2FC=1.64$, $p_{adj}=3.58 \times 10^{-9}$, **Figure 2-6D and 2-6E**) and *LRP6* ($\log_2FC=1.63$, $p_{adj}=1.17 \times 10^{-12}$, **Figure 2-6G and 2-6H**), 4 hypomethylated DNAm sites for *LRP5* ($\Delta\beta$ -3.5% to -8.0%, $FDR<0.05$) (**Figure 2-6F**), 2 of which reside in a CpG island, and 2 hypomethylated DNAm sites for *LRP6* ($\Delta\beta$ -2.7% to -4.0%, $FDR<0.05$), both of which reside in a CpG island (**Figure 2-6I**).

Moreover, we identified transcriptional upregulation of *AXIN2* ($\log_2FC=0.95$) and downregulation of *GSK3 β* ($\log_2FC=-0.28$), consistent with canonical Wnt signaling activation. *AXIN2*, a key mediator, is directly induced by Wnt pathway activation, and its expression serves as a proxy reporter for Wnt pathway signaling (62, 63). Additionally, in BOS patient samples, the Wnt pathway transcription factors, *TCF7* ($\log_2FC=0.81$) and *LEF1* ($\log_2FC=0.86$), were significantly transcriptionally upregulated.

Altered epigenetic profile in *VANGL2*, a non-canonical Wnt signaling gene, in BOS.

VANGL2 emerged as a prominently dysregulated gene in BOS, displaying significant overexpression in both blood and fibroblast RNA-seq datasets (**Figure 2-3G**). The integration of RNA-seq, ATAC-seq, and DNAm data consistently identified *VANGL2* as differentially regulated across all datasets (**Figure 2-4C and 2-5A**).

In BOS blood samples, *VANGL2* exhibited transcriptional upregulation accompanied by hypomethylation ($\Delta\beta$ = -7.6%) at the TSS of *VANGL2* CpG site cg17024258 (**Figure 2-6J**) and an inverse relationship between hypermethylation and transcriptional upregulation in individual samples. The 5' UTR of *VANGL2* displayed a 2.3-fold increase in chromatin accessibility as observed in ATAC-seq data (**Figure 2-6K**).

These findings were supported by differential gene expression of *VANGL2* in BOS patients compared to control samples. Cross-tissue RNA-seq integration (**Figure 2-3G**) showed that *VANGL2* was significantly overexpressed in both BOS patient blood ($\log_2FC=3.8$, **Figure 2-6L**) and in fibroblasts ($\log_2FC=2.55$, **Figure 2-6M**).

Additionally, the visualization of ATAC-seq data in IGV revealed a clear increase in reads for BOS patients at the *VANGL2* TSS, indicating more open chromatin (**Figure 2-6N**). CUT&RUN for H3K4me3, a transcription-activating histone mark, further supported these findings, showing significantly stronger H3K4me3 binding at the *VANGL2* TSS in BOS compared to controls (**Figure 2-6O**). Notably, *ASXL1*'s role in histone modifications, particularly H3K4me3, H3K27me3, and H2AK119ub, has been established, and our data further corroborates that truncating mutations in *ASXL1* result in the loss of its gene-repressive functions.

Discussion

Comprehensive multi-omics studies establish a direct link between *ASXL1* mutations and the Wnt signaling pathway.

We employed a comprehensive multi-omics approach to profile primary samples from BOS patients. This study marks the first direct link of *ASXL1* mutations to the activation of both canonical and the non-canonical Wnt-signaling pathway, specifically the planar cell polarity (PCP) pathway (**Figure 2-7**). These pathways are essential in developmental patterning, with the PCP pathway playing a key role in defining cell polarity and regulating morphogenesis and diverse cellular processes⁷². The data demonstrate that upregulation of Wnt pathways in BOS samples can be attributed to disruption of the epigenetic landscape, ultimately resulting in transcriptomic dysregulation.

In contrast to most studies focusing on human *ASXL1* function in the context of myeloid leukemia, which propose haploinsufficiency^{10,35}, dominant negative, or gain-of-function effects^{30,36,37}, our findings present a unique perspective. Previous reports on *in vivo* deletion of *ASXL1* showed mild phenotypes in myeloid and lymphoid cell frequencies⁷, while other studies with *ASXL1* knockout mice demonstrated systemic developmental defects and MDS-like presentation, suggesting a potential loss-of-function effect in hematopoietic progenitor cells^{10,35}.

However, it is essential to note that the majority of *ASXL1* mutations identified in human myeloid malignancies and BOS do not involve a deletion but rather harbor C-terminal nonsense or frameshift mutations. These mutations are hypothesized to escape nonsense-mediated decay and may potentially exert a dominant-negative or gain-of-function effect ³⁰.

Our findings, revealing comparable *ASXL1* transcript and protein expression between BOS and control samples (**Figures 2-2D, 2-3A, and 2-3B**) are consistent with this hypothesis and show no evidence of degradation or nonsense mediated decay. The dysregulation observed in BOS may be a result of the combination of haploinsufficiency of the full-length transcript and translation of a short and non-functional protein, leading to the mistargeting of ASXL1 functions. Furthermore, the developmental context may also play a role, a hypothesis supported by iPSC studies conducted by Matheus et al. ⁸.

Additionally, our findings of comparable levels of wild-type and pathogenic transcripts in BOS patients add further evidence to the existing literature, refuting the role of nonsense-mediated decay and haploinsufficiency in BOS pathogenesis ^{36,73}. Instead, the pathogenesis appears to be mediated through mis-expression and impact, potentially through gain of function, of a truncated ASXL1 protein.

These findings highlight the discovery of the direct linkage between *ASXL1* mutations and Wnt-signaling pathways, challenges prevailing theories regarding *ASXL1* mutations, and provides crucial insights into the complex mechanisms underlying BOS pathogenesis.

High-penetrance mutations in severe monogenic disorders supersede effects of age, sex and genetic background in molecular assays.

Despite the well-established tissue-specific gene expression ⁷⁴, we propose that the high-penetrance effect of truncating *ASXL1* mutations leads to a set of dysregulated genes across tissues. The identification of 25 common DEGs in both BOS blood and fibroblast data supports the overarching hypothesis that truncating *ASXL1* mutations exert a strong

tissue-independent effect across tissues expressing *ASXL1*. The functions of these key genes likely contribute to BOS pathophysiology, suggesting that targeting these genes could provide a systemic treatment for BOS.

Addressing the inherent challenges in studying rare diseases, including heterogeneity in ages at sample collection and diverse genetic backgrounds of patients, is crucial. While controlling for age-specific effects and genetic background proves challenging with fewer than 100 reported patients worldwide, our study took steps to mitigate these challenges. DNA methylation controls were matched for age and sex effects, while ATAC-seq and RNA-seq controls were matched for sex and genetic background effects. The integration of ATAC-seq, RNA-seq, and DNA methylation data, along with the identification of common dysregulated genes, suggests that the impact of truncating *ASXL1* mutations supersedes the effects of age and genetic background.

Integration of ATAC-seq and RNA-seq identified seven genes that were commonly dysregulated (**Table 2-4**). We tested these genes specifically for age-specific effects since that was a variable we were unable to sufficiently match given the young age of the BOS patient cohort, and restricted access to samples from healthy young children who do not require procedures. In our ATAC-seq and RNA-seq assays, we primarily used related parental controls to account for genetic variation in our population. The seven genes that were commonly dysregulated in our ATAC-seq and RNA-seq integration (**Table 2-4**), were all tested as part of a larger study by Lee et al.⁷⁵ through transcriptome analysis to classify human genes based on age-specific differential expression analysis. None of these genes showed an age-specific signature. These seven genes have important functions in body patterning and organogenesis, nervous system development and ubiquitination regulation. While this does not rule out age-specific differences in potential interacting proteins, it suggests that truncating *ASXL1* effects are not significantly impacted by age.

ASXL1 represses canonical and non-canonical Wnt signaling in normal development

A pivotal finding of this study is the association between *ASXL1* mutations and the activation of both canonical and non-canonical Wnt-signaling pathways. The canonical Wnt signaling pathway plays an important role in embryonic development, stem cell maintenance, and differentiation of cells in adults and is highly conserved across species ⁷⁶. The Wnt signaling pathway is under complex regulation, and functions in context-specific manners dependent on the receptors present on the cell membrane and Wnt ligand-receptor interactions during pathway activation ⁷⁷.

The molecular mechanisms and specific developmental stages of the interactions between the canonical and non-canonical Wnt pathways are not well elucidated. In the absence of Wnt signaling, β -catenin undergoes degradation by the 26S proteasome, comprising two subcomplexes, the 20S proteasome and the regulatory particle ^{78,79} (**Figure 2-6A**). We observed hypermethylation and transcriptional downregulation of *PSMA8*, a subunit of the 20S proteasome, in BOS patient samples (**Figure 2-5C and 2-5D**). This suggests downregulation of the 26S proteasome-mediated degradation of β -catenin in BOS patient tissues ⁷⁸. While *PSMA8* is involved in numerous cellular pathways, considering the context of other differentially regulated Wnt signaling pathway genes identified, the hypermethylation and transcriptional downregulation of *PSMA8* aligns with our findings of aberrant Wnt activation in RNA-seq (**Figure 2-3E**) and DNA methylation studies (**Figure 2-5B**) of BOS.

The “ β -catenin destruction complex” comprising AXIN1, AXIN2, Dishevelled1/2/3 (DVL1, DVL2, DVL3), GSK3 β and Adenomatous polyposis coli (APC), plays a crucial role in Wnt signaling. BOS patient fibroblast samples exhibited increased protein levels of AXIN1, AXIN2, DVL2, and DVL3. Notably, Pt3 and Pt4, who harbor truncating mutations within close proximity; at amino acid 635 and 642 respectively, exhibited control-like levels of AXIN2, DVL2 and, to some extent, DVL3. This hints at a location specific effect of *ASXL1* mutations, necessitating further investigation. The heterogeneity of BOS, influenced by patient specific factors such as

sex and age, or environmental effects, was evident in our analysis. Fibroblasts derived from female BOS patients (Pt2, Pt14 and Pt15) exhibited much higher levels of Wnt pathway proteins than fibroblasts from male patients (Pt3 and Pt4). No significant difference was identified between male and female control samples, suggesting a potential sex-linked differential effect of truncating *ASXL1* mutations on Wnt pathway protein expression, warranting further study and replication with an increased number of patient samples.

Contrary to expectations, Western blot analysis did not reveal differences in total whole cell β -catenin or VANGL2 levels between BOS and control fibroblasts (**Figure 2-6B and 2-6C**). Given that both genes are both post-translationally regulated, further studies are required to unveil their role in driving phenotypic effects. Investigating the cellular localization of canonical and non-canonical Wnt proteins is essential to understanding their role in BOS pathogenesis.

Notably, the nuclear translocation of β -catenin is important in identifying Wnt signaling pathway activation. In the Wnt-active state, the “ β -catenin destruction complex” is dissociated and leads to nuclear translocation of β -catenin ^{42–44} (**Figure 2-6A**). In the nucleus, β -catenin binds to the TCF and LEF families, initiating transcription of Wnt pathway targets ^{80,81}. *TCF7*, upregulated in BOS samples, is an important regulator of self-renewal and differentiation in multipotential hematopoietic cell lines ⁸², potentially shedding light on the pathogenesis of myeloid cancers with truncating *ASXL1* mutations.

In contrast, the non-canonical Wnt pathways, including the PCP pathway which defines cell polarity and migration, exhibit β -catenin-independent regulation of crucial events during embryonic development ⁸³. However, experiments suggesting binary utilization of these pathways may be model-system and context-specific ⁴⁷. Given the complex interplay and cross-talk at various levels of the Wnt signaling pathways, we endorse a more integrated view, rather than the outdated binary classification ⁴⁷.

Dysregulation of the Wnt signaling pathway identified in BOS samples may explain phenotypic presentations in BOS patients.

The identified dysregulation of the Wnt signaling pathway in BOS samples provides a potential explanation for the observed phenotypic presentations in BOS patients. Wnt signaling plays a key role in hair growth and the development of hair follicles. Notably, *WNT10B*, transcriptionally upregulated in BOS samples, promotes differentiation of primary skin epithelial cells towards the hair shaft in mice and elongation of the hair shaft in isolated rabbit whisker hair follicles^{84–86}. Intriguingly, clinicians and parents have frequently noted excessive hypertrichosis with rapidly growing hair and nails in BOS patients²⁴. This suggests the possibility that aberrant activation of Wnt signaling may play a crucial role in the pathophysiology of BOS patients.

Several other significant Wnt signaling DEGs in our datasets further support the intricate role of Wnt signaling in BOS. *WNT1*, known for its importance in regulating cell fate and patterning during embryogenesis, and *WNT7A*, a key gene in embryonic dorsal vs ventral patterning gene, are among these genes. *WNT7A* also serves as a key regulator of normal neural stem cell renewal, proliferation, and differentiation. The activation of the PCP pathway *WNT7A*, coupled with its overexpression inducing *Vangl2* overexpression and impairing neurulation in mouse neural stem cells through aberrant Vangl2 polarized distribution^{87,88}, adds a layer of complexity to the role of Wnt signaling in BOS pathophysiology.

The comprehensive integration of our epigenomic, transcriptional and protein data regarding the Wnt signaling pathways consistently indicates a distinct activation of canonical Wnt signaling in BOS samples. These findings not only provide valuable insights into the molecular mechanisms underlying BOS but also offer a potential connection to the observed phenotypes in patients.

Upregulation of *VANGL2* in BOS drives clinical phenotypes.

The profound dysregulation of *VANGL2* was notable in all our datasets, with a remarkable 9-fold increase in gene expression driven by epigenetic dysregulation (**Figures 2-6J to 2-6O**). *VANGL2*, known for its role in regulating Wnt protein distribution ⁸⁹, also plays non-canonical roles in regulating various developmental processes, including neurulation, cardiac development, kidney-branching morphogenesis and the regulation of hematopoiesis ^{48,90–92}. Although we identified no significant difference in total *VANGL2* protein at the whole cell level between BOS and control fibroblasts (**Figures 2-6B, 2-6C**), the functional impact of *VANGL2* is strongly dependent upon cellular localization. The asymmetric localization of *VANGL2* on the plasma membrane is needed for signal transduction, polarization and cell organization ⁹³. The PCP pathway is also crucial for neural tube closure ⁹⁴ and studies suggest that a gradient of Wnt activity helps establish *VANGL2* polarity in the neural plate during neurulation ⁹⁵, with canonical Wnt signaling required for neural crest induction, and non-canonical Wnt pathway required for neural crest migration ⁹⁶.

Studies involving knockout and loss-of-function of *Vangl2* identified significant reduction in spine density and dendritic branching in primary culture rat hippocampal neurons ⁴⁸. In humans, homozygous mutations in *VANGL2* result in embryonic lethality and cause craniorachischisis - a severe neural tube defect encompassing anencephaly and bony defects of the spine - in mice ⁹⁷. Heterozygous *VANGL2* mutations are embryonic lethal and have been identified in miscarried human fetuses with severe cranial neural-tube defects ⁹⁸. Considering the essential role of neural crest cells in nervous system and craniofacial development, along with their critical contribution to cardiac outflow septation and alignment ⁹⁹, it is noteworthy that *Vangl2* knockout mice exhibit cardiac outflow tract malformations and septal defects ⁹⁰, mirroring findings in BOS patients ²⁴. These observations suggest that dysregulation of *VANGL2*, with downstream effects on neural crest migration, may contribute to the severe neural phenotype, distinctive craniofacial features, and cardiac defects observed in BOS patients ²⁴.

While the interplay between canonical Wnt and non-canonical Wnt pathways is not well understood, recent literature suggests reciprocal pathway inhibition through competition for a common downstream coreceptor *FZD*^{83,100}, most evident during tissue regeneration and development^{101,102}. We identified upregulation of *FZD3* in our data. The asymmetric localization of *FZD3*, required for proper function, is dependent on anchoring by *VANGL2* through physical interaction of the two proteins¹⁰³. *FZD3* regulates the establishment of the non-canonical Wnt/PCP pathway, playing a role neural crest cell migration and neural tube closure^{104,105}, findings that are supported by molecular characterization of truncating *ASXL1* mutations conducted in iPSCs⁸.

Limitations and future directions.

A limitation of our study is its focus on differentiated cell types, whereas BOS pathophysiology is inherently shaped by cellular context and developmental stage. The molecular effects of epigenetic modifiers like *ASXL1* are influenced by cellular and developmental context, and the effects of *ASXL1* are known to begin in early embryogenesis¹⁰⁶. To capture cell types where *ASXL1* maximally impacts clinical phenotypes, embryonic stem cells would represent an ideal model. In future studies, multi-omics approaches would benefit from being performed in patient-derived stem-cell models and through samples generated with introduction of truncating *ASXL1* mutations using DNA- editing systems. This shift towards stem-cell models holds promise for gaining additional insights into the role of *ASXL1* in stem cell homeostasis and differentiation.

Conclusion

In conclusion, we present the largest multi-omics analysis of patients with Bohring-Opitz syndrome (BOS), a rare genetic disorder. Conducting the first analyses of RNA-seq and ATAC-seq on primary BOS patient samples, we contribute to the growing literature on *ASXL1* and the pathogenesis of *ASXL1* truncating mutations. Our integrated methods reveal the

dysregulation of key transcripts such as *VANGL2*, through direct chromatin modifications, suggesting that physiological levels of truncating *ASXL1* mutations lead to the loss of the repressive function of *ASXL1*. Importantly, pathways dysregulated across multiple assays are involved in neural development, shedding light on the severe intellectual and neural features in BOS patients. The dysregulation of Wnt/ β -catenin and non-canonical Wnt PCP pathways emerges as a critical finding, representing key, drug-targetable pathways. Our results hold major implications for identifying potential therapeutics for diseases harboring *ASXL1* mutations.

Figures and Tables

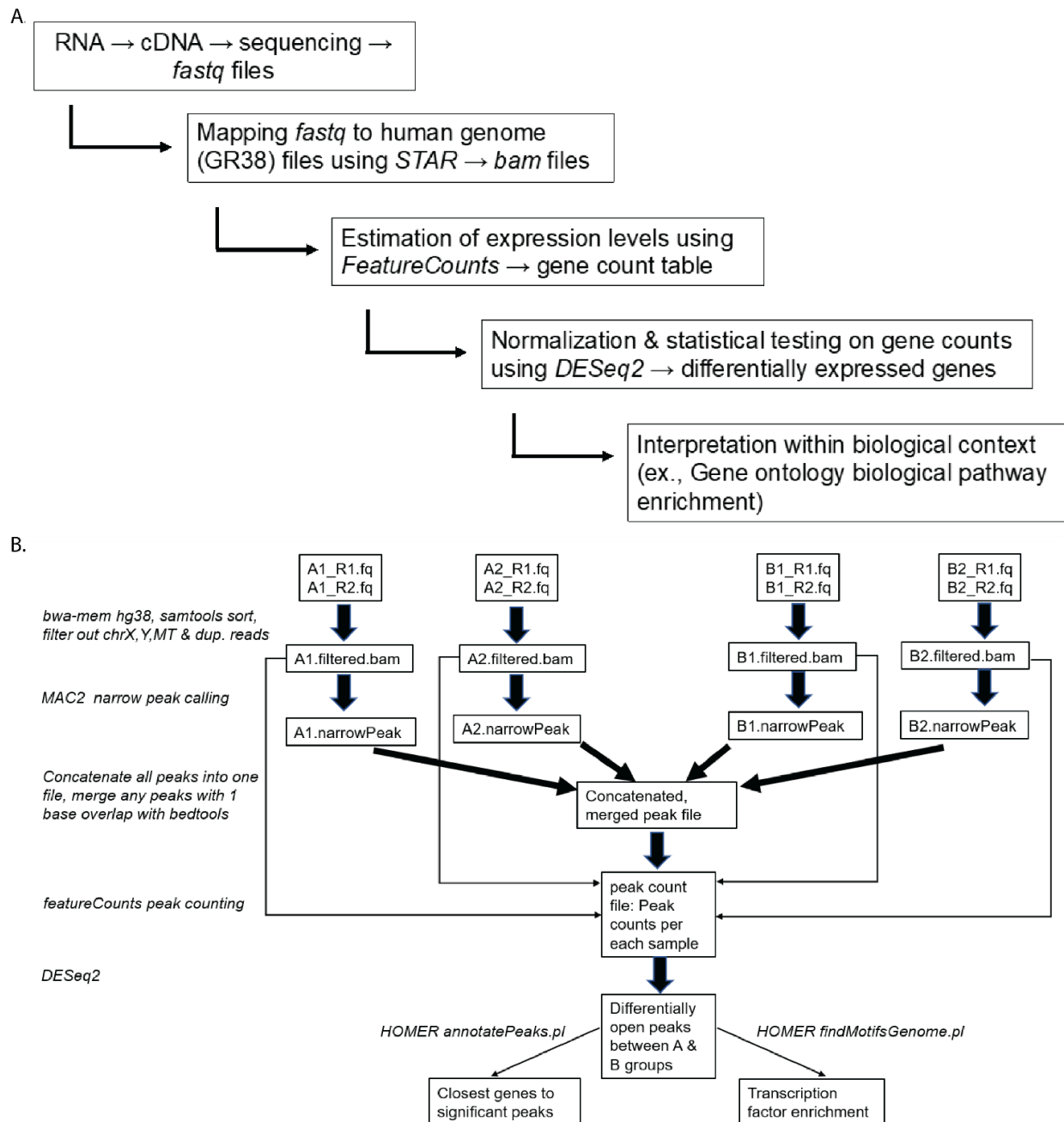


Figure 2-1: Arboleda lab RNA-seq and ATAC-seq analysis pipelines

(A) Arboleda Lab RNA-seq analysis pipeline.

(B) Arboleda Lab ATAC-seq analysis pipeline.

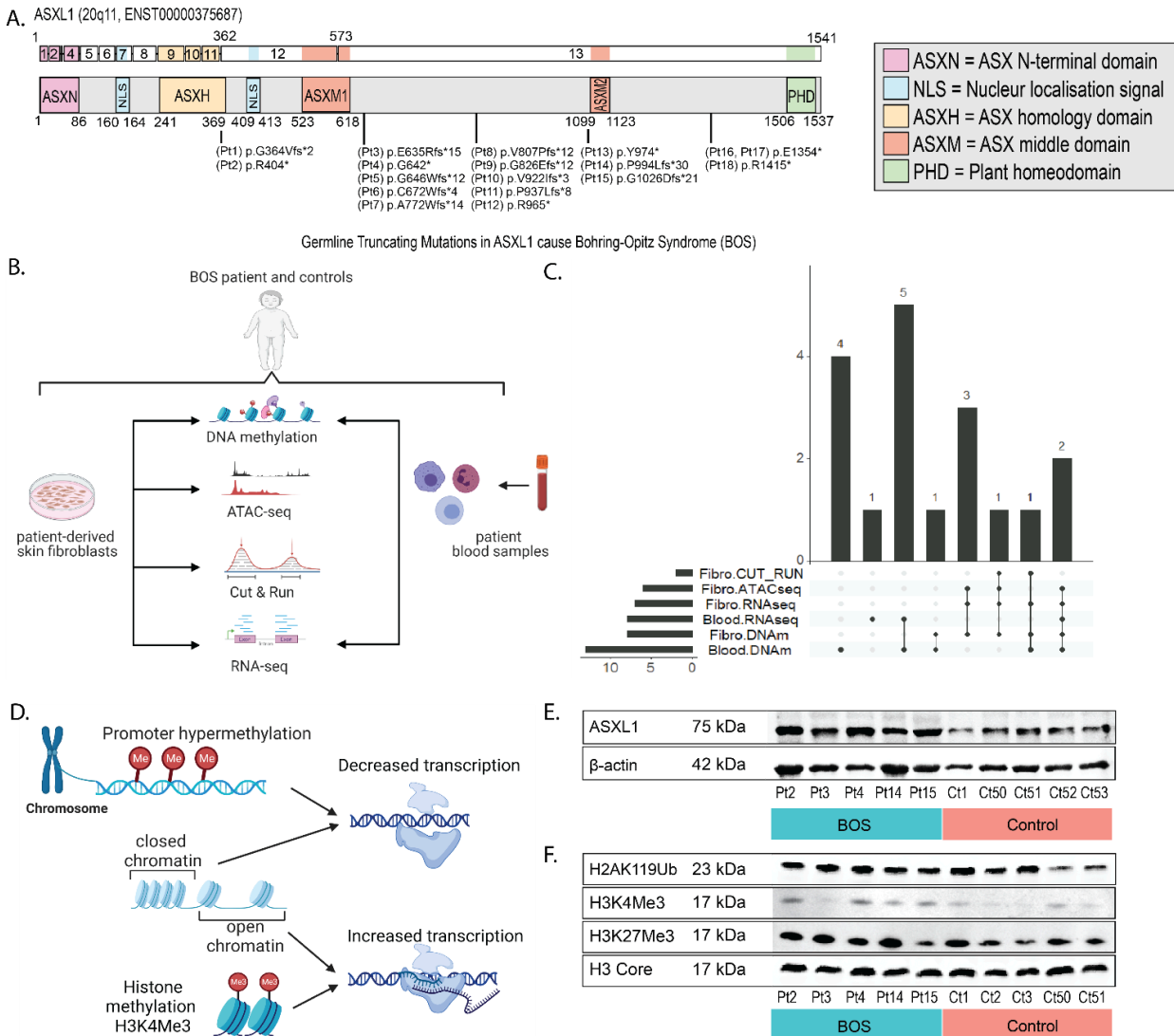


Figure 2-2: Multiomics study design for Bohring-Opitz syndrome (BOS) caused by pathogenic mutations in ASXL1

(A) Schematic representation of the *ASXL1* transcript (ENST00000375687.10) and protein (GenBank: NM_015338.6; GRCh37), depicting its functional domains and mutations causing BOS. Mutations listed correspond to patients in this study, tagged with a Pt identifier.

(B) Peripheral blood and dermal fibroblasts underwent epigenomic assays for ATAC-seq, CUT&RUN, DNA methylation, and global transcriptome analysis using RNA-seq.

(C) Across multiomics assays and two specimen types, we had 8 of 18 BOS samples with fibroblast assays and 14 of 18 with blood assays. Four of 18 BOS patients had data from assays across both specimen types.

(D) Promoter hypermethylation and closed chromatin, examined through DNA methylation and ATAC-seq analysis, respectively, are associated with decreased transcription, while activating histone methylation, such as H3K4Me3 at promoters and open chromatin, are associated with increased transcription.

(E) Western blot for representative BOS (n = 5) and representative control (n = 5) fibroblast whole-cell lysate extracts shows no significant difference in total ASXL1 protein. This experiment was repeated three times.

(F) Western blot for representative BOS (n = 5) and representative control (n = 5) fibroblast histone extracts reveals no significant difference in H2AK119ub, H3K4me3, and H3K27me3. This experiment was repeated three times.

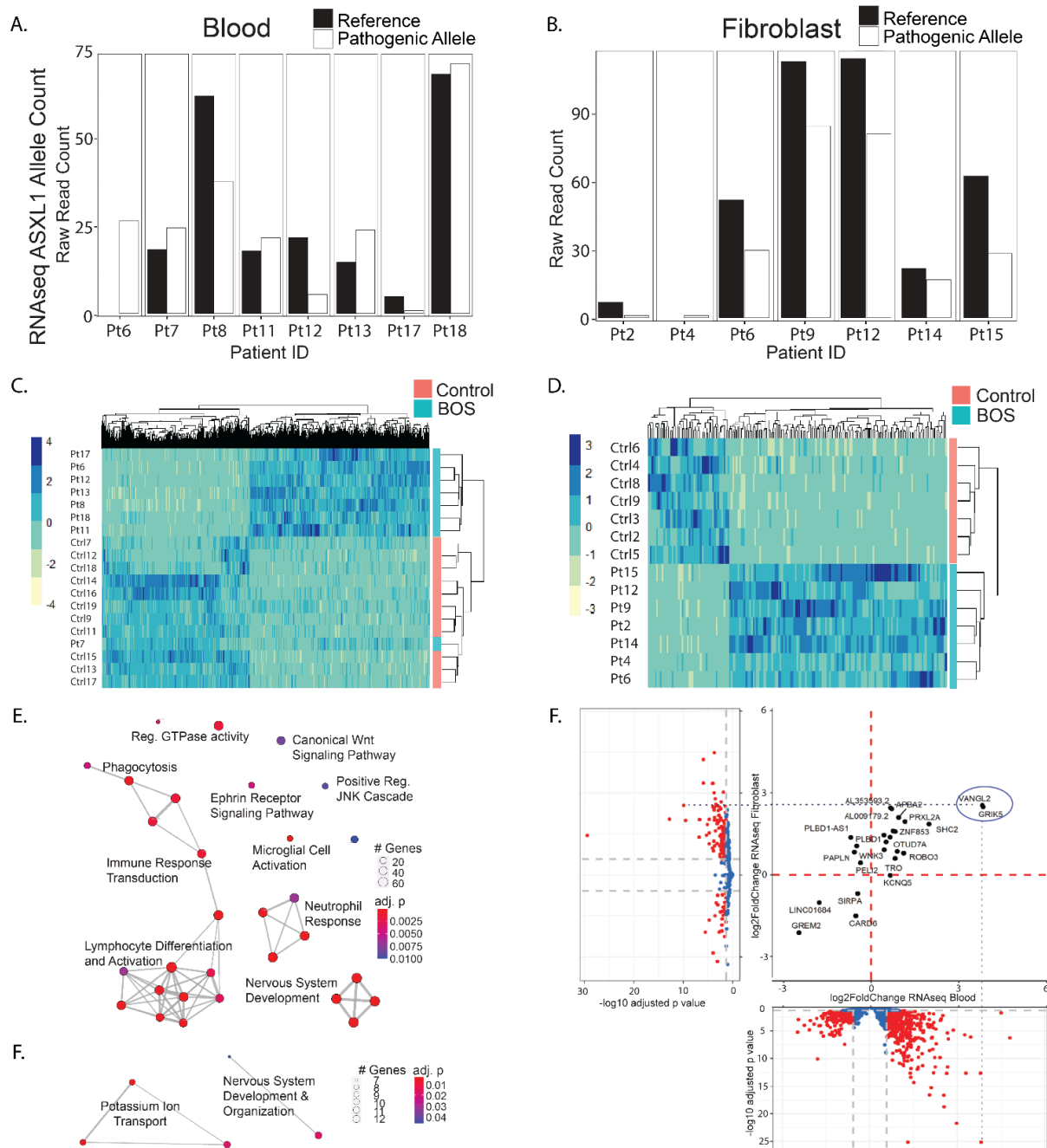


Figure 2-3: Pathogenic mutations in *ASXL1* induce both tissue-specific and tissue-independent effects on gene expression

(A) and **(B)**: RNA-seq raw read counts illustrate the *ASXL1* reference allele (black) and pathogenic allele (white, outlined in black) at mutation sites in blood (n=8) and fibroblast (n=7) samples, providing insights into allele-specific expression patterns.

(C) and (D): RNA-seq heatmaps showcase all significant Differentially Expressed Genes (DEGs) with $P_{adj} < 0.05$ and $abs(log_2FC) \geq 0.58$. In blood, 1097 DEGs are identified (blue for BOS, n=8; pink for control, n=11), with 53.8% being more upregulated. In fibroblasts, 155 DEGs are highlighted (BOS, n=7; control, n=7), with 80.6% being more upregulated in BOS patients.

(E) and (F): Gene ontology analyses reveal enriched biological processes in blood and fibroblast RNA-seq. Blood samples exhibit genes related to nervous system development and the canonical Wnt signaling pathway, while fibroblast samples highlight genes associated with nervous system development.

(G): Volcano plots for BOS vs. control RNA-seq in blood (x-axis) and fibroblast (y-axis) identify a core subset of 25 shared dysregulated transcripts. Remarkably, 21 of these 25 DEGs (84%) display dysregulation in the same direction, emphasizing a strong tissue-independent effect of *ASXL1* mutations on gene expression.

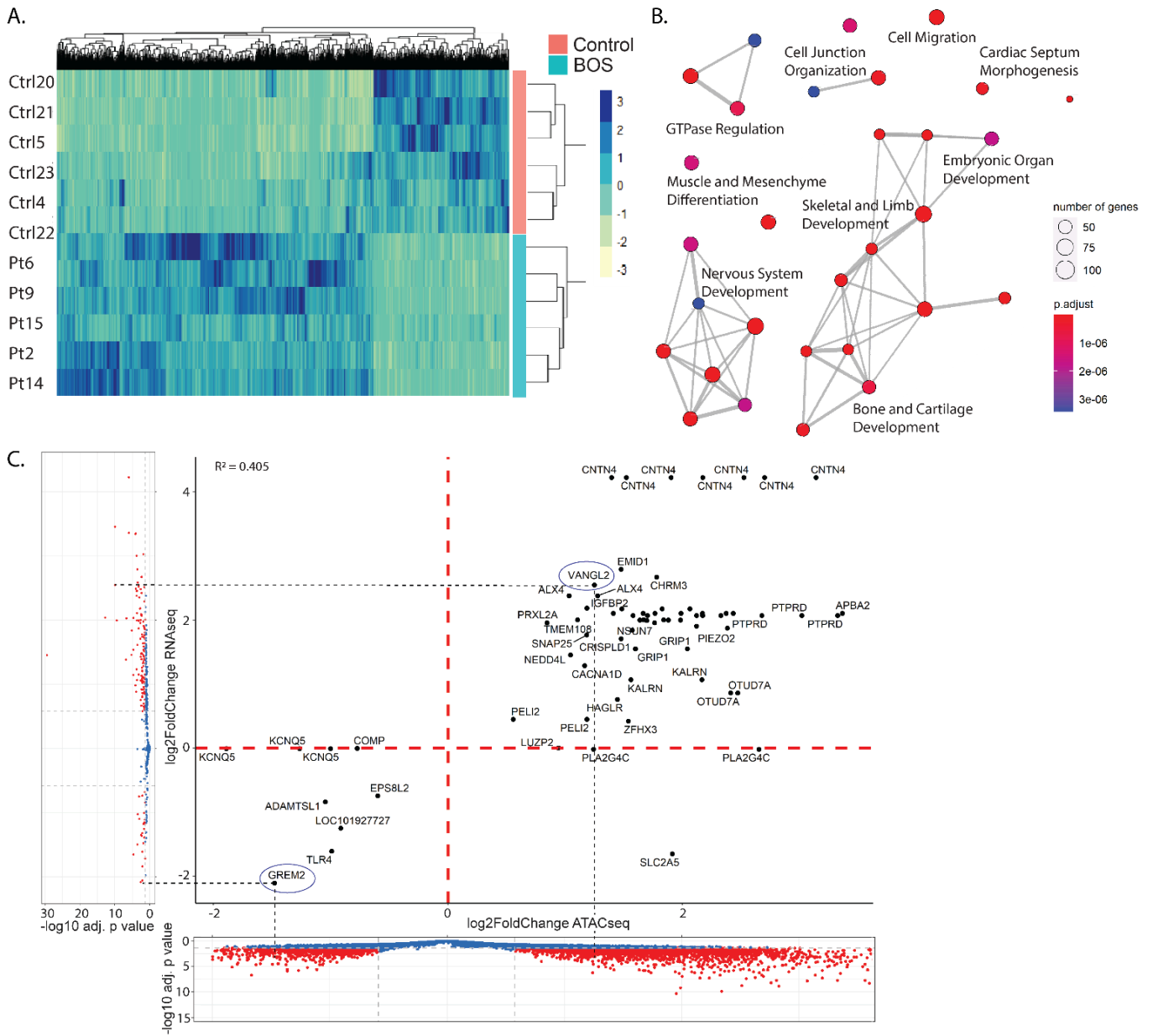


Figure 2-4: Epigenetic dysregulation in fibroblasts from BOS patients precipitates transcriptomic dysregulation.

(A) ATAC-seq heatmap showcases significant DEGs with $p_{\text{adj}} < 0.05$ and $\text{abs}(\log_2\text{FC}) \geq 0.58$ in fibroblasts. In BOS (blue, $n=7$) vs. control (pink, $n=7$), 4336 DEGs are identified, with 70.0% being more upregulated.

(B) Gene set enrichment unveils pivotal dysregulated pathways, notably including nervous system development.

(C) Integration of chromatin accessibility (ATAC-seq, x-axis) and gene expression (RNA-seq, y-axis) in BOS (n=7) compared to control fibroblasts (n=7) identified a positive correlation ($R^2 = 0.405$) between chromatin accessibility and gene expression. A subset of 37 common dysregulated transcripts is highlighted (right), signifying their shared epigenetic and transcriptomic alterations. DEGs deemed significant (red) met the criteria of $\text{abs}(\log_2\text{FC}) \geq 0.58$ and $p_{\text{adj}} > 0.05$.

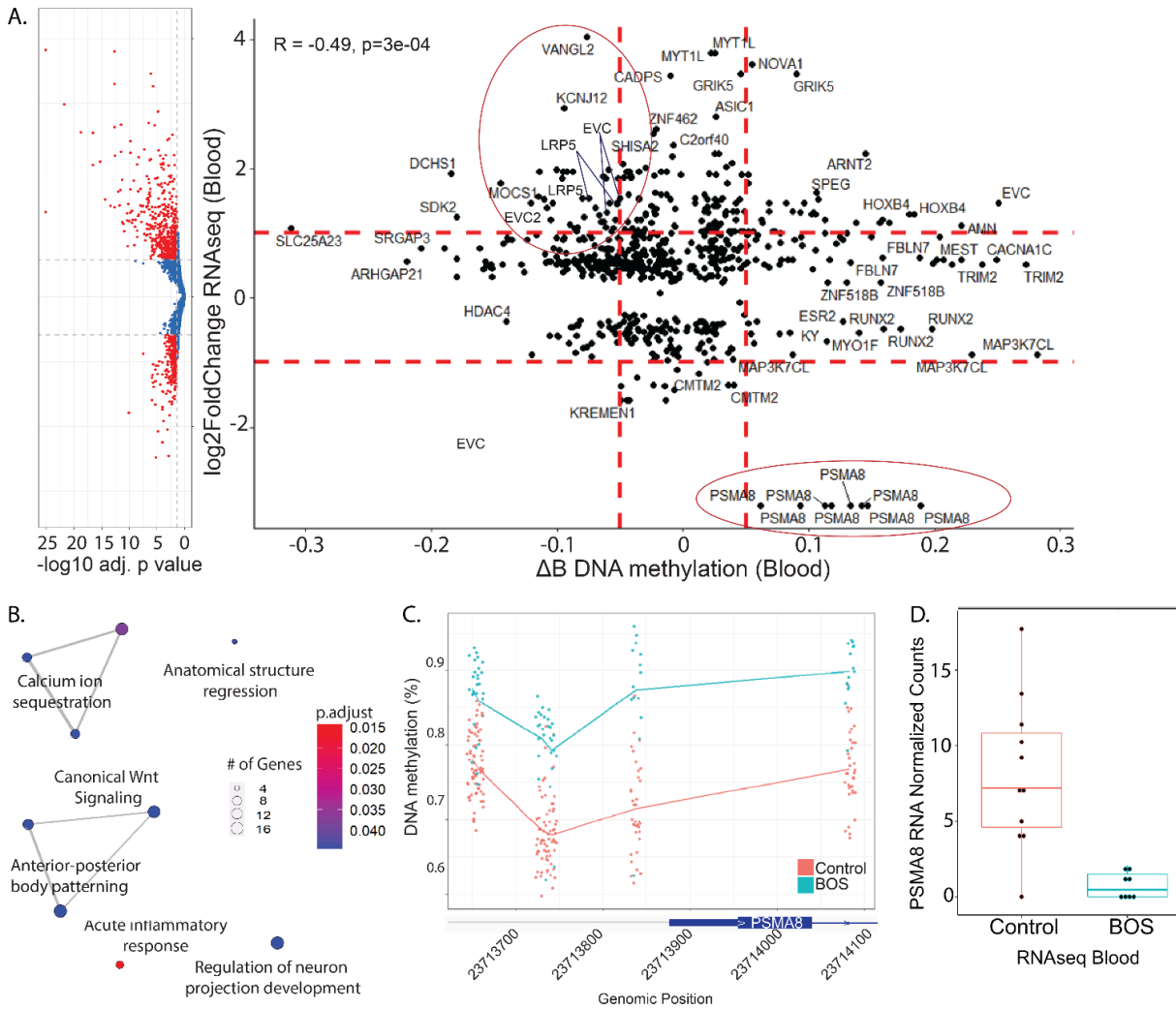


Figure 2-5: DNA methylation drives transcriptomic dysregulation in BOS samples and identifies common dysregulated transcripts enriched in Wnt signaling genes

(A) Integration of patient blood samples across DNA methylation (DNAm) (BOS n=13, control n=26) and RNA transcriptomic (BOS n=8, control n=11) dysregulation identifies 672 differentially methylated CpG sites (adjusted $p < 0.05$) that correlate with 341 RNA-seq Differentially Expressed Genes (DEGs) (adjusted $p < 0.05$). These significant DEGs were further filtered for RNA-seq $|\log_2FC| \geq 0.58$ and DNAm $|\Delta\beta| \geq 0.05$, shown by the dotted red lines. After filtering, we retained 50/672 CpG sites (7.44%) and 24/341 unique genes (7.04%).

(B) Analysis of enriched biological processes identified canonical Wnt signaling, anterior-posterior body patterning, regulation of neuron projection development, and other biologically relevant pathways.

(C) In BOS patients, *PSMA8*, a key component of the β -catenin destruction complex, is hypermethylated in blood DNAm across 8 CpG sites ($\Delta\beta$ 6.1% to 18.9%).

(D) *PSMA8* shows strong downregulation in blood RNA-seq ($\log_2FC=-2.92$). Control sample normalized counts ranged from 0 to 17.7, with a mean of 8.1, and quartile bounds of 4.6 and 10.8. BOS normalized sample counts ranged from 0 to 2.1, with a mean of 0.8, and quartile bounds of 0 and 1.5.

(B) Whole cell lysate (15ug) of representative BOS (n=5) and control-derived (n=5) fibroblasts show downstream Wnt pathway activation at the protein level through staining for VANGL2, β -catenin, axis-inhibition protein 1 (AXIN1), AXIN2, DVL2, and DVL3.

(C) ImageJ quantification identified a fold increase of 1.5x for AXIN1, 2.8x for AXIN2, 3.5x for DVL2, and 1.5x for DVL3 in BOS. This was repeated 2 times.

In BOS patient samples, the Wnt pathway co-receptor *LRP5* (green) shows transcriptional upregulation in **(D)** blood ($\log_2FC=1.64$, $p_{adj}=3.58 \times 10^{-9}$) and **(E)** fibroblast RNA-seq and **(F)** DNA hypomethylation in blood ($\Delta\beta$ -3.5% to -8.0%, $FDR<0.05$) at multiple CpG sites. Similarly, *LRP6* (green) shows transcriptional upregulation in **(G)** blood RNA-seq ($\log_2FC=1.63$, $p_{adj}=1.17 \times 10^{-12}$) and **(H)** fibroblast RNA-seq and **(I)** DNA hypomethylation in blood ($\Delta\beta$ -2.7% to -4.0%, $FDR<0.05$).

BOS samples exhibit strong dysregulation of *VANGL2* across tissue and assay types (pink). *VANGL2* is **(J)** hypomethylated ($\Delta\beta$ -7.6%) at CpG site cg17024258, and shows **(K)** increased chromatin accessibility at the 5' UTR ($\log_2FC=1.20$), and significant transcriptional upregulation in **(L)** blood RNA-seq ($\log_2FC=3.80$) and **(M)** fibroblast RNA-seq ($\log_2FC=2.55$). The *VANGL2* promoter shows **(N)** increased chromatin accessibility and **(O)** increased H3K4me3 marks in BOS.

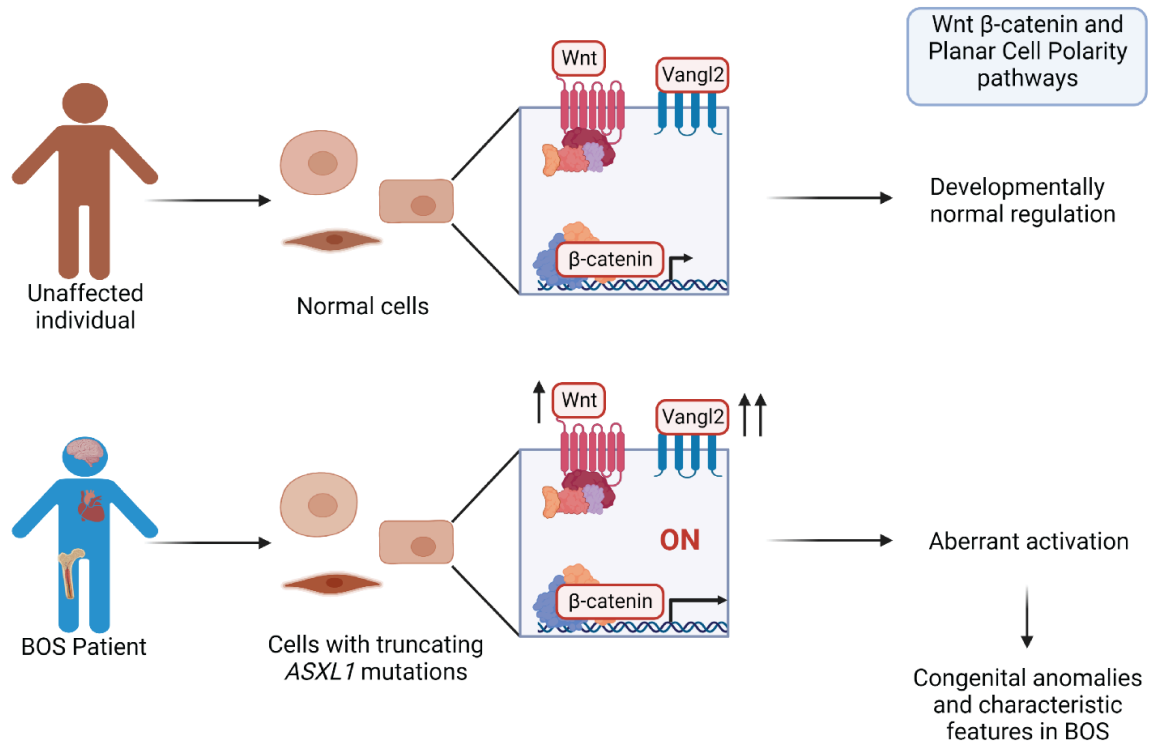


Figure 2-7: Multi-omics analyses identify broad activation of canonical Wnt signaling and planar cell polarity pathways across BOS patient tissues

Pt#	Condition	ASXL1 Mutation		Severity	Sex	Age	Assay						
		cDNA Change NM_015338.6	Protein Change Q8IXJ9				RNAseq		ATACseq	DNAm	DNAm	CUT&Run	Western Blot
							Fibro	Blood	Fibro	Blood	Fibro	Fibro	Fibro
Pt1	BOS	c.1091delG	p.G364Vfs*2	Severe	M	31	x	x	x	x	x		
Pt2	BOS	c.1210C>T	p.R404*	Moderate	F	9	x		x		x		x
Pt3	BOS	c.1910_1922del	p.E635Rfs*15	Moderate	M	24				x			x
Pt4	BOS	c.1924G>T	p.G642*	Moderate	M	8	x			x	x	x	x
Pt5	BOS	c.1934_1935insG	p.G646Wfs*12	Mild	M	2				x			
Pt6	BOS	c.2013_2014delAG	p.C672Wfs*4	Moderate	F	15	x	x	x	x	x		
Pt7	BOS	c.2313_2318del_insTTGG	p.A772Wfs*14	Mod. / Severe	F	29		x		x			
Pt8	BOS	c.2416_2417dupAC	p.V807Pfs*12	Severe	M	15		x		x			
Pt9	BOS	c.2477delG	p.G826Efs*12	?	F	3	x		x		x	x	
Pt10	BOS	c.2759_2762 dupCATC	p.V922Ifs*3	Moderate	M	7				x			
Pt11	BOS	c.2810delC	p.P937Lfs*8	?	F	9		x		x			
Pt12	BOS	c.2893C>T	p.R965*	?	F	10	x	x	x	x	x		
Pt13	BOS	c.2922C>A	p.Y974*	Moderate	M	7		x					
Pt14	BOS	c.2981delC	p.P994Lfs*30	Severe	F	5	x		x		x		x
Pt15	BOS	c.3077delG	p.G1026Dfs*21	Severe	F	7	x		x		x		x
Pt16	BOS	c.4060G>T	p.E1354*	?	F	4				x			
Pt17	BOS	c.4060G>T	p.E1354*	?	M	12		x		x			
Pt18	BOS	c.4243C>T	p.R1415*	Mild	M	4		x		x			

Table 2-1: Bohring-Opitz syndrome (BOS) patient information and assays

Bohring-Opitz syndrome (BOS) patients included in our assays are identified with their patient IDs. ASXL1 mutation, age at collection, sex, and assays conducted are shown.

Ctrl#	Condition	Sex	Age	Assay						
				RNAseq		ATACseq	CUT&Run	DNAm	DNAm	Western Blot
				Fibro	Blood	Fibro	Fibro	Blood	Fibro	Fibro
Ctrl1	Control	F	37					x	x	x
Ctrl2	Control	M	38	x					x	x
Ctrl3	Control	F	40	x					x	x
Ctrl4	Control	M	44	x		x			x	
Ctrl5	Control	F	41	x		x		x	x	x
Ctrl6	Control	F	42	x					x	
Ctrl7	Control	M	42		x					
Ctrl8	Control	M	46	x						x
Ctrl9	Control	F	43	x	x					
Ctrl10	Control	F	60	x	x	x				x
Ctrl11	Control	F	54		x					
Ctrl12	Control	M	49		x					
Ctrl13	Control	M	44		x					
Ctrl14	Control	F	41		x					
Ctrl15	Control	M	58		x					
Ctrl16	Control	F	60		x					
Ctrl17	Control	M	40		x					
Ctrl18	Control	F	48		x					
Ctrl19	Control	F	28		x					
Ctrl20	Control	M	11			x				
Ctrl21	Control	F	13			x				
Ctrl22	Control	M	45			x				
Ctrl23	Control	F	45			x				
Ctrl24	Control	F	34.7				x			
Ctrl25	Control	M	8				x			
Ctrl26	Control	F	14				x			
Ctrl27	Control	F	21				x			
Ctrl28	Control	M	16				x			
Ctrl29	Control	M	20				x			
Ctrl30	Control	M	12				x			
Ctrl31	Control	F	28				x			
Ctrl32	Control	F	18				x			
Ctrl33	Control	F	21				x			
Ctrl34	Control	F	21				x			
Ctrl35	Control	F	11				x			
Ctrl36	Control	M	11				x			
Ctrl37	Control	M	10				x			
Ctrl38	Control	M	15				x			
Ctrl39	Control	F	36				x			
Ctrl40	Control	M	22				x			

Ctrl41	Control	F	28				x			
Ctrl42	Control	F	15				x			
Ctrl43	Control	F	19				x			
Ctrl44	Control	F	21				x			
Ctrl45	Control	M	10				x			
Ctrl46	Control	M	37				x			
Ctrl47	Control	M	4				x			
Ctrl48	Control	M	4				x			
Ctrl49	Control	M	4				x			
Ctrl50	Control	F	56							x
Ctrl51	Control	M	55							x
Ctrl52	Control	F	45							x
Ctrl53	Control	M	44							x
Ctrl54	Control	F	29						x	
Ctrl55	Control	F	46						x	

Table 2-2: Control sample information and assays

Control samples included in our assays are identified with their control IDs. Age at collection, sex, and assays conducted are shown.

Antibody	Clone	Raised In	Expected Size	Vendor	Catalog #
ASXL1	Polyclonal	Rabbit	150kDa	Thermofisher	PA5-71475
ASXL1	Monoclonal (12F9)	Rat	165kDa	Millipore	MABE1933
ASXL1	Polyclonal	Rabbit	200kDa	abcam	ab228009
ASXL1	Polyclonal	Rabbit	N/A	abcam	ab55825
ASXL1	Polyclonal	Rabbit	N/A	Abnova	PAB5254
ASXL1	Monoclonal (11I4)	Mouse	~155kDa	Invitrogen	MA536238
ASXL1	Monoclonal (2049C2a)	Mouse	N/A	abcam	ab50817
ASXL1	Monoclonal (6E2)	Mouse	165kDa	Santa Cruz	sc-293204
Rhodamine Anti-B-Actin			42kDa	Bio-Rad	12004163
FLAG	Monoclonal (M2)	Mouse	Not Listed	Sigma-Aldrich	F1804
H3K4me3	Monoclonal (C42D8)	Rabbit	17kDa	Cell Signaling	9751
H3c	Monoclonal (D1H2)	Rabbit	17kDa	Cell Signaling	4499
H3K27me3	Monoclonal (C36B11)	Rabbit	17kDa	Cell Signaling	9733
H2AK119ub	Monoclonal (D27C4)	Rabbit	23kDa	Cell Signaling	8240
DVL2	Monoclonal (20D2)	Rabbit	90 to 95 kDa	Cell Signaling	3224
DVL3	Polyclonal	Rabbit	88 to 93 kDa	Cell Signaling	3218
AXIN1	Monoclonal (C76H11)	Rabbit	110 kDa	Cell Signaling	2087
AXIN2	Polyclonal	Rabbit	95kDa	Thermofisher	PA521093
VANGL2	Polyclonal	Rabbit	60 to 70kDa	Proteintech	21492-1-AP
β-Catenin	Polyclonal	Rabbit	92kDa	Invitrogen	71-2700
Anti-Rabbit Alexa Fluor 680		Goat		Invitrogen	A-32734
Anti-Mouse Alexa Fluor 488		Goat		Invitrogen	A-11001

Table 2-3: List of antibodies used

Antibodies used in this paper are shown here with their purchasing information.

Assays		All	RNAseq	RNAseq	ATACseq	CUT&RUN	DNAm
Sample Type		Blood + Fibro	Fibro	Blood	Fibro	Fibro	Blood
Total # Samples		18	7	8	6	2	13
Sex							
	Male	9 (50%)	1 (14.3%)	4 (50%)	0 (0%)	1 (50%)	8 (61.5%)
	Female	9 (50%)	6 (85.7%)	4 (50%)	6 (100%)	1(50%)	5 (38.5%)
Age (yr)							
	Avg (Range)	11.2 (2 - 31)	8.1 (3 - 15)	12.6 (4 - 29)	8.2 (3 - 15)	5.5 (3 - 8)	13.1 (2 - 31)
Mutation Location							
	Coding (bp)	2617 (1091 - 4243)	2368 (1210 - 3077)	2630 (2013 - 4243)	2442 (1210 - 3077)	2200 (1924 - 2477)	2649 (1091 - 4243)
	Protein (AA)	872 (364 - 1415)	790 (404 - 1026)	987 (672 - 1415)	815 (404 - 1026)	734 (642 - 826)	883 (364 - 1415)
Mutation Type							
	Substitution	7 (38.9%)	4 (57.1%)	4 (50%)	2 (33.3%)	1 (50%)	5 (38.5%)
	In/Del/Dup	11 (61.1%)	3 (42.9%)	4 (50%)	4 (66.7%)	1 (50%)	8 (61.5%)
	Nonsense	7 (38.9%)	3 (42.9%)	4 (50%)	2 (33.3%)	1 (50%)	5 (38.5%)
	Frameshift	11 (61.1%)	4 (57.1%)	4 (50%)	4 (66.7%)	1 (50%)	8 (61.5%)

Table 2-4: BOS patient summary table

Summary data (age, sex, mutation) for BOS patients in our cohort that are included in each -omics assay.

Gene Name	RefSeq	Exon Location	Probe (nmoles)	Primer (nmoles)	Probe Sequence	Primer 1 Sequence	Primer 2 Sequence
ASXL1	NM_015338	2 - 4	2.5	5	5'-/56-FAM/AGTCCTTCT/ZEN/ GCCTCTATGACCTGCA/3IABkFQ/-3'	5'-GCTCCAATGA CACC AAAACAG-3'	5'-AAACAACCCCT CTCCTCCT-3'
ASXL1	NM_015338	10 - 12	2.5	5	5'-/56-FAM/ACCTTCTTT/ZEN/ TCCTTCTCCATTTCTGTCG/3IABkFQ/-3'	5'GCCTGGCTGA TGGTGAA-3'	5'-GACTCTTCTTTG GTCAAACCCA-3'
ACTB	NM_001101	1 - 2	2.5	5	5'-5HEX/TCATCCATG/ZEN/ GTGAGCTGGCGG/3IABkFQ/-3'	5'-ACAGAGCCTC GCCTTTG-3	5'-CCTTGACATGC CGGAG-3'

Table 2-5: RT-qPCR primers

PrimeTime™ RT-qPCR primers were purchased from IDT to assay ASXL1 expression. Pre-designed primer pairs and fluorescently labeled 5' nuclease probe that were used in our assays are listed. These are two primer sets for ASXL1, each targeting a different exonic range. and one set targeting B-actin (ACTB) which was used as a control.

Gene ID	log2FC	pvalue	padj	Gene Name	Approved Name
ENSG00000168671.10	4.81	5.00E-15	3.74E-11	UGT3A2	UDP glycosyltransferase family 3 member A2
ENSG00000000005.6	4.47	2.65E-07	2.05E-04	TNMD	tenomodulin
ENSG00000144619.15	4.23	4.36E-10	1.09E-06	CNTN4	contactin 4
ENSG00000260230.3	3.45	2.12E-14	1.19E-10	FRRS1L	ferric chelate reductase 1 like
ENSG00000087258.15	3.36	5.24E-10	1.18E-06	GNAO1	G protein subunit alpha o1
ENSG00000182132.13	3.35	9.18E-08	1.03E-04	KCNIP1	potassium voltage-gated channel interacting protein 1
ENSG00000134115.13	3.34	6.71E-08	8.36E-05	CNTN6	contactin 6
ENSG00000104435.13	-3.17	1.64E-06	8.19E-04	STMN2	stathmin 2
ENSG00000246379.8	3.03	6.34E-05	1.31E-02	AC040168.1	NA
ENSG00000006432.15	3.00	2.12E-07	1.80E-04	MAP3K9	mitogen-activated protein kinase kinase kinase 9
ENSG00000266554.2	-2.86	1.04E-07	1.11E-04	LINC01443	long intergenic non-protein coding RNA 1443
ENSG00000186998.16	2.79	4.63E-07	2.97E-04	EMID1	EMI domain containing 1
ENSG00000133019.11	2.67	1.54E-05	4.73E-03	CHRM3	cholinergic receptor muscarinic 3
ENSG00000132622.11	2.59	6.75E-07	3.98E-04	HSPA12B	heat shock protein family A (Hsp70) member 12B
ENSG00000162738.6	2.55	2.79E-14	1.25E-10	VANGL2	VANGL planar cell polarity protein 2
ENSG00000181444.13	2.52	1.08E-06	6.06E-04	ZNF467	zinc finger protein 467
ENSG00000286457.1	2.52	1.93E-04	2.99E-02	AC009560.2	NA
ENSG00000120337.8	2.49	1.10E-07	1.13E-04	TNFSF18	TNF superfamily member 18
ENSG00000105737.9	2.49	6.02E-06	2.46E-03	GRIK5	glutamate ionotropic receptor kainate type subunit 5
ENSG00000270110.2	2.46	2.89E-07	2.09E-04	AL353593.2	NA
ENSG00000156689.7	2.46	2.85E-07	2.09E-04	GLYATL2	glycine-N-acyltransferase like 2
ENSG00000267886.1	2.46	1.66E-07	1.55E-04	AC074135.1	NA
ENSG00000287252.1	2.42	1.15E-05	3.91E-03	AL009179.2	NA
ENSG00000052850.8	2.38	2.16E-07	1.80E-04	ALX4	ALX homeobox 4
ENSG00000138769.11	2.35	2.66E-06	1.24E-03	CDKL2	cyclin dependent kinase like 2
ENSG00000140279.12	2.29	2.04E-04	3.12E-02	DUOX2	dual oxidase 2
ENSG00000156475.18	-2.26	1.28E-04	2.23E-02	PPP2R2B	protein phosphatase 2 regulatory subunit Bbeta
ENSG00000115457.10	2.19	4.16E-06	1.76E-03	IGFBP2	insulin like growth factor binding protein 2
ENSG00000158258.16	2.17	3.31E-06	1.52E-03	CLSTN2	calsyntenin 2
ENSG00000180875.5	-2.11	1.51E-09	3.09E-06	GREM2	gremlin 2, DAN family BMP antagonist

Table 2-6: 30 most significant differentially expressed genes in BOS fibroblasts identified through RNA-seq

DESeq2 was used to analyze RNA-seq gene counts from our cohort of BOS and control fibroblast samples. Samples were adjusted for the co-variate of sex based on principal component analysis. Genes were considered significant if Bonferroni adjusted $p < 0.05$ and the top 30 dysregulated genes are listed here with their adjusted p value, log₂Fold Change, and gene annotations.

Gene ID	log2FC	pvalue	padj	Gene Name	Approved Name
ENSG00000165349.12	4.80	1.80E-09	5.56E-07	SLC7A3	solute carrier family 7 member 3
ENSG00000267313.7	4.51	7.66E-04	1.80E-02	AC021504.1	NA
ENSG00000105737.9	3.83	5.56E-30	7.04E-26	GRIK5	glutamate ionotropic receptor kainate type subunit 5
ENSG00000162738.6	3.80	1.15E-16	2.25E-13	VANGL2	VANGL planar cell polarity protein 2
ENSG00000139910.20	3.46	3.06E-09	8.44E-07	NOVA1	NOVA alternative splicing regulator 1
ENSG00000173894.11	3.30	9.77E-17	2.25E-13	CBX2	chromobox 2
ENSG00000186487.19	3.26	9.46E-09	2.12E-06	MYT1L	myelin transcription factor 1 like
ENSG00000110881.11	2.98	2.05E-26	1.73E-22	ASIC1	acid sensing ion channel subunit 1
ENSG00000107447.8	2.87	8.35E-08	1.42E-05	DNTT	DNA nucleotidylexotransferase
ENSG00000287268.1	2.70	1.02E-09	3.44E-07	AL050350.1	NA
ENSG00000180730.4	2.67	2.12E-15	2.98E-12	SHISA2	shisa family member 2
ENSG00000147257.14	2.66	2.05E-08	4.29E-06	GPC3	glypican 3
ENSG00000119147.10	2.63	8.33E-17	2.11E-13	C2orf40	ECRG4 augurin precursor
ENSG00000148143.13	2.55	2.87E-23	1.82E-19	ZNF462	zinc finger protein 462
ENSG00000163520.14	2.54	4.83E-21	2.45E-17	FBLN2	fibulin 2
ENSG00000171227.7	2.51	1.42E-04	5.83E-03	TMEM37	transmembrane protein 37
ENSG00000180875.5	-2.48	3.26E-08	6.02E-06	GREM2	gremlin 2, DAN family BMP antagonist
ENSG00000000971.15	-2.46	1.77E-05	1.15E-03	CFH	complement factor H
ENSG00000261025.1	2.45	2.42E-05	1.46E-03	AL445471.2	NA
ENSG00000140986.8	2.43	9.89E-06	7.33E-04	RPL3L	ribosomal protein L3 like

Table 2-7: 20 most significant differentially expressed genes in BOS blood identified through RNA-seq

DESeq2 was used to analyze RNA-seq gene counts from our cohort of BOS and control blood samples. Samples were adjusted for the co-variate of sex based on principal component analysis. Genes were considered significant if Bonferroni adjusted $p < 0.05$ and the top 20 dysregulated genes are listed here with their adjusted p value, log₂Fold Change, and gene annotations.

Gene ID	log2FC	pvalue	padj	Annotation	Gene Name	Approved Name
ENSG0000020633	4.39	8.36E-14	1.05E-09	Promoter-TSS	RUNX3	RUNX family transcription factor 3
ENSG00000118113	2.51	1.53E-13	1.66E-09	Promoter-TSS	MMP8	matrix metalloproteinase 8
ENSG00000122378	1.76	1.68E-08	2.21E-05	Promoter-TSS	PRXL2A	peroxiredoxin like 2A
ENSG00000229676	1.34	2.32E-08	2.83E-05	Promoter-TSS	ZNF492	zinc finger protein 492
ENSG00000115594	-1.92	2.66E-07	1.79E-04	Promoter-TSS	IL1R1	interleukin 1 receptor type 1
ENSG00000154143	2.36	3.34E-07	2.06E-04	Promoter-TSS	PANX3	pannexin 3
ENSG00000132639	1.18	9.14E-07	4.23E-04	Promoter-TSS	SNAP25	synaptosome associated protein 25
ENSG00000003400	-1.43	2.18E-06	7.43E-04	Promoter-TSS	CASP10	caspase 10
N/A	1.56	2.75E-06	8.87E-04	Promoter-TSS	LOC150051	Homeobox protein cut-like 1
ENSG00000204442	-1.44	2.93E-06	9.21E-04	5'UTR	FAM155A	family with sequence similarity 155 member A
ENSG00000146859	-1.05	3.05E-06	9.45E-04	Promoter-TSS	TMEM140	transmembrane protein 140
ENSG00000156265	-1.58	6.47E-06	1.61E-03	5'UTR	MAP3K7CL	MAP3K7 C-terminal like
ENSG00000180660	-1.74	7.74E-06	1.79E-03	Promoter-TSS	MAB21L1	mab-21 like 1
ENSG00000030419	-1.53	1.21E-05	2.48E-03	Promoter-TSS	IKZF2	IKAROS family zinc finger 2
ENSG00000142698	1.17	1.38E-05	2.65E-03	5'UTR	C1orf94	chromosome 1 open reading frame 94
ENSG00000124212	-1.12	1.40E-05	2.68E-03	Promoter-TSS	PTGIS	prostaglandin I2 synthase
ENSG00000162738	1.25	4.54E-05	5.97E-03	5'UTR	VANGL2	VANGL planar cell polarity protein 2
ENSG00000120093	1.89	5.07E-05	6.37E-03	Promoter-TSS	HOXB3	homeobox B3
ENSG00000114757	1.18	5.80E-05	6.94E-03	Promoter-TSS	PEX5L	peroxisomal biogenesis factor 5 like
ENSG00000253953	-1.03	1.12E-04	1.06E-02	Promoter-TSS	PCDHGB4	protocadherin gamma subfamily B, 4

Table 2-8: Significant differentially expressed genes in BOS fibroblasts identified through ATAC-seq

DESeq2 was used to analyze ATACseq gene counts from our cohort of BOS and control fibroblast samples. Samples were adjusted for the co-variate of sex based on principal component analysis. Genes were considered significant if Bonferroni adjusted $p < 0.05$ and the top 20 dysregulated genes are listed here with their adjusted p value, log₂Fold Change, and gene annotations.

Gene Name	Gene Function	log2FC		
		RNA-seq Blood	RNA-seq Fibro	ATACseq Fibro
VANGL2	Tissue polarity and patterning, and neural plate development	3.80	2.55	1.25
PRXL2A	Antioxidant activity and osteoclast differentiation	1.15	1.96	1.30
APBA2	Neuronal adapter protein	0.94	2.11	2.15
OTUD7A	Deubiquitinating enzyme and potential tumour suppressor	0.90	0.86	2.44
KCNQ5	Potassium channel critical in neurons	0.65	-0.01	-1.39
PELI2	MAPK cascade, ubiquitin ligase activity	-0.37	0.45	0.87
GREM2	BMP antagonist, organogenesis and body patterning	-2.48	-2.11	-1.48

Table 2-9: Integration of datasets across tissue types and across sample types identified a subset of overlapping genes that are significantly dysregulated in BOS

These 7 genes were identified across blood and fibroblast RNA-seq as well as fibroblast ATACseq. Gene functions and their respective significant log2FC are shown.

GO ID	Description	GeneRatio	BgRatio	pvalue	p.adjust
GO:0002675	positive regulation of acute inflammatory response	6/306	27/18637	4.13E-06	0.01
GO:0051235	maintenance of location	17/306	318/18637	2.18E-05	0.04
GO:0060070	canonical Wnt signaling pathway	17/306	333/18637	3.91E-05	0.04
GO:0050729	positive regulation of inflammatory response	10/306	133/18637	7.00E-05	0.04
GO:0007389	pattern specification process	19/306	426/18637	8.60E-05	0.04
GO:0010975	regulation of neuron projection development	19/306	429/18637	9.43E-05	0.04
GO:0031349	positive regulation of defense response	17/306	361/18637	1.06E-04	0.04
GO:0051282	regulation of sequestering of calcium ion	9/306	115/18637	1.18E-04	0.04
GO:0002673	regulation of acute inflammatory response	6/306	48/18637	1.28E-04	0.04
GO:0060033	anatomical structure regression	4/306	17/18637	1.43E-04	0.04
GO:0009952	anterior/posterior pattern specification	12/306	204/18637	1.46E-04	0.04
GO:0051208	sequestering of calcium ion	9/306	119/18637	1.54E-04	0.04
GO:0051651	maintenance of location in cell	12/306	208/18637	1.75E-04	0.05
GO:0005516	calmodulin binding	13/316	198/18181	4.49E-05	0.02

Table 2-10: Gene ontology analysis for BOS blood RNA-seq and DNA methylation integration

Integration of DEGs from BOS RNA-seq blood and BOS blood DNA methylation were analyzed using clusterProfiler v3.12.0 to identify gene set enrichments. clusterProfiler v3.12.0 uses all genes from the Gencode hg38 annotation, version 31, as background. We filtered for gene ontologies with padj < 0.05 for significance.

References

1. Fisher, C. L., Berger, J., Randazzo, F. & Brock, H. W. A human homolog of Additional sex combs, ADDITIONAL SEX COMBS-LIKE 1, maps to chromosome 20q11. *Gene* **306**, 115–126 (2003).
2. Aravind, L. & Iyer, L. M. The HARE-HTH and associated domains: novel modules in the coordination of epigenetic DNA and protein modifications. *Cell Cycle* **11**, 119–131 (2012).
3. Katoh, M. Functional proteomics of the epigenetic regulators ASXL1, ASXL2 and ASXL3: a convergence of proteomics and epigenetics for translational medicine. *Expert Rev. Proteomics* **12**, 317–328 (2015).
4. Jain, K. *et al.* Characterization of the plant homeodomain (PHD) reader family for their histone tail interactions. *Epigenetics Chromatin* **13**, 3 (2020).
5. Sanchez, R. & Zhou, M.-M. The PHD finger: a versatile epigenome reader. *Trends Biochem. Sci.* **36**, 364–372 (2011).
6. Schuettengruber, B., Bourbon, H.-M., Di Croce, L. & Cavalli, G. Genome Regulation by Polycomb and Trithorax: 70 Years and Counting. *Cell* **171**, 34–57 (2017).
7. Fisher, C. L. *et al.* Additional sex combs-like 1 belongs to the enhancer of trithorax and polycomb group and genetically interacts with Cbx2 in mice. *Dev. Biol.* **337**, 9–15 (2010).
8. Matheus, F. *et al.* Pathological ASXL1 Mutations and Protein Variants Impair Neural Crest Development. *Stem Cell Reports* **12**, 861–868 (2019).
9. McGinley, A. L., Li, Y., Deliu, Z. & Wang, Q. T. Additional sex combs-like family genes are required for normal cardiovascular development. *Genesis* **52**, 671–686 (2014).
10. Abdel-Wahab, O. *et al.* Deletion of Asxl1 results in myelodysplasia and severe developmental defects in vivo. *J. Exp. Med.* **210**, 2641–2659 (2013).
11. Yamato, G. *et al.* ASXL2 mutations are frequently found in pediatric AML patients with t(8;21)/RUNX1-RUNX1T1 and associated with a better prognosis. *Genes Chromosomes*

- Cancer* **56**, 382–393 (2017).
12. Pereira, J. D. *et al.* Ezh2, the histone methyltransferase of PRC2, regulates the balance between self-renewal and differentiation in the cerebral cortex. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 15957–15962 (2010).
 13. Sahtoe, D. D., van Dijk, W. J., Ekkebus, R., Ovaa, H. & Sixma, T. K. BAP1/ASXL1 recruitment and activation for H2A deubiquitination. *Nat. Commun.* **7**, 10292 (2016).
 14. Inoue, D., Nishimura, K., Kozuka-Hata, H., Oyama, M. & Kitamura, T. The stability of epigenetic factor ASXL1 is regulated through ubiquitination and USP7-mediated deubiquitination. *Leukemia* **29**, 2257–2260 (2015).
 15. Rothberg, J. L. M. *et al.* Mtf2-PRC2 control of canonical Wnt signaling is required for definitive erythropoiesis. *Cell Discov* **4**, 21 (2018).
 16. Chen, J. *et al.* The microtubule-associated protein PRC1 promotes early recurrence of hepatocellular carcinoma in association with the Wnt/ β -catenin signalling pathway. *Gut* **65**, 1522–1534 (2016).
 17. Asada, S. *et al.* Mutant ASXL1 cooperates with BAP1 to promote myeloid leukaemogenesis. *Nature Communications* vol. 9 Preprint at <https://doi.org/10.1038/s41467-018-05085-9> (2018).
 18. Asada, S. & Kitamura, T. Aberrant histone modifications induced by mutant ASXL1 in myeloid neoplasms. *Int. J. Hematol.* **110**, 179–186 (2019).
 19. Kweon, S.-M. *et al.* An Adversarial DNA N6-Methyladenine-Sensor Network Preserves Polycomb Silencing. *Mol. Cell* **74**, 1138–1147.e6 (2019).
 20. Srivastava, A., McGrath, B. & Bielas, S. L. Histone H2A Monoubiquitination in Neurodevelopmental Disorders. *Trends in Genetics* vol. 33 566–578 Preprint at <https://doi.org/10.1016/j.tig.2017.06.002> (2017).
 21. Srivastava, A. *et al.* De novo dominant ASXL3 mutations alter H2A deubiquitination and transcription in Bainbridge–Ropers syndrome. *Hum. Mol. Genet.* **25**, 597–608 (2015).

22. Bainbridge, M. N. *et al.* De novo truncating mutations in ASXL3 are associated with a novel clinical phenotype with similarities to Bohring-Opitz syndrome. *Genome Med.* **5**, 11 (2013).
23. Fang, X. *et al.* Asxl1 C-terminal mutation perturbs neutrophil differentiation in zebrafish. *Leukemia* vol. 35 2299–2310 Preprint at <https://doi.org/10.1038/s41375-021-01121-8> (2021).
24. Russell, B., Tan, W.-H. & Graham, J. M., Jr. Bohring-Opitz syndrome. in *GeneReviews®* (eds. Adam, M. P. *et al.*) (University of Washington, Seattle, Seattle (WA), 2018).
25. Hoischen, A. *et al.* De novo nonsense mutations in ASXL1 cause Bohring-Opitz syndrome. *Nat. Genet.* **43**, 729–731 (2011).
26. Shashi, V. *et al.* De Novo Truncating Variants in ASXL2 Are Associated with a Unique and Recognizable Clinical Phenotype. *Am. J. Hum. Genet.* **99**, 991–999 (2016).
27. Russell, B. & Graham, J. M. Expanding our knowledge of conditions associated with the ASXL gene family. *Genome Med.* **5**, 1–3 (2013).
28. Pratcorona, M. *et al.* Acquired mutations in ASXL1 in acute myeloid leukemia: prevalence and prognostic value. *Haematologica* **97**, 388–392 (2012).
29. Genovese, G. *et al.* Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N. Engl. J. Med.* **371**, 2477–2487 (2014).
30. Gelsi-Boyer, V. *et al.* Mutations in ASXL1 are associated with poor prognosis across the spectrum of malignant myeloid diseases. *J. Hematol. Oncol.* **5**, 12 (2012).
31. Micol, J.-B. & Abdel-Wahab, O. The Role of Additional Sex Combs-Like Proteins in Cancer. *Cold Spring Harb. Perspect. Med.* **6**, (2016).
32. Jaiswal, S. *et al.* Age-related clonal hematopoiesis associated with adverse outcomes. *N. Engl. J. Med.* **371**, 2488–2498 (2014).
33. Dawoud, A. A. Z., Tapper, W. J. & Cross, N. C. P. Clonal myelopoiesis in the UK Biobank cohort: ASXL1 mutations are strongly associated with smoking. *Leukemia* **34**, 2660–2672 (2020).

34. Nagase, R. *et al.* Expression of mutant Asxl1 perturbs hematopoiesis and promotes susceptibility to leukemic transformation. *J. Exp. Med.* **215**, 1729–1747 (2018).
35. Wang, J. *et al.* Loss of Asxl1 leads to myelodysplastic syndrome-like disease in mice. *Blood* **123**, 541–553 (2014).
36. Inoue, D. *et al.* Truncation mutants of ASXL1 observed in myeloid malignancies are expressed at detectable protein levels. *Experimental Hematology* vol. 44 172–176.e1 Preprint at <https://doi.org/10.1016/j.exphem.2015.11.011> (2016).
37. Yang, H. *et al.* Gain of function of ASXL1 truncating protein in the pathogenesis of myeloid malignancies. *Blood* **131**, 328–341 (2018).
38. Guo, M., Peng, Y., Gao, A., Du, C. & Herman, J. G. Epigenetic heterogeneity in cancer. *Biomark Res* **7**, 23 (2019).
39. Nusse, R. & Clevers, H. Wnt/ β -Catenin Signaling, Disease, and Emerging Therapeutic Modalities. *Cell* **169**, 985–999 (2017).
40. Malhotra, S. & Kincade, P. W. Canonical Wnt pathway signaling suppresses VCAM-1 expression by marrow stromal and hematopoietic cells. *Exp. Hematol.* **37**, 19–30 (2009).
41. Tamai, K. *et al.* LDL-receptor-related proteins in Wnt signal transduction. *Nature* **407**, 530–535 (2000).
42. Wu, D. & Pan, W. GSK3: a multifaceted kinase in Wnt signaling. *Trends Biochem. Sci.* **35**, 161–168 (2010).
43. Mao, J. *et al.* Low-density lipoprotein receptor-related protein-5 binds to Axin and regulates the canonical Wnt signaling pathway. *Mol. Cell* **7**, 801–809 (2001).
44. Kim, S.-E. *et al.* Wnt stabilization of β -catenin reveals principles for morphogen receptor-scaffold assemblies. *Science* **340**, 867–870 (2013).
45. Wang, B., Tian, T., Kalland, K.-H., Ke, X. & Qu, Y. Targeting Wnt/ β -catenin signaling for cancer immunotherapy. *Trends Pharmacol. Sci.* **39**, 648–658 (2018).
46. Gajos-Michniewicz, A. & Czyz, M. WNT signaling in melanoma. *Int. J. Mol. Sci.* **21**, 4852

(2020).

47. van Amerongen, R. & Nusse, R. Towards an integrated view of Wnt signaling in development. *Development* **136**, 3205–3214 (2009).
48. Hagiwara, A., Yasumura, M., Hida, Y., Inoue, E. & Ohtsuka, T. The planar cell polarity protein Vangl2 bidirectionally regulates dendritic branching in cultured hippocampal neurons. *Mol. Brain* **7**, 79 (2014).
49. Vogelauer, M., Krall, A. S., McBrian, M. A., Li, J.-Y. & Kurdistani, S. K. Stimulation of Histone Deacetylase Activity by Metabolites of Intermediary Metabolism*. *J. Biol. Chem.* **287**, 32006–32016 (2012).
50. Corces, M. R. *et al.* An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat. Methods* **14**, 959–962 (2017).
51. Meers, M. P., Bryson, T. D., Henikoff, J. G. & Henikoff, S. Improved CUT&RUN chromatin profiling tools. *Elife* **8**, (2019).
52. Awamleh, Z. *et al.* DNA methylation signature associated with Bohring-Opitz syndrome: a new tool for functional classification of variants in ASXL genes. *Eur. J. Hum. Genet.* (2022) doi:10.1038/s41431-022-01083-0.
53. McLean, C. Y. *et al.* GREAT improves functional interpretation of cis-regulatory regions. *Nat. Biotechnol.* **28**, 495–501 (2010).
54. Andrews, S. & Others. FastQC: a quality control tool for high throughput sequence data. Preprint at (2010).
55. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
56. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
57. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
58. The Gene Ontology resource: enriching a GOld mine. *Nucleic Acids Res.* **49**, D325–D334

(2021).

59. Ashburner, M. *et al.* Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **25**, 25–29 (2000).
60. Yu, G., Wang, L.-G., Han, Y. & He, Q.-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284–287 (2012).
61. Heinz, S. *et al.* Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
62. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv [q-bio.GN]* (2013).
63. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
64. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
65. Leung, S. K. *et al.* Full-length transcript sequencing of human and mouse cerebral cortex identifies widespread isoform diversity and alternative splicing. *Cell Rep.* **37**, 110022 (2021).
66. Russell, B. E. *et al.* Clinical findings in 39 individuals with Bohring-Opitz syndrome from a global patient-driven registry with implications for tumor surveillance and recurrence risk. *Am. J. Med. Genet. A* **191**, 1050–1058 (2023).
67. Buenrostro, J. D., Wu, B., Chang, H. Y. & Greenleaf, W. J. ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Curr. Protoc. Mol. Biol.* **109**, 21.29.1–21.29.9 (2015).
68. Skene, P. J., Henikoff, J. G. & Henikoff, S. Targeted in situ genome-wide profiling with high efficiency for low cell numbers. *Nat. Protoc.* **13**, 1006–1019 (2018).
69. Abdel-Wahab, O. *et al.* ASXL1 mutations promote myeloid transformation through loss of PRC2-mediated gene repression. *Cancer Cell* **22**, 180–193 (2012).

70. Guo, H. *et al.* DNA methylation and chromatin accessibility profiling of mouse and human fetal germ cells. *Cell Res.* **27**, 165–183 (2017).
71. Yan, F., Powell, D. R., Curtis, D. J. & Wong, N. C. From reads to insight: a hitchhiker's guide to ATAC-seq data analysis. *Genome Biol.* **21**, 22 (2020).
72. Jones, C. & Chen, P. Planar cell polarity signaling in vertebrates. *Bioessays* **29**, 120–132 (2007).
73. Inoue, D. *et al.* Myelodysplastic syndromes are induced by histone methylation–altering ASXL1 mutations. *J. Clin. Invest.* **123**, 4627–4640 (2013).
74. Dezso, Z. *et al.* A comprehensive functional analysis of tissue specificity of human gene expression. *BMC Biol.* **6**, 49 (2008).
75. Lee, G. & Lee, M. Classification of Genes Based on Age-Related Differential Expression in Breast Cancer. *Genomics Inform.* **15**, 156–161 (2017).
76. Kim, J. H. *et al.* Wnt signaling in bone formation and its therapeutic potential for bone diseases. *Ther. Adv. Musculoskelet. Dis.* **5**, 13–31 (2013).
77. van Amerongen, R., Mikels, A. & Nusse, R. Alternative wnt signaling is initiated by distinct receptors. *Sci. Signal.* **1**, re9 (2008).
78. Park, H.-B., Kim, J.-W. & Baek, K.-H. Regulation of Wnt Signaling through Ubiquitination and Deubiquitination in Cancers. *Int. J. Mol. Sci.* **21**, (2020).
79. Mao, Y. Structure, Dynamics and Function of the 26S Proteasome. in *Macromolecular Protein Complexes III: Structure and Function* (eds. Harris, J. R. & Marles-Wright, J.) 1–151 (Springer International Publishing, Cham, 2021).
80. Rao, T. P. & Kühl, M. An updated overview on Wnt signaling pathways: a prelude for more. *Circ. Res.* **106**, 1798–1806 (2010).
81. van de Wetering, M. *et al.* The beta-catenin/TCF-4 complex imposes a crypt progenitor phenotype on colorectal cancer cells. *Cell* **111**, 241–250 (2002).
82. Wu, J. Q. *et al.* Tcf7 is an important regulator of the switch of self-renewal and

- differentiation in a multipotential hematopoietic cell line. *PLoS Genet.* **8**, e1002565 (2012).
83. Grumolato, L. *et al.* Canonical and noncanonical Wnts use a common mechanism to activate completely unrelated coreceptors. *Genes Dev.* **24**, 2517–2530 (2010).
 84. Montcouquiol, M. *et al.* Identification of Vangl2 and Scrb1 as planar polarity genes in mammals. *Nature* **423**, 173–177 (2003).
 85. Cetera, M., Leybova, L., Woo, F. W., Deans, M. & Devenport, D. Planar cell polarity-dependent and independent functions in the emergence of tissue-scale hair follicle patterns. *Dev. Biol.* **428**, 188–203 (2017).
 86. Wu, Z., Zhu, Y., Liu, H., Liu, G. & Li, F. Wnt10b promotes hair follicles growth and dermal papilla cells proliferation via Wnt/ β -Catenin signaling pathway in Rex rabbits. *Biosci. Rep.* **40**, (2020).
 87. Shariatmadari, M. *et al.* Increased Wnt levels in the neural tube impair the function of adherens junctions during neurulation. *Mol. Cell. Neurosci.* **30**, 437–451 (2005).
 88. Le Grand, F., Jones, A. E., Seale, V., Scimè, A. & Rudnicki, M. A. Wnt7a activates the planar cell polarity pathway to drive the symmetric expansion of satellite stem cells. *Cell Stem Cell* **4**, 535–547 (2009).
 89. Brunt, L. *et al.* Vangl2 promotes the formation of long cytonemes to enable distant Wnt/ β -catenin signaling. *Nat. Commun.* **12**, 2058 (2021).
 90. Ramsbottom, S. A. *et al.* Vangl2-regulated polarisation of second heart field-derived cells is required for outflow tract lengthening during cardiac development. *PLoS Genet.* **10**, e1004871 (2014).
 91. Yates, L. L. *et al.* The planar cell polarity gene Vangl2 is required for mammalian kidney-branching morphogenesis and glomerular maturation. *Hum. Mol. Genet.* **19**, 4663–4676 (2010).
 92. Bouali, S., Héту-Arbour, R., Gardet, C. & Heinonen, K. M. Vangl2 Promotes Hematopoietic Stem Cell Expansion. *Front Cell Dev Biol* **10**, 760248 (2022).

93. Torban, E. *et al.* Tissue, cellular and sub-cellular localization of the Vangl2 protein during embryonic development: effect of the Lp mutation. *Gene Expr. Patterns* **7**, 346–354 (2007).
94. Ueno, N. & Greene, N. D. E. Planar cell polarity genes and neural tube closure. *Birth Defects Res. C Embryo Today* **69**, 318–324 (2003).
95. Ossipova, O., Kim, K. & Sokol, S. Y. Planar polarization of Vangl2 in the vertebrate neural plate is controlled by Wnt and Myosin II signaling. *Biol. Open* **4**, 722–730 (2015).
96. De Calisto, J., Araya, C., Marchant, L., Riaz, C. F. & Mayor, R. Essential role of non-canonical Wnt signalling in neural crest migration. *Development* **132**, 2587–2597 (2005).
97. Torban, E., Wang, H.-J., Groulx, N. & Gros, P. Independent mutations in mouse Vangl2 that cause neural tube defects in looptail mice impair interaction with members of the Dishevelled family. *J. Biol. Chem.* **279**, 52703–52713 (2004).
98. Lei, Y.-P. *et al.* VANGL2 mutations in human cranial neural-tube defects. *The New England journal of medicine* vol. 362 2232–2235 (2010).
99. Kirby, M. L. Cellular and molecular contributions of the cardiac neural crest to cardiovascular development. *Trends Cardiovasc. Med.* **3**, 18–23 (1993).
100. Corda, G. & Sala, A. Non-canonical WNT/PCP signalling in cancer: Fzd6 takes centre stage. *Oncogenesis* **6**, e364 (2017).
101. van Amerongen, R. Alternative Wnt pathways and receptors. *Cold Spring Harb. Perspect. Biol.* **4**, (2012).
102. Angers, S. & Moon, R. T. Proximal events in Wnt signal transduction. *Nat. Rev. Mol. Cell Biol.* **10**, 468–477 (2009).
103. Montcouquiol, M. *et al.* Asymmetric localization of Vangl2 and Fz3 indicate novel mechanisms for planar cell polarity in mammals. *J. Neurosci.* **26**, 5265–5275 (2006).
104. Hua, Z. L., Jeon, S., Caterina, M. J. & Nathans, J. Frizzled3 is required for the development of multiple axon tracts in the mouse central nervous system. *Proc. Natl. Acad. Sci. U. S. A.*

111, E3005–14 (2014).

105. Hua, Z. L., Smallwood, P. M. & Nathans, J. Frizzled3 controls axonal development in distinct populations of cranial and spinal motor neurons. *Elife* **2**, e01482 (2013).
106. Perez-Garcia, V. *et al.* BAP1/ASXL complex modulation regulates epithelial-mesenchymal transition during trophoblast differentiation and invasion. *Elife* **10**, (2021).

CHAPTER 3:

Dissecting the Impact of *ASXL1* Mutations on
Neurodevelopmental Processes in Bohring-Opitz syndrome

Abstract

Bohring-Opitz syndrome (BOS) is caused by truncating mutations in the *ASXL1* (additional sex combs like 1) gene and is characterized by significant developmental delays. This study aims to elucidate the molecular and cellular mechanisms underlying BOS by examining the expression of *ASXL* gene family members and conducting a comprehensive transcriptomic analysis across different tissue types derived from BOS patients. Our findings reveal that truncated *ASXL1* transcripts do not undergo nonsense-mediated decay but highlight a potential compensatory role played by *ASXL3* in neural progenitor cells (NPCs). Furthermore, significant transcriptomic alterations in induced pluripotent stem cells (iPSCs) and NPCs indicate disrupted neurodevelopmental processes and suggest a role for these genes in the pathology of BOS. Additionally, gene ontology analysis linked significant differentially expressed genes in BOS NPCs to crucial neurodevelopmental pathways, such as axogenesis and pattern specification. We also observed pronounced defects in neural crest development driven by significant downregulation of the *ZIC1* and *ZIC4* genes in BOS neural lineages, providing insights into the developmental anomalies associated with the syndrome. This study provides a foundation for understanding the complex interplay between genetic mutations and neurodevelopmental dysregulation in BOS, offering potential targets for future therapeutic interventions.

Introduction

Bohring-Opitz syndrome (BOS; OMIM #605039) is a classic example of how genetic mutations can disrupt normal human development, leading to a spectrum of physical, developmental, and neurological challenges. Characterized by distinctive facial features, growth delays, and a range of congenital anomalies, BOS is most commonly associated with truncating mutations in the *ASXL1* gene, a crucial regulator of epigenetic modifications and gene expression. Despite its clear genetic linkages, the pathophysiological mechanisms through which *ASXL1* mutations lead to the varied manifestations of BOS remain poorly understood, highlighting a pressing need to examine the developmental timeline of dysregulatory effects.

ASXL1 is highly expressed in progenitor cells of the developing brain. Recent single cell RNA-sequencing studies of mouse organogenesis have clearly identified cell-type specific expression of *ASXL1*¹ which is most highly expressed across early neural progenitors and glial subtypes².

Cellular models offer unparalleled insights into human developmental biology, enabling the recapitulation of disease progression from the earliest stages of embryogenesis. Induced Pluripotent Stem Cells (iPSCs), derived from patient somatic cells, can be differentiated into a myriad of cell types, including Neural Progenitor Cells (NPCs) and neural crest cells (NCCs), providing a versatile platform for studying the cellular and molecular underpinnings of rare diseases such as BOS.

Patient-derived iPSCs allow for the modeling of human genetic diseases by capturing the pluripotent state of early development, crucial for studying disorders with developmental origins³. Moreover, iPSCs capture the entire genetic makeup of the individual from whom they are derived, including any disease-specific mutations. This aspect is critical for understanding how genetic abnormalities contribute to disease pathology. For example, iPSCs can be used to model neural tube defects by differentiating them into NPCs and NCCs, allowing researchers to study the molecular and cellular consequences of genetic mutations observed in patients.

NPCs, which give rise to neurons and glial cells, and NCCs, are critical for the development of many structures including the peripheral nervous system and facial cartilage. NPCs serve as a model to examine the early stages of neural development and the impact of genetic mutations on neurodevelopmental processes, and are thus particularly relevant for understanding the neurological and craniofacial abnormalities seen in BOS.

The neural crest and its implications in vertebrate development and Bohring-Opitz syndrome

Similarly, NCCs, derived from the transient and highly migratory population in vertebrate embryos, play a significant role in neurodevelopment. Recent studies using patient-derived iPSCs identified *ASXL1* as a critical regulator of neural crest development, with mutations perturbing the NC regulatory network and resulting in impaired neural crest delamination and emigration ⁴. Many of the tissues and anatomical structures that the NC contribute to overlap with those that are disrupted in BOS, such as cardiac septal defects, cleft palate and lip, and seizure ⁵.

The neural crest is a pivotal component in the evolutionary history and diversification of vertebrates. These transient, multipotent cells originate from the dorsal aspect of the neural tube and embark on a remarkable journey, migrating to various regions of the embryo. This migration is intricately timed with the closure of the neural tube, a process essential for proper embryonic development. The NC cells give rise to a diverse array of cell types, including neurons, glia, cartilage, and melanocytes, underscoring their critical role in forming the complex structures necessary for vertebrate life ^{6,7}.

Central to the process of NC cell migration and differentiation are the signaling pathways that guide these cells to specific locations and functions. Among these, the planar cell polarity (PCP) pathway and canonical Wnt signaling are paramount. The PCP pathway, involving core components such as Frizzled, Dishevelled, Prickle, and Vang-like (VANGL1/2) proteins, regulates cell-cell adhesion and is crucial for tissue morphogenesis and neural tube closure ⁸.

Concurrently, Wnt/ β -Catenin signaling plays a vital role in NC induction, lineage specification, and differentiation, triggering the expression of key transcriptional factors necessary for the transition from neural borders to distinct NC cell types⁹. Intriguingly, multi-omics studies on BOS patient-derived fibroblasts and blood samples have identified an aberrant upregulation of canonical Wnt and PCP pathways at the epigenetic, transcriptomic, and protein levels¹⁰.

In BOS, recent findings have shown significant dysregulation of VANGL2 in patient-derived samples¹⁰. VANGL2, a core component of the planar cell polarity (PCP) pathway, is essential for establishing cell polarity and guiding tissue morphogenesis, particularly during neural tube closure. This protein's function is so critical that disruptions in its signaling pathway have been directly linked to developmental anomalies, including neural tube defects that are clinically significant in human pathologies^{11,12}. This suggests a direct link between PCP pathway disruption and the neural crest related developmental anomalies observed in BOS^{5,10}.

The Zic gene family and its implications in neural crest function and Bohring-Opitz syndrome

The Zic gene family, comprising five transcription factors with zinc finger domains in mice and humans, plays a pivotal role in early embryonic development, including neural crest induction and neurulation^{13,14}. The *Zic* gene family (Zinc finger protein of the cerebellum) regulates various developmental processes, from cell fate determination to cranial development, to negative regulation of the Wnt pathway^{15,16}. Notably, mutations in the *ZIC* gene family have been implicated in severe brain malformations and neural tube closure defects^{17–21}, conditions that mirror some clinical aspects of BOS. In particular, *ZIC1* and *ZIC4* heterozygous deletion are responsible for the Dandy-Walker malformation cerebellar phenotype in deletion 3q patients²¹. The role of Zic genes in neural crest development suggests a possible pathway through which BOS manifests, highlighting the critical need to study these genes within this specific developmental context.

Rationale for this study

Here, we delve into the molecular intricacies of BOS by exploring the transcriptional dysregulation in patient-derived iPSCs and NPCs, as well as the functional impacts of these gene regulatory networks. Through detailed studies of these cells and integration with our previous study in differentiated cells derived from BOS patients, we seek to illuminate the developmental trajectories altered by *ASXL1* mutations, offering insights into the molecular cascades disrupted in this syndrome.

Methods

Patient recruitment

Study approval. This project was approved by the UCLA IRB# 11-001087. In conjunction with Dr. Bianca Russell and the ASXL Biobank at the University of California, Los Angeles, written informed consent from participants and guardians were obtained for skin punch biopsies. Deep phenotyping and review of medical records was also collected.

Sample collection and derivation

All individual level-data was de-identified prior to analysis and samples were collected for follow-up experiments. We collected blood from 12 BOS individuals and 11 controls, and skin punch biopsies from 7 individuals and 7 controls. Patient-derived skin fibroblasts were generated by the Translational Pathology Core Laboratory at UCLA. Blood and fibroblasts were processed as per methods described in ¹⁰.

A subset of the patient-derived fibroblasts (n=2) were reprogrammed by the Cedars-Sinai Core to generate induced pluripotent stem cells (iPSCs) (**Figure 3-1A**). iPSC controls were purchased from Cedars-Sinai Core (n=3). Neural progenitor cells (NPCs) and neural crest models were differentiated from these iPSCs using StemDIFF Neural Induction protocol and 2D neuroectoderm protocol from the Fu lab respectively.

Cell culture.

General conditions:

All cell lines were cultured in a controlled environment at 37°C with 5% CO₂. The following specific conditions were applied for each cell type:

Fibroblasts were maintained in Dulbecco's Modified Eagle Medium (DMEM, Gibco™, Cat No. 11-995-073) supplemented with 10% fetal bovine serum (FBS), 1% Non-essential Amino Acids (100X, Gibco™, Cat No. 11140-050), and 1% Penicillin-Streptomycin (Pen/Strep). Cells were cultured in T75 flasks.

iPSCs were cultured in mTeSR™1 medium (STEMCELL Technologies, #85850) atop a Vitronectin-coated surface (Invitrogen, #A14700) in 6-well plates to support pluripotency.

NPCs were maintained in StemDIFF™ Neural Progenitor Medium (STEMCELL Technologies, #05839) on a Vitronectin-coated surface (Invitrogen, #A14700) in 6-well plates to promote progenitor cell growth.

NC models were cultured on vitronectin-coated coverslips in neural induction medium enriched with dual SMAD inhibitors: SB 431542 (10 µM, Sigma Aldrich, Cat No. S4317-5MG), a TGF-β inhibitor, and LDN 193189 (500 nM, Sigma Aldrich, Cat No. SML0559-5MG), a bone morphogenetic protein 4 (BMP4) inhibitor ²².

Neural crest induction and differentiation

These models were developed according to established protocol ²². In brief, circular patterns of 400 µm diameter were created using silicon stamps provided by the Fu Lab and used to spin-coat #1.5 thickness coverslips (Thomas Scientific, #1217N79) with vitronectin (20 µg/ml; Trevigen). A single cell suspension from each iPSC line was plated onto the patterned coverslips at a density of 20,000 cells per cm². Neural induction was initiated using the aforementioned neural induction medium containing dual SMAD inhibitors for a period of 9 days.

Additionally, CHIR 99021 (3 μ M; a WNT pathway activator and β -catenin stabilizer) was supplemented on Days 3 and 4 to promote neural plate border cell specification ²².

Neural crest immunofluorescence.

Fixation and permeabilization:

Coverslips with neural crest models were fixed in 4% paraformaldehyde (PFA) at room temperature (RT) for 10 minutes. After fixation, coverslips were washed with phosphate-buffered saline (PBS). Coverslips were incubated in 0.1% Triton X-100 in PBS for 5 minutes at RT to permeabilize cells. Following permeabilization, coverslips were washed again with PBS.

Antibody incubation:

We used a blocking solution of 2.5% FBS in PBS to reduce non-specific binding. Primary antibodies were diluted in the blocking solution as follows: Nestin (Cell Signaling, #33475, 1:100) at a dilution of 1:100, PAX6 (Abcam, #ab78545, 1:200) at a dilution of 1:200, ZIC1 (Novus, #NB600-488, 1:200) at a dilution of 1:200. 50 μ L of primary antibody dilution was applied to a piece of parafilm taped to the bench per sample. Coverslips were flipped cell-side down and incubated for an hour at RT before washing 3x in blocking solution.

Secondary antibodies were diluted in blocking solution as follows: Anti-mouse AF555 (Invitrogen, #PIA32727, 1:400), Anti-rabbit AF488 (Invitrogen, #PIA32731TR, 1:400). 50 μ L of secondary antibody dilution was applied to a piece of parafilm taped to the bench per sample. Coverslips were flipped cell-side down and incubated for an hour at RT shielded from light before washing 3x in blocking solution.

Mounting:

A drop of Invitrogen™ ProLong™ Glass Antifade Mountant with NucBlue™ Stain (ThermoFisher, #P36985) was added to microscope slides just before mounting the coverslip cell-side down. Slides were cured overnight at 4°C before imaging.

RNA-sequencing and analysis.

All cell lines were grown to 80-90% confluency before total RNA extraction via RNeasy Plus Mini Kit (Qiagen, #74136). All samples were then prepared in technical triplicate with QIAseq FastSelect -rRNA HMR (Qiagen, #334386) and Illumina Truseq® Stranded Total RNA Library Prep Gold (Illumina, 20020599), before sequencing to a minimum of 30 million paired end 150bp reads per sample. A detailed protocol and analysis are described ¹⁰.

Results

Methodological overview of sample collection and cellular reprogramming

To study stem cell differentiation during early development in the genetic background of BOS, we employed a comprehensive approach to examine the transcriptomic and functional dysregulation driven by truncating *ASXL1* mutations.

Skin punch biopsies were collected from patients and controls to generate patient-derived fibroblasts, followed by reprogramming of fibroblasts into induced pluripotent stem cells (iPSCs) (**Figure 3-1A**). We derived and characterized two BOS-patient iPSCs harboring truncations at p.G642*, G826Efs*12 (**Figure 3-2A**), and controls lines through the iPSC Core at Cedars Sinai Medical Center in Los Angeles. These iPSC lines were differentiated to NPCs using the StemDIFF Neural Induction protocol and also to a neural crest model. These cell types, and blood samples collected from patients, were assayed with RNA-sequencing in technical triplicates (**Figure 3-1A**).

The detailed schematic of the *ASXL1* gene highlights the truncating mutations specifically identified in this patient cohort across various tissue types including fibroblasts, blood, iPSCs, and NPCs (**Figure 3-1B**).

Principal Component Analysis (PCA) of the transcriptomic data from BOS patient-derived samples revealed distinct clustering by tissue type (PC1 explaining 68% of

variance) and partial segregation of BOS and control samples within each tissue (PC2 explaining 25% of variance), suggesting a strong tissue-specific transcriptional identity and disease-specific transcriptional signatures (**Figure 3-2A**).

Key genes and differential expression in iPSCs: insights into BOS pathophysiology

In iPSCs, we identified 73 significantly upregulated and 54 downregulated differentially expressed genes (DEGs) (**Figure 3-2B and Table 3-1**). Noteworthy among the upregulated genes was *Maternally Expressed Gene 3 (MEG3)*, an imprinted long non-coding RNA that plays a crucial role in cellular differentiation and has been implicated in osteogenesis. Another upregulated gene, *Nuclear Factor I/B (NFIB)*, is essential for brain development, specifically in myelination and neurogenesis.

Conversely, *Salt-Inducible Kinase 1 (SIK1)* showed significant downregulation. SIK1 interacts with histone deacetylases and the *TORC (transducer of regulated CREB activity)* signaling pathways, suggesting its potential role in BOS pathology through modulation of gene expression and stress response pathways.

Key genes and differential expression in NPCs: functional insights into BOS pathophysiology

In NPCs, we identified 1660 upregulated DEGs and 2241 downregulated DEGs in BOS patient-derived samples (**Figure 3-2C and Table 3-1**). Gene expression analysis showed dramatic downregulation of crucial neural crest genes *ZIC1* and *ZIC4*, both of which play significant roles in early neural development and have been implicated in craniofacial formation and neural tube closure. Also downregulated were *NFIB*, which supports neural development, and *Gremlin 2 (GREM2)*, a BMP antagonist involved in embryonic development and tissue differentiation. Conversely, *VANGL2*, part of the planar cell polarity pathway that is crucial for neurodevelopmental processes, was found to be upregulated.

Previous studies examining iPSC-derived neural crest cells with truncating ASXL1 mutations also identified significant changes in *ZIC1* and *ZIC4*, highlighting their role in neural

crest defects (Matheus, 2019). Additionally, our previous study on BOS patient-derived fibroblast and blood samples also noted the significant impact of truncating *ASXL1* mutations on *GREM2* and *VANGL2* dysregulation, emphasizing the cross-tissue and developmental effects of these mutations ¹⁰.

The gene ontology analysis of significant DEGs in NPCs has linked these genes to several key developmental processes, providing deep insights into the pathophysiology of BOS. Among these processes, pattern specification (Gene Ontology: GO:007389, padj 8.46E-08) is particularly notable, as it relates to the basic body layout and the precise positioning of neural structures—a function directly influenced by *ASXL1* through its regulation of *Homeobox (HOX)* genes, which are pivotal in the anterior-posterior body patterning.

Moreover, the ontology highlighted the significance of axogenesis (Gene Ontology: GO:0007409, padj 8.78E-08), the process by which axons form during neurodevelopment. Axogenesis is critical for the establishment of neural circuits and proper brain function, making its disruption a likely contributor to the neurological symptoms observed in BOS patients. The development of specific brain regions was also prominently featured in the ontology results, including forebrain development (Gene Ontology: GO:0030900, padj 4.90E-07) and telencephalon development (Gene Ontology: GO:0021537, padj 6.92E-07). The forebrain is responsible for a variety of complex behaviors and cognitive functions, and its development is crucial for overall brain structure and function. The telencephalon, part of the forebrain, plays a key role in motor and sensory activities and cognitive abilities.

Implications of *ASXL* gene family expression in BOS: comparing across tissues and conditions

ASXL1 expression did not differ significantly between BOS and control samples across all tissues (**Figure 3-2E**). These, and our previous findings of germline levels of the wild type and the truncated *ASXL1* allele in BOS patient derived samples across different tissues ¹⁰, suggest that nonsense-mediated decay (NMD) is not a contributing factor in BOS.

However, *ASXL3* was notably decreased in BOS NPCs (**Figure 3-2F**), pointing to a specific pathway through which *ASXL1* truncating mutations may exert their effects, potentially mediated by *ASXL3* in neurodevelopmental processes. There was no significant difference in *ASXL3* expression in BOS in any of the other cell types.

Neural crest specification and defects in Bohring-Opitz syndrome derived models

We utilized a 2D neural crest model to investigate early specification events in BOS. Developed by Dr. Jianping Fu, this model incorporates pre-patterned geometrical confinement to induce emergent patterning of neuroepithelial and neural plate border (NPB) cells (**Figure 3-3A**) (Xue, 2018). The process and timeline of induction is depicted in **Figure 3-3A**. This induction is illustrated in **Figure 3-3B** with the pre-patterned coverslips and micropatterned colonies, and neural crest colonies at Days 2, 5, and 7 imaged at 40X magnification, showing the transition from columnar to cuboidal cell shapes with rounded nuclei at the colony periphery—indicative of neuroectoderm thickening similar to in vivo neural induction by Day 7.

Control colonies successfully undergoing neural crest formation were stained on Day 9 with ZIC1 and PAX6, alongside DAPI, to examine neural plate border and neuroectoderm specification respectively. The localisation of PAX6+ cells at the colony center and ZIC1+ cells to the colony periphery indicated successful induction and patterning (**Figure 3-3C**). Conversely, BOS-derived neural crest models exhibited profound specification defects, displaying impaired attachment and disorganized growth by Day 7, and failure to maintain a monolayer structure (**Figure 3-3D**). Unlike controls that preserved a distinct neural plate border even past the typical collection timepoint of Day 9, BOS colonies involuted into a disorganized cell mass, highlighting the severe impact of the disease on cellular organization and stability, and inhibiting our ability to characterize these cells by immunofluorescence.

Our findings correlate with those by Matheus et al. (2019), who noted defects in neural crest lamination and emigration in BOS due to *ASXL1* mutations ⁴. Their study identified

reduced activation of transcription factor *ZIC1* and *ZIC4*, which are critical in the NC gene regulatory network. Our transcriptomic analysis further underscores this, showing significant downregulation of *ZIC1* (log2FC -5.92, padj 3.21E-52) and *ZIC4* (log2FC -5.28, padj 7.92E-46) in BOS NPCs (**Figures 3-3E and 3-3F**). This specific defect in neural specification in BOS NPCs likely contributes to the failed development of a neural plate border in the disease models.

Comparative transcriptomic analysis in Bohring-Opitz syndrome across multiple tissue types identify potential biomarkers

We undertook a comprehensive transcriptomic comparison using patient-derived iPSC, NPC, blood, and fibroblast samples to understand the gene expression profile in BOS. Significant upregulation and downregulation of genes were identified across these tissues (**Table 3-1**). A crucial intersection of DEGs was identified, with 441 DEGs significantly regulated in two or more tissues, demonstrating a broad impact of BOS across different cell types (**Figure 3-4A**). Furthermore, 19 genes showed significant differential regulation across three of the four tissues. Notably, 10 of these genes exhibited significant dysregulation specifically in NPCs, blood, and fibroblasts but not in iPSCs, highlighting differentiation-associated genetic changes in BOS (**Table 3-2**).

We verified iPSC and NPC types using gene expression markers, such as *SRY-box transcription factor 2* (*SOX2*) and *Nestin* (*NES*), complemented by immunofluorescence staining (**Figures 3-4B and 3-4C**). These findings corroborate the cell identity and showed no significant differences between BOS and control samples in these specific markers, indicating accurate cell type identification and consistent baseline expression of pluripotency and neural lineage markers.

Our analysis particularly highlighted *VANGL2* and *GREM2*; *VANGL2* was significantly upregulated, while *GREM2* was significantly downregulated across NPC, blood, and fibroblast

samples (**Figures 3-4D and 3-4E, Table 3-2**). Previously, our research demonstrated that the dysregulation of these genes in BOS is driven by epigenetic alterations ¹⁰. This consistent dysregulation across various tissues suggests that *VANGL2* and *GREM2* may serve as potential biomarkers for identifying BOS in both progenitor and differentiated cell types.

Discussion

Expression of ASXL genes and implications for BOS pathophysiology

The expression of *ASXL1*, *ASXL2*, and *ASXL3* in various tissue types from BOS patients provides critical insights into their roles in the syndrome's development. Notably, *ASXL1* expression levels did not significantly differ between BOS and control samples across all tissues (**Figure 3-2E**). This observation aligns with findings from Lin et al. (2023), indicating that NMD is not a contributing factor in BOS ¹⁰. However, a marked decrease in *ASXL3* expression was noted specifically in BOS NPCs (**Figure 3-2F**), underscoring a potentially unique role of *ASXL3* in neurodevelopmental processes associated with BOS. *ASXL3*, as an epigenetic regulator, influences chromatin remodeling and gene expression essential for neural cell fate determination and development ^{23,24}. The clinical manifestations of *ASXL3*-related syndromes, such as Bainbridge-Ropers syndrome, include intellectual disability and developmental delays, further supporting the critical role of *ASXL3* in neural development ²⁵.

Differential expression in iPSCs and NPCs and clinical implications

Significant transcriptional changes were observed in BOS-derived iPSCs and NPCs, highlighting a dysregulated core transcriptional program. *MEG3*, a neuron-specific long noncoding RNA, was notably upregulated in BOS iPSCs and is known to induce necroptosis in human neurons, indicating its potential involvement in neurodegeneration observed in Alzheimer's disease ²⁶. Additionally, *NFIB*, critical for brain development and myelination, displayed altered expression levels in BOS, being upregulated in BOS iPSCs and

downregulated in NPCs ²⁷. *NFIB*'s role is further complicated by its expression across various tissues and its implication in different neurodevelopmental disorders and cancers, suggesting a multifaceted role in cellular development and pathology ^{28,29}.

Gene ontology analysis of NPCs from BOS patients revealed significant links to developmental processes, notably pattern specification and axogenesis, which are crucial for the correct positioning of neural structures and axon formation, respectively and overall brain function. The dysregulation of these pathways likely contributes directly to the neurological manifestations of BOS, including developmental delays and cognitive impairments.

Neural crest impairment and aberrant downregulation of *ZIC1* and *ZIC4* in BOS

The observed defects in neural crest development in BOS models provide profound insights into the developmental anomalies associated with the syndrome. Significant reductions in *ZIC1* and *ZIC4* expression in BOS-derived NPCs highlight their pivotal role in neural crest and early neural development (**Figure 3-3D and 3-3E**). These genes are essential for the specification and migration of neural crest cells, with their dysregulation potentially leading to the development of craniofacial anomalies and neurological disorders observed in BOS patients. The impact of *ZIC1* and *ZIC4* dysregulation extends to conditions like Dandy-Walker malformation, where these genes play critical roles in cerebellar development and size regulation, mirroring some of the structural brain abnormalities seen in BOS ^{4,21}.

In conclusion, this study provides a comprehensive overview of the genetic and cellular disruptions caused by *ASXL1* mutations in BOS. Our analyses suggest that while *ASXL1* itself does not undergo significant expression changes, the downstream effects mediated through *ASXL3* and changes in other critical genes like *MEG3* and *NFIB* contribute substantially to the syndrome's pathology. The specific downregulation of neurodevelopmentally important genes, such as *ZIC1* and *ZIC4*, in BOS-derived neural progenitor cells further elucidates the impact of these mutations on neural crest development and the resultant neurodevelopmental defects.

These findings underscore the complexity of BOS and suggest that future research should focus on the detailed mechanisms through which *ASXL1* mutations affect neural development. Ultimately, understanding these pathways may provide novel insights into potential therapeutic strategies aimed at mitigating the severe developmental and neurological symptoms observed in patients with BOS. This study not only advances our understanding of BOS pathology but also highlights the critical role of epigenetic and transcriptional regulation in human neurodevelopmental disorders.

Figures and Tables

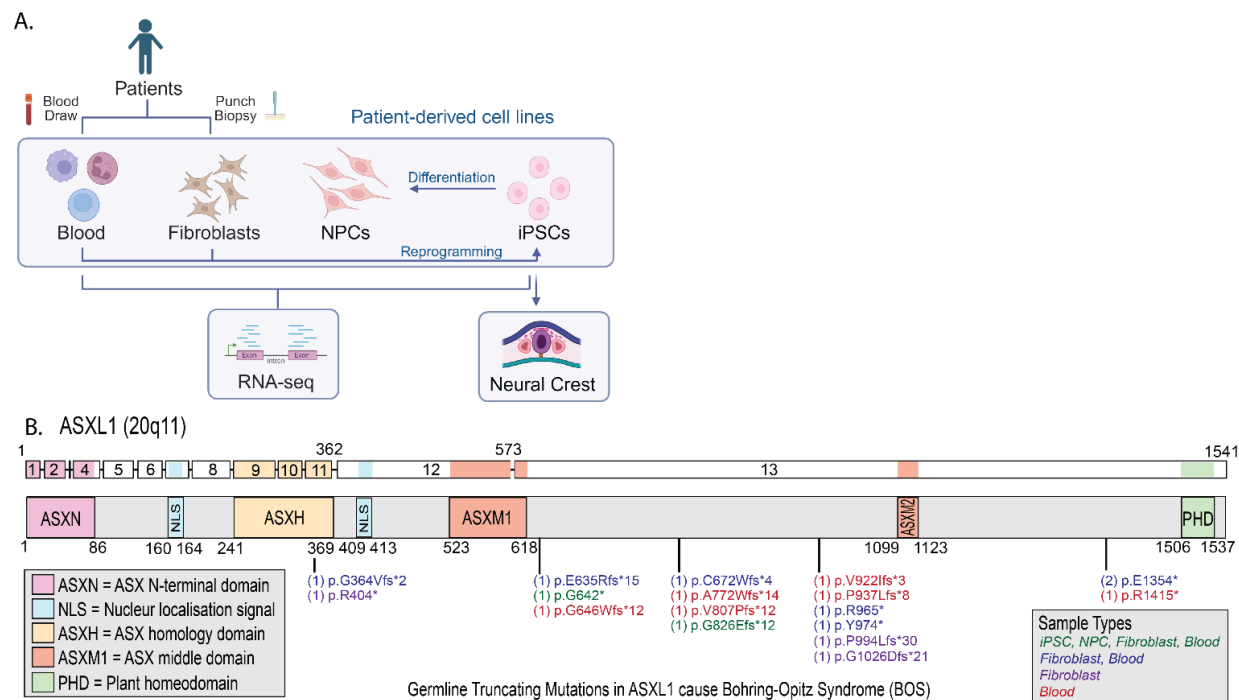


Figure 3-1: Investigating genetic and cellular mechanisms in Bohring-Opitz syndrome

(A) Workflow illustrating sample collection to analyses in Bohring-Opitz syndrome (BOS).

(B) Schematic of the *ASXL1* gene showing exons and protein domains. Truncating mutations identified in BOS patient samples in this cohort are highlighted, with distinctions made across different sample types including fibroblasts, blood, iPSCs, and NPCs.

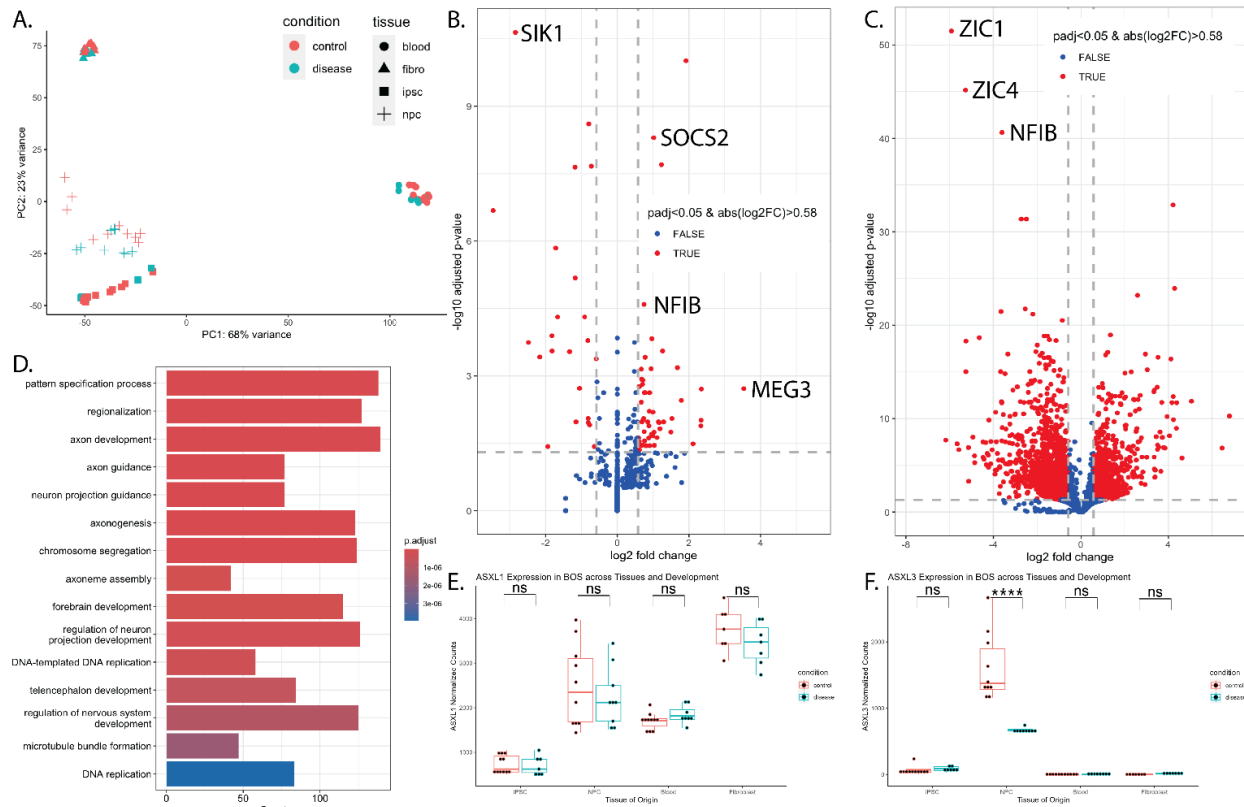


Figure 3-2: Molecular profiling of Bohring-Opitz syndrome via RNA-seq analysis

(A) Principal component analysis (PCA) of RNA-seq data from iPSCs, NPCs, blood, and fibroblast samples. PC1 accounts for 68% of the variance, distinguishing between tissue types, while PC2 (25% variance) separates control and BOS samples within each tissue category.

(B) Volcano plot displaying differential gene expression between BOS compared to control iPSCs. Key genes include upregulation of MEG3 and NFIB and downregulation of SIK1 and ZNF667.

(C) Volcano plot illustrating differential gene expression in BOS compared to control NPCs, with notable downregulation of ZIC1, ZIC4, NFIB, and GREM2, and upregulation of VANGL2.

(D) Gene ontology analysis of significant DEGs in BOS NPCs, highlighting involvement in pattern specification, axon development, neuron projection guidance, and forebrain development.

(E) Boxplot depicting *ASXL1* gene expression across different sample types (iPSC, NPC, blood, fibroblast) and groups (control in red, BOS in blue). No significant differences in transcript levels were identified between BOS and control across the sampled tissues.

(F) Boxplot depicting *ASXL3* gene expression across different sample types (iPSC, NPC, blood, fibroblast), showing a significant reduction in *ASXL3* expression in BOS NPCs, suggesting a potential mechanism for neurodevelopmental defects observed in BOS through *ASXL3* depletion. No significant differences in transcript levels were identified between BOS and control across the other sampled tissues.

ns or no stars denote $p\text{-value} > .05$, * denote $p\text{-value} \leq .05$, ** denote $p\text{-value} < .01$, *** denote $p\text{-value} < .001$, **** denote $p\text{-value} < .0001$

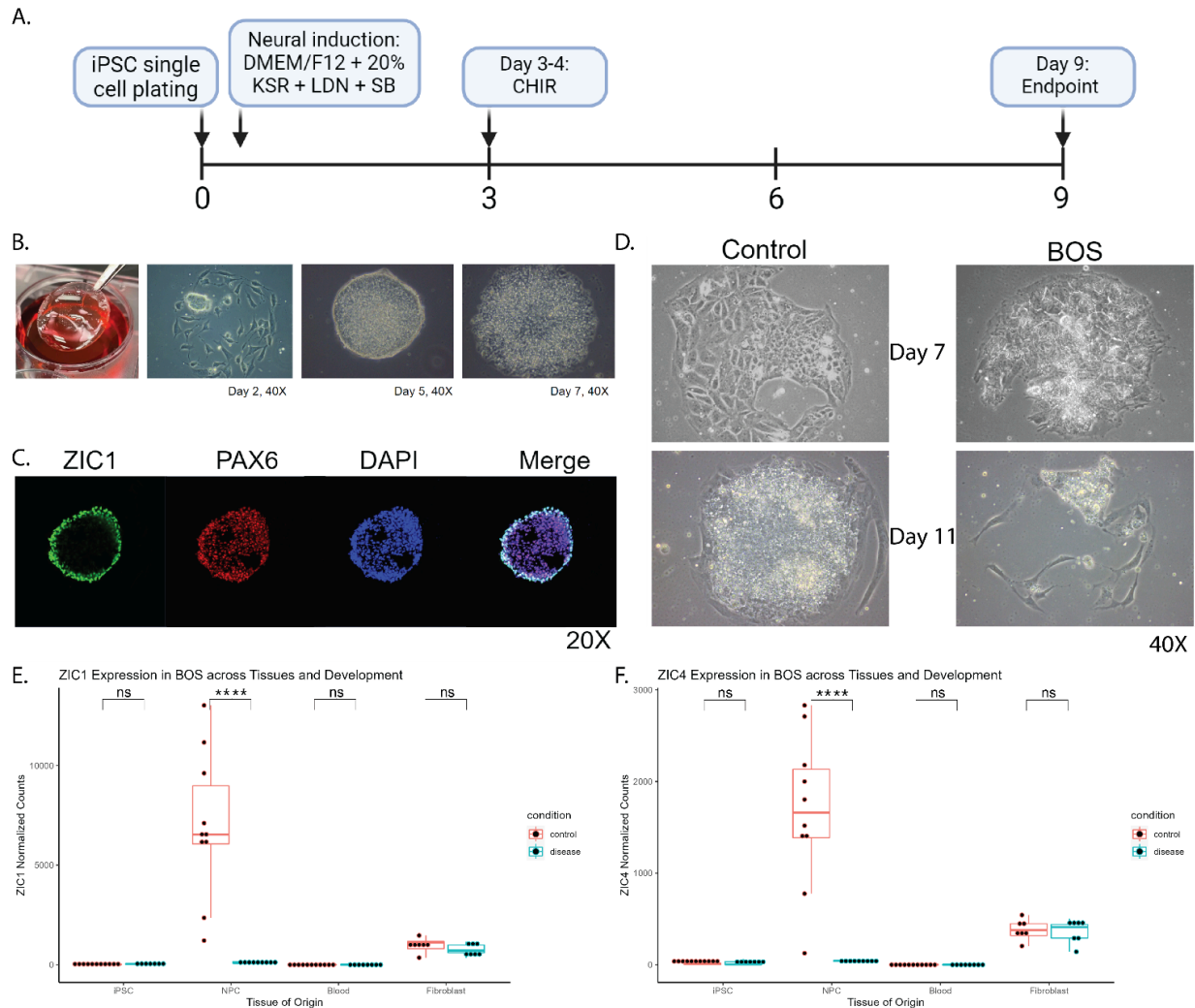


Figure 3-3: Neural crest specification and defects in Bohring-Opitz syndrome derived models

(A) Timeline of neural crest induction in the 2D model showing emergent patterning of neuroepithelial and NPB cells.

(B) Micrographs of micropatterned colonies on pre-patterned coverslips, illustrating the morphological transitions of iPSC cells to neuroectoderm cells with a neural plate border through day-wise progression (Days 2, 5, and 7). Colonies were imaged at 40X magnification.

(C) Immunofluorescence staining of control colonies at Day 9 for neural plate border (ZIC1) and neuroectoderm (PAX6) markers, with cellular localizations indicative of successful specification. Colonies were imaged at 40X magnification.

(D) Control and BOS-derived neural crest models at Day 7 and Day 11 showing disorganized growth and inability to maintain monolayer formation, contrasting sharply with control models.

Transcriptomic analysis of BOS-derived tissues showing significant downregulation of **(E)** ZIC1 and **(F)** ZIC4 in BOS-derived NPCs, critical for neural crest specification.

ns or no stars denote $p\text{-value} > .05$, * denote $p\text{-value} \leq .05$, ** denote $p\text{-value} < .01$, *** denote $p\text{-value} < .001$, **** denote $p\text{-value} < .0001$

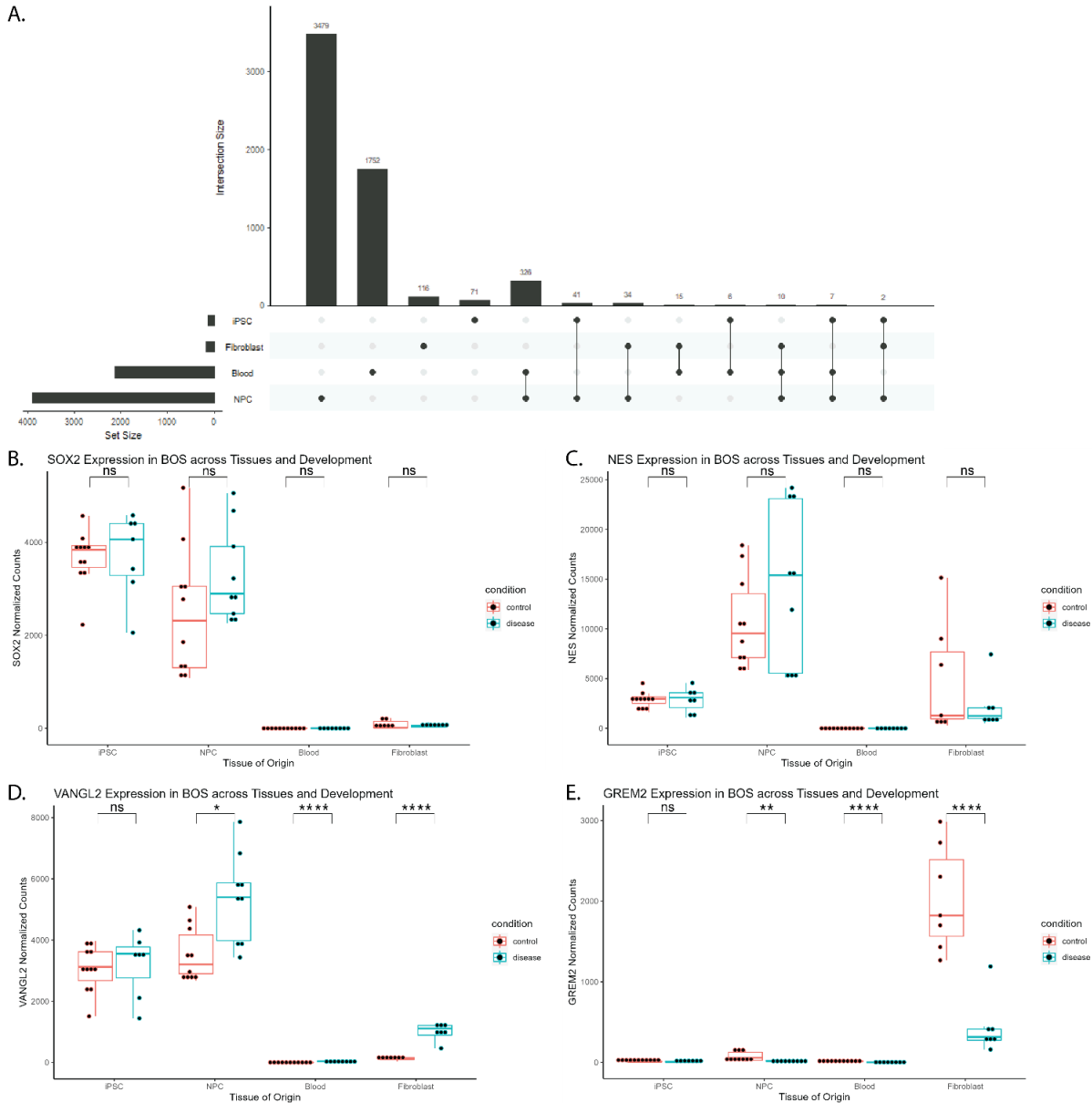


Figure 3-4: Comparative transcriptomic analysis in Bohring-Opitz syndrome across multiple tissue types identify potential biomarkers

(A) Intersection of differential gene expression across BOS-derived tissue types, illustrating widespread genetic dysregulation across tissues driven by truncating *ASXL1* mutations.

(B) Gene expression analysis of *SOX2* in iPSCs and NPCs, confirming the pluripotent and neural progenitor cell identity respectively.

(C) *NES* expression profile in NPCs, indicating no significant differential expression between BOS and control cells and supporting the maintenance of neural characteristics.

(D) Expression patterns of **(D)** *VANGL2* and **(E)** *GREM2* across iPSC, NPC, blood, and fibroblast tissues highlighting their significant upregulation and downregulation respectively in BOS across the latter three tissues.

ns or no stars denote $p\text{-value} > .05$, * denote $p\text{-value} \leq .05$, ** denote $p\text{-value} < .01$, *** denote $p\text{-value} < .001$, **** denote $p\text{-value} < .0001$

Tissue Type	Significant Upregulated Genes	Significant Downregulated Genes
iPSC	73	54
NPC	1660	2241
Blood	1166	952
Fibroblast	129	48

Table 3-1: Number of significant differentially expressed genes (DEGs) across tissues in BOS

Gene Name	Log2FC_iPSC	padj_iPSC	Log2FC_NPC	padj_NPC	Log2FC_Fibro	padj_Fibro	Log2FC_Blood	padj_Blood
AC005224.3	-0.57	1.36E-03	2.56	1.11E-09	1.49E-03	9.76E-01	0.44	8.08E-03
ADGRF3	0.60	0.02	-0.78	1.95E-02	-2.35E-04	9.98E-01	0.68	5.51E-03
CAND2	-4.30E-06	1.00	-0.51	2.05E-03	1.38	1.61E-07	0.65	2.83E-05
GREM2	-8.24E-07	1.00	-1.81	7.92E-04	-2.11	3.09E-06	-2.48	6.02E-06
IQCA1	-3.46E-05	0.04	-0.54	1.93E-02	0.83	8.56E-02	1.76	6.52E-05
KCNQ5	2.41E-06	1.00	-1.30	1.98E-02	-0.01	3.51E-02	0.65	6.02E-03
NIPAL3	0.52	2.37E-03	-0.92	2.35E-03	1.74E-03	9.70E-01	0.37	2.86E-02
NLRP2	-2.22E-06	0.01	-2.32	2.15E-03	-9.19E-04	NA	-0.68	4.42E-02
OBSCN-AS1	-2.67E-06	1.00	-0.43	4.22E-02	1.61	3.29E-04	0.74	8.77E-03
OTUD7A	-2.87E-06	1.00	-0.70	3.24E-03	0.86	7.48E-03	0.90	3.94E-06
PELI2	6.54E-06	1.00	0.79	3.64E-05	0.45	5.56E-03	-0.37	4.92E-02
PLBD1	8.56E-06	1.00	0.76	1.07E-03	1.06	7.52E-03	-0.50	2.39E-02
ROBO3	3.88E-07	1.00	-0.53	2.63E-03	0.80	7.48E-03	1.11	1.36E-04
SOCS2	1.02	5.01E-09	0.92	4.35E-02	-0.02	1.03E-02	-0.02	9.35E-01
STMN2	1.79	3.53E-03	-1.70	4.32E-04	-3.17	8.19E-04	0.00	NA
TRO	6.78E-09	1.00	-0.47	2.00E-02	0.60	2.59E-02	0.83	6.77E-03
VANGL2	1.76E-06	1.00	0.52	1.06E-02	2.55	1.25E-10	3.80	2.25E-13
ZNF667	2.34	0.01	2.37	9.64E-05	1.36E-03	9.90E-01	1.09	7.67E-07
ZNF667-AS1	2.11	0.03	2.41	1.87E-04	-4.02E-03	9.49E-01	1.01	3.18E-05

Table 3-2: Common DEGs across BOS tissues that are significantly dysregulated in at least 3 tissues

References

1. Cao, J. *et al.* The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566**, 496–502 (2019).
2. Home :: BrainSpan: Atlas of the developing human brain. <http://www.brainspan.org/>.
3. Ebert, A. D. & Svendsen, C. N. Human stem cells and drug screening: opportunities and challenges. *Nat. Rev. Drug Discov.* **9**, 367–372 (2010).
4. Matheus, F. *et al.* Pathological ASXL1 Mutations and Protein Variants Impair Neural Crest Development. *Stem Cell Reports* **12**, 861–868 (2019).
5. Russell, B., Tan, W.-H. & Graham, J. M., Jr. Bohring-Opitz syndrome. in *GeneReviews®* (eds. Adam, M. P. *et al.*) (University of Washington, Seattle, Seattle (WA), 2018).
6. Kulesa, P. M. & Gammill, L. S. Neural crest migration: patterns, phases and signals. *Dev. Biol.* **344**, 566–568 (2010).
7. Sauka-Spengler, T. & Bronner-Fraser, M. A gene regulatory network orchestrates neural crest formation. *Nat. Rev. Mol. Cell Biol.* **9**, 557–568 (2008).
8. Nagaoka, T., Katsuno, T., Fujimura, K., Tsuchida, K. & Kishi, M. Functional interaction between Vangl2 and N-cadherin regulates planar cell polarization of the developing neural tube and cochlear sensory epithelium. *Sci. Rep.* **13**, 3905 (2023).
9. Hari, L. *et al.* Temporal control of neural crest lineage generation by Wnt/ β -catenin signaling. *Development* **139**, 2107–2117 (2012).
10. Lin, I. *et al.* Multi-omics of Bohring-Opitz syndrome truncating ASXL1 mutations identify canonical and non-canonical Wnt signaling dysregulation. *JCI Insight* (2023) doi:10.1172/jci.insight.167744.
11. Galea, G. L. *et al.* Cell non-autonomy amplifies disruption of neurulation by mosaic Vangl2 deletion in mice. *Nat. Commun.* **12**, 1159 (2021).
12. Lei, Y.-P. *et al.* VANGL2 mutations in human cranial neural-tube defects. *The New England*

journal of medicine vol. 362 2232–2235 (2010).

13. Aruga, J. *Zic Family: Evolution, Development and Disease*. (Springer, 2018).
14. Aruga, J. *et al.* A novel zinc finger protein, zic, is involved in neurogenesis, especially in the cell lineage of cerebellar granule cells. *J. Neurochem.* **63**, 1880–1890 (1994).
15. Fujimi, T. J., Hatayama, M. & Aruga, J. *Xenopus* Zic3 controls notochord and organizer development through suppression of the Wnt/ β -catenin signaling pathway. *Dev. Biol.* **361**, 220–231 (2012).
16. Pourebrahim, R. *et al.* Transcription factor Zic2 inhibits Wnt/ β -catenin protein signaling. *J. Biol. Chem.* **286**, 37732–37740 (2011).
17. Nakata, K., Koyabu, Y., Aruga, J. & Mikoshiba, K. A novel member of the *Xenopus* Zic family, Zic5, mediates neural crest development. *Mech. Dev.* **99**, 83–91 (2000).
18. Inoue, T. *et al.* Mouse Zic5 deficiency results in neural tube defects and hypoplasia of cephalic neural crest derivatives. *Dev. Biol.* **270**, 146–162 (2004).
19. Elms, P., Siggers, P., Napper, D., Greenfield, A. & Arkell, R. Zic2 is required for neural crest formation and hindbrain patterning during mouse development. *Dev. Biol.* **264**, 391–406 (2003).
20. Purandare, S. M. *et al.* A complex syndrome of left-right axis, central nervous system and axial skeleton defects in Zic3 mutant mice. *Development* **129**, 2293–2302 (2002).
21. Blank, M. C. *et al.* Multiple developmental programs are altered by loss of Zic1 and Zic4 to cause Dandy-Walker malformation cerebellar pathogenesis. *Development* **138**, 1207–1216 (2011).
22. Xue, X. *et al.* Mechanics-guided embryonic patterning of neuroectoderm tissue from human pluripotent stem cells. *Nat. Mater.* **17**, 633–641 (2018).
23. Katoh, M. & Katoh, M. Identification and characterization of ASXL3 gene in silico. *Int. J. Oncol.* **24**, 1617–1622 (2004).
24. Srivastava, A. *et al.* De novo dominant ASXL3 mutations alter H2A deubiquitination and

- transcription in Bainbridge–Ropers syndrome. *Hum. Mol. Genet.* **25**, 597–608 (2015).
25. Schirwani, S. *et al.* Expanding the phenotype of ASXL3-related syndrome: A comprehensive description of 45 unpublished individuals with inherited and de novo pathogenic variants in ASXL3. *Am. J. Med. Genet. A* **185**, 3446–3458 (2021).
 26. Balusu, S. *et al.* Long noncoding RNA MEG3 activates neuronal necroptosis in Alzheimer’s disease. *bioRxiv* (2022) doi:10.1101/2022.02.18.480849.
 27. Schanze, I. *et al.* NFIB Haploinsufficiency Is Associated with Intellectual Disability and Macrocephaly. *Am. J. Hum. Genet.* **103**, 752–768 (2018).
 28. Consortium, T. G. *et al.* The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. *Science* **348**, 648–660 (2015).
 29. Fane, M. E. *et al.* NFIB Mediates BRN2 Driven Melanoma Cell Migration and Invasion Through Regulation of EZH2 and MITF. *EBioMedicine* **16**, 63–75 (2017).

CHAPTER 4

ASXL1 Mutations Drive Metabolic Dysregulation and Mitochondrial Dysfunction in Bohring-Opitz syndrome

Abstract

Bohring-Opitz syndrome (BOS, OMIM#605309) is a rare neurodevelopmental disorder caused by heterozygous truncating mutations in the *ASXL1* gene. Patients are identified at birth with significant neurodevelopmental impairments and distinctive craniofacial anomalies. While previous research has primarily focused on the genetic and molecular dimensions of BOS, our study reveals novel insights into the metabolic dysfunctions and mitochondrial defects characterizing this condition.

We conducted a comprehensive study using patient-derived fibroblasts to investigate the metabolic and mitochondrial consequences of *ASXL1* mutations in BOS. Our findings reveal a significant decrease in mitochondrial pyruvate carrier proteins (MPC1 and MPC2) in BOS fibroblasts compared to controls, coupled with reduced mitochondrial oxygen consumption rates (OCR). Metabolomic analysis highlighted a shift towards increased glycolysis and amino acid biosynthesis, indicative of a metabolic reprogramming in response to mitochondrial dysfunction. Further functional studies demonstrated a dependency on specific amino acids for cellular proliferation and migration.

This study provides novel insights into the metabolic disturbances driven by *ASXL1* mutations in BOS, illustrating significant mitochondrial dysfunction and compensatory metabolic reprogramming. These findings open new avenues for investigating targeted therapies and suggest broader implications for the pathophysiology of BOS and possibly other similar genetic neurodevelopmental disorders.

Introduction

The intersection between epigenetic regulation and metabolism represents a vital area of study in understanding cellular adaptation and disease pathogenesis. Epigenetic modifications, which alter gene expression without changing the DNA sequence, are crucial for cellular responses to environmental stimuli and maintaining identity. Conversely, metabolism encompasses all biochemical processes vital for life, supplying substrates and energy for epigenetic modifications. The dynamic interplay between these processes is evident in various diseases, including cancer, where metabolic and epigenetic reprogramming are fundamental to tumor survival under stress conditions ^{1,2}.

In this context, mutations in the *additional sex combs-like 1* (*ASXL1*) gene present a unique lens for exploring how disruptions in epigenetic regulation can lead to metabolic dysregulation. *ASXL1* is integral to the regulation of gene expression, and its mutations are associated with a range of hematological malignancies and a rare neurodevelopmental germline disorder, Bohring-Opitz syndrome (BOS). These mutations lead to altered chromatin states and dysregulated gene expression ³, and are thought to impact cellular metabolism and contribute to disease progression ⁴. The reciprocal relationship between metabolism and epigenetics suggests that metabolic changes induced by *ASXL1* mutations could further exacerbate epigenetic dysregulation, creating a feedback loop that promotes disease progression.

BOS is a rare genetic disorder first described in 1999 by Bohring et al., characterized by profound developmental delays, distinct craniofacial anomalies, growth retardation, and a spectrum of other clinical manifestations ⁵⁻⁷. Seizures and respiratory complications are common, severely impacting the quality of life for affected individuals ^{5,6}. The genetic underpinnings of BOS are primarily *de novo* mutations in the *ASXL1* gene, essential for chromatin remodeling and normal development. Despite advancements in genetic diagnostics, the rarity of BOS poses challenges in accumulating comprehensive research data, making each individual sample critical for understanding the disorder's complex nature. Current management

strategies focus on symptomatic care, highlighting the need for targeted therapies that could alleviate the diverse symptoms associated with BOS ^{5,6}.

Current understanding and research gaps in BOS

While significant strides have been made in understanding the neurodevelopmental and systemic manifestations of BOS, our knowledge of its metabolic landscape remains limited. There is a critical need for comprehensive studies to elucidate the molecular dysregulations underpinning BOS and other rare diseases. Our use of patient-derived samples has been instrumental in advancing our understanding of BOS. These cell models not only carry the genetic and molecular aberrations inherent to BOS but also allow for elucidating disease mechanisms, identifying potential therapeutic targets, and screening effective treatments through high-throughput methods.

Recent multi-omics studies on primary BOS patient samples identified significant upregulation of Vang-like 2 (VANGL2), which plays a critical role in cell polarity and directed cell migration ^{3,8}. VANGL2 dysregulation is linked to impaired cellular motility and tissue morphogenesis ⁸, neural crest closure and morphogenesis in mice ^{9,10} and in humans ¹¹, and impaired neural crest in BOS cell models ¹².

Role of ASXL1 mutations in mTOR signaling and metabolic dysregulation

The mechanistic target of rapamycin (mTOR) pathway is a central regulator of growth, proliferation, and metabolism, responding to external cues, nutrient availability, and cellular energy status ¹³. The mTOR pathway regulates glycolysis, a fundamental cellular process involved in energy production, by promoting expression of key glycolytic pathway enzymes. mTOR is a conserved serine/threonine kinase that exists in two distinct complexes: the mTOR Complex 1 (mTORC1) and mTOR Complex 2 (mTORC2) orchestrates a myriad of processes from protein synthesis to mitochondrial activity ^{14,15}. Upon activation of mTORC1, S6 kinase becomes phosphorylated (pS6K) and subsequently activates downstream targets involved in

ribosomal biogenesis and protein translation. This process is tightly regulated and plays a crucial role in maintaining cellular homeostasis by balancing the energy and nutrient availability with the cellular demand for protein synthesis ¹⁴.

Truncating germline mutations in *ASXL1* cause BOS, while somatic mutations in the hematopoietic stem cell drive acute myeloid leukemia (AML) progression. Fujino et al. (2021) and Youn et al. (2017) demonstrated the implications of *ASXL1* mutations on the Akt/mTOR pathway, suggesting a potential mechanism for the observed proliferative disorders in BOS ^{16,17}. Wild-type and truncated mutant *ASXL1* was found to bind to BRCA1 Associated Protein 1 (BAP1) and AKT serine threonine kinase 1 (AKT1) in both the cytoplasm and nucleus in HEK293T cells, suggesting complex formation ¹⁷. This disruption not only affects glycolysis and mitochondrial function but also influences cellular processes like growth and migration, essential for tissue development and repair.

The mTOR pathway is highly responsive to the availability of amino acids, particularly asparagine, and dysregulation can contribute to disease pathology. Asparagine is a key modulator of mTORC1 and activating transcription factor 4 (ATF4), a transcription factor that modulates the expression of genes involved in amino acid metabolism, antioxidant responses, and stress response ^{18,19}. When cellular levels of asparagine are depleted, it leads to the activation of the integrated stress response (ISR) pathway, a cellular stress response mechanism, and impairment of mTORC1 activity ¹⁸. This activation, in turn, induces the expression of *ATF4* ^{18,20} in order to increase *de novo* asparagine synthesis via asparagine synthetase (ASNS) ^{19,21}.

Objective and rationale for our study

Understanding the role of *ASXL1* mutations in both epigenetic regulation and metabolic disruptions is crucial for identifying novel therapeutic targets. Our research aims to dissect these interactions, providing insights that could lead to novel therapeutic strategies for BOS and

potentially other related disorders. By identifying key metabolic pathways affected by *ASXL1* mutations and exploring their link to disease pathology, our work seeks to pave the way for metabolic interventions that could enhance patient outcomes and quality of life.

Methods

Sample collection

We collected samples as outlined in Lin et al. (2023)³, which provides comprehensive details on the selection and characterization of cell lines. To ensure representative diversity, we incorporated both male and female subjects, and strove to include an equal ratio, to minimize the potential for gender-specific variations in our analysis (**Table 4-1**). Furthermore, we deliberately included cell lines derived from patients with varied genetic variants to capture the heterogeneity inherent in BOS. Additionally, we sampled from a broad age range, encompassing individuals across different age groups affected by BOS, thus ensuring that our analysis reflects the spectrum of ages typically observed in this condition. This enhances the robustness and relevance of our findings in elucidating the complexities of BOS pathogenesis.

Cell culture and media conditions

Fibroblasts were maintained under typical growth conditions, as previously described³. Medias used in each assay are described in detail (**Table 4-2**). Cells were maintained in Glucose media (+NEAA) unless otherwise specified. Cells used in galactose assays were cultured in Galactose media (+NEAA) for at least 10 days prior to assay. Catalogue numbers for each reagent are listed (**Table 4-3**).

Scratch wound assay and analysis

We used scratch wound assays to evaluate fibroblast growth and wound healing dynamics. Cells were seeded at a density of 40,000 cells per well in a 96-well Sartorius ImageLock Plate (Sartorius BA-04856) containing full growth media. Four hours post-seeding, a

scratch was introduced using the Incucyte® 96-Well Woundmaker Tool (Sartorius, Cat #4563). Following a gentle wash with PBS, cells were changed to and incubated in Glucose media (+NEAA), Glucose media (-NEAA) or Glucose media (+L-Asn), and the plate was imaged for 48 hours using an Incucyte SX5 system located within the UCLA Broad Stem Cell Research Center Microscopy Core Facility. Images were captured every 2 hours over the 48-hour period with phase-contrast microscopy at 10X wide mode objective. Subsequently, the data were analyzed using the Incucyte Scratch Wound Analysis Software Module (Sartorius 9600-0012).

The rate of wound closure was calculated by determining the change in wound area or distance between the wound edges over time. The closure rate was then expressed as the percentage of wound closure per unit time, allowing for the comparison of different experimental conditions or treatments. The data was then graphed using RStudio *ggplot2* to visualize the changes in wound closure over time and the differences between experimental conditions. The mean and standard deviation (SD) were calculated and plotted at each time point for each condition.

Western blots

Cell cultures were subjected to three distinct growth conditions: Glucose media (+NEAA), Glucose media (-NEAA), and Glucose media (+L-Asn). These conditions were maintained for predetermined durations of 16, 48, and 72 hours in 15 cm² culture dishes. After incubation, whole cell lysates were prepared as previously described ³, using a 10× Cell Lysis Buffer (Cell Signaling Technology 9803S) supplemented with Halt Protease and Phosphatase Inhibitor Cocktail (100X, Thermo Fisher Scientific, Catalog #78440). Lysates were quantified using the Pierce BCA Protein Assay Kit (Thermo Fisher Scientific 23225).

Western blotting was performed using standard protocols ¹⁹. The following commercial antibodies were used as probes: ASNS (Proteintech 14681-1-AP, 1:1000), ATF4 (Cell Signaling 11815, 1:500), phospho-T389 S6 kinase (Cell Signaling 9234, 1:500), S6 kinase (Cell Signaling

2708, 1:1000), phospho-S235/235 S6 ribosomal protein (Cell Signaling 4858, 1:3000), S6 ribosomal protein (Cell Signaling 2217, 1:1000), 4E-BP1 (Cell Signaling 9644, 1:500), MPC1 (Sigma HPA045119, 1:1000), MPC2 (Cell Signaling 46141S, 1:1000), LC3A/B (Cell Signaling 4108, 1:1000) and β -Actin (Cell Signaling 4970, 1:1000).

Protein bands were quantified using *ImageJ* software, facilitating the analysis of protein intensity. Data analysis was performed in RStudio, where protein intensity values were normalized against actin intensity values. Statistical evaluations, including pairwise t-tests, were employed to discern significant differences between the experimental conditions. Visualization of the results was performed using *ggplot2* in R. Boxplots depict the normalized protein intensity relative to actin and the fold changes of target protein levels as a result of nutrient depletion or replenishment compared to their baseline expression within the same sample.

Seahorse assay

Fibroblast cell lines were cultured under standard conditions and were seeded at 12,000 cells per well in an Agilent Seahorse 96 well XFe96 / XF Pro Microplate one day prior to analysis (Agilent 103794-100). Oxygen consumption rates (OCR) were completed on a Seahorse XF96 Analyzer. The Seahorse XF Analyzer was loaded with pre-calibrated sensor cartridges containing compounds targeting mitochondrial respiration: oligomycin (2 μ M round 1 and round 3, 2.5uM round 2), FCCP (3 μ M x2), and a mixture of rotenone (2 μ M) and antimycin A (2 μ M). The Seahorse assay was conducted according to the manufacturer's guidelines, measuring OCR to assess mitochondrial function and other variables to evaluate mitochondrial bioenergetics.

Mitochondrial imaging

To examine mitochondrial dynamics, we plated 12,000 cells per well in a 96 well plate (Corning Costar 3603). We used tetramethylrhodamine, ethyl ester (TMRE) to assess mitochondrial membrane potential, and Mitotracker Green (MTG) to estimate mitochondrial

mass. Initially, we added a carbonyl cyanide-4 (trifluoromethoxy) phenylhydrazone (FCCP) control to each well to induce depolarization of the mitochondrial membrane, serving as a crucial experimental control. Changes in mitochondrial membrane potential were quantified by measuring the fluorescence intensity of TMRE before and after FCCP treatment. This approach allowed us to analyze FCCP-induced depolarization effectively, providing insights into mitochondrial functionality under stress conditions.

RNA sequencing (RNA-seq)

Fibroblasts were cultured per established protocol ³ in Glucose media (+NEAA) or Glucose media (-NEAA). Cells were harvested, RNA was extracted, and RNA-seq was conducted and analyzed according to established protocol ³. RNA-seq data are deposited in the NCBI Gene Expression Omnibus database (GEO GSEXXXX).

Intracellular metabolite extraction and analysis

Fibroblast cells were seeded in six-well plates at 200,000 cells per well, in Glucose media (+NEAA) or Glucose media (-NEAA) accordingly. One day prior, media was changed and glucose tracing or glutamine tracing media were added, with or without NEAA as needed. Metabolites were extracted at 70–80% confluence per established protocol ¹⁸. Dried metabolites were stored at -80°C prior to mass spectrometry analysis. Analysis was conducted using the Christofk Lab Grapha platform.

Dot blot assays

Cell lysate

Cell lysate was prepared according to R&D Proteome Profiler Human Cell Stress Array Kit (R&D ARY178) and R&D Systems Proteome Profiler Human Ubiquitin Array Kit (R&D ARY027) protocols.

Cell stress assay

The evaluation of cellular stress responses was conducted using the R&D Systems Proteome Profiler Human Cell Stress Array Kit (R&D ARY178), which enables the simultaneous detection of 26 different stress-related proteins. The assessment of relative expression levels of stress-induced proteins in human samples was performed according to the manufacturer's guidelines. In brief, nitrocellulose membranes included in the array kit, each spotted with antibodies specific to different stress-related proteins, were incubated with prepared cell lysates overnight at 4°C on a rocking platform shaker. Following extensive washing steps to remove non-specific binding, the membranes were exposed to a biotinylated detection antibody cocktail targeting stress-related proteins for 1 hour. Subsequently, thorough washing was performed, and streptavidin-HRP was added. The resulting signal, proportional to the amount of stress-induced protein bound, was visualized using a ChemiDoc system per manufacturer guidelines, and analyzed.

Ubiquitination screen assay

The evaluation of ubiquitination of 49 target proteins were determined using an R&D Systems Proteome Profiler Human Ubiquitin Array Kit (R&D ARY027). Relative expression levels of ubiquitination of human proteins in samples were determined as per the manufacturer's protocol. Briefly, the human ubiquitin array nitrocellulose membranes spotted with 49 different antibodies to human ubiquitin target proteins were incubated with prepared cell lysates overnight at 4C on a rocking platform shaker. After thorough washing to remove unbound proteins, the membranes were incubated with biotinylated pan-anti-ubiquitin detection antibody cocktail for 1 hour. Then the membrane was washed thoroughly, followed by addition of streptavidin-HRP. The signal produced at each capture spot corresponding to the relative amount of ubiquitinated protein bound was imaged using a ChemiDoc system per manufacturer guidelines, and analyzed.

Cleavage Under Targets and Release Using Nuclease (CUT&RUN)

Cells were seeded at 200,000 cells per well in 6 well plates in Glucose media (+NEAA) or Glucose media (-NEAA), and cultured for 4 days. Upon reaching confluence, four pellets of 50,000 cells per cell line, per media condition, were harvested to a 1.5 mL Eppendorf tube. Pellets were washed with PBS and pelleted at 300xg for 5 minutes.

CUT&RUN libraries were prepared using the SimpleChIP ChIP-seq DNA Library Prep Kit for Illumina (Cell Signaling 56795S) along with SimpleChIP ChIP-seq Multiplex Oligos for Illumina (Cell Signaling 47538S), according to manufacturer's guidelines, with adaptations for CUT&RUN as specified in the CUT&RUN Assay Kit (Cell Signaling 86652S). The following antibodies were utilized for CUT&RUN libraries: IgG isotype (DA1E) (Cell Signaling 66362, negative control, 1:20 dilution), H3K4me3 (Cell Signaling 9751S, 1:50) and ATF-4 (D4B8) Rabbit mAb (Cell Signaling 11815, 1:200), as per protocol. Subsequently, all libraries were quantified using the Qubit 4.0 fluorometer and assessed for quality using the Agilent High Sensitivity DNA Tapestation (Agilent 5067-5584) before multiplexing and sequencing on a Nextseq 2000 P2 Kit (100 cycles), targeting a minimum of 5 million paired-end reads per sample with 50 bp read length.

Samples were analyzed using the established NF-Core CUT&RUN best practices bioinformatics pipeline ²².

Quantification and statistical analysis

Each experimental condition was repeated at least three times. Fibroblast lines from controls (n=3) and BOS patient (n=3) were seeded in technical replicates of at least three in each experimental condition. All experiments were performed in biological triplicate unless otherwise indicated.

Graphs present the mean +/-SD, and p values were generated by two-tailed Student's t test. Asterisks indicate the significance of the p value: *p < 0.05; **p < 0.01; ***p < 0.001. Data were analyzed and visualized using GraphPad Prism and *ggplot2*.

Results

In our investigation of the metabolic dysregulation present in Bohring-Opitz syndrome (BOS), we employed an integrated approach using patient-derived fibroblasts as a model system. **Figure 4-1A** illustrates collection of skin punch biopsies from BOS patients to culturing of fibroblasts that serve as the foundation for our investigations. This schematic representation illustrates the array of analytical assays utilized to unravel the complex metabolic and molecular landscape of BOS.

We utilized the scratch wound assay to evaluate the migration capabilities of BOS cells in various nutrient conditions. This assay provided valuable insights into cell motility and wound healing, and allows for high-throughput manipulation of drug and compound testing, such as repletion or depletion of nutrients. Subsequent analysis with the Seahorse assay examined critical alterations in mitochondrial bioenergetics, including significant shifts in the OCR and glycolytic activity, underscoring the metabolic reprogramming in BOS cells.

To delve deeper into the molecular aspects, western blot analyses were conducted to quantify the regulation of the mTOR pathway, expression levels of mitochondrial proteins, and the activation of stress pathways. Our study extended to both untargeted metabolomics and targeted tracing to probe the metabolic pathway intermediates, revealing a disrupted metabolic flux indicative of the metabolic inflexibility in BOS cells. Multi-omics assays, including CUT&RUN for histone modifications (H3K4me3) and transcription factor (ATF4) localization, as well as RNA-seq analysis under different amino acid conditions, provided a comprehensive view into the epigenetic and transcriptional shifts occurring in BOS fibroblasts.

Mitochondrial pyruvate carrier levels are significantly reduced in BOS

Western blot analyses were conducted to examine the expression levels of the mitochondrial pyruvate carriers MPC1 and MPC2, which are crucial for mitochondrial pyruvate uptake and, by extension, mitochondrial function. These assays were performed on fibroblast

cells derived from both control individuals (n=4) and patients diagnosed with BOS (n=3). Our findings revealed a significant deficiency in the expression levels of both MPC1 and MPC2 in the BOS fibroblast samples compared to controls (**Figure 4-1B**), indicating a direct impact of *ASXL1* mutations on mitochondrial pyruvate transport. This difference was not identified at the RNA transcript level.

Quantification of the protein levels, normalized against actin as a loading control, further confirmed the reduced expression of MPC1 and MPC2 in BOS cells (**Figure 4-1C**). Statistical analysis showed a significant decrease in MPC1 ($p=8.32E-04$) and MPC2 ($p=1.70E-03$) expression levels in BOS fibroblasts, suggesting that *ASXL1* mutations may lead to impaired mitochondrial function by affecting the expression of key mitochondrial pyruvate carriers.

We extended our findings to other BOS patient-derived sample types. Our findings revealed a stark contrast in the impact of *ASXL1* mutations on MPC1 and MPC2 protein levels between cell types, with a significant decrease observed in BOS fibroblasts but not in BOS-induced pluripotent stem cells (iPSCs, **Figure 4-2A**) or neural progenitor cells (NPCs, **Figure 4-2B**). This differential expression suggests that the cellular state—specifically, the transition from pluripotency to a differentiated state—plays a critical role in mediating the phenotype. While BOS iPSCs and NPCs maintained normal levels of these MPCs, the differentiated fibroblasts exhibited diminished levels, implying a possible loss of mitochondrial function associated with differentiation in the presence of *ASXL1* mutations.

Metabolomic analyses unveil significant enrichment in glycolysis in BOS cells

Untargeted metabolomic analyses highlighted significant disruptions in mitochondrial function and cellular metabolism in BOS fibroblasts. This was evident from the dysregulation of critical metabolites, including dihydroorotate, FGAR, glucose, glyceraldehyde-3-phosphate, UMP, AMP, CMP, and pyruvate (Student's t test, $p=1.92E-02$), emphasizing the profound metabolic disturbance in BOS cells (**Figure 4-3A**). Mummichog 2.0 analysis further supported

these findings, revealing significant pathway enrichments. Notably, we identified enrichment of glycolysis/gluconeogenesis (enrichment factor 2.05, $-\log_{10}$ p-value 2.16), and pyruvate metabolism (enrichment factor 2.34, $-\log_{10}$ p-value 1.96), suggesting metabolic reprogramming or compensation in BOS patient-derived fibroblasts (**Figure 4-3B**).

A heatmap comparison between control and BOS samples showed a significant increase in glycolytic metabolites in BOS samples, indicating enhanced glycolytic flux (**Figure 4-3C**). Pyruvate, a key glycolysis metabolite, was significantly elevated in BOS fibroblasts (Student's t test, $p=1.92E-02$) (**Figure 4-3D**).

BOS patient-derived fibroblasts utilise increased glycolysis and decreased oxidative phosphorylation

Using the Seahorse assay, we identified a distinctive metabolic profile in BOS cells, characterized by a marked decrease in the basal OCR and increased extracellular acidification rate (ECAR). The former is a proxy measure for mitochondrial oxidative phosphorylation and the latter is a proxy measure for glycolytic processes. Quantitative analysis of ATP production sources revealed significantly decreased percentage of ATP generated by oxidative phosphorylation in BOS fibroblasts (mean $35.37\% \pm 10.95\%$) compared to controls (mean $52.77\% \pm 16.93\%$) ($p=1.16E-03$), indicating a diminished usage of mitochondrial ATP generation and reliance on alternative pathways such as glycolysis (**Figure 4-3E**).

The normalized OCR data from the Seahorse assay (**Figure 4-3F**) highlights the significantly reduced mean basal OCR in BOS fibroblasts (2.45 pmol/min) compared to controls (2.93 pmol/min) ($p=4.23E-02$) (**Figure 4-3G**). Surprisingly, BOS cells had comparable levels of maximal OCR to controls, suggesting that BOS cells have an adaptive decrease of OCR and adaptive increase of glycolysis, with potential for mitochondrial engagement when glycolysis is less favorable.

The normalized ECAR data highlighted the upregulated glycolytic metabolism in BOS cells compared to controls (**Figure 4-3H**). BOS cells had significantly increased mean basal ECAR (3.00 mpH/min) compared to controls (1.90 mpH/min) ($p=1.01E-04$) (**Figure 4-3I**).

Depletion of nutrients exacerbates metabolic inflexibility in BOS cells

To further investigate the adaptive responses of BOS cells to metabolic stress, we assessed the effects of nutrient alterations, with a specific focus on glucose, NEAA and pyruvate depletion. Availability of glucose regulates the glycolytic pathways and the availability of NEAAs and pyruvate regulates the TCA cycle.

We cultured control and BOS fibroblast cells under conditions designed to challenge their metabolic flexibility, namely substituting glucose with galactose to push cells towards oxidative phosphorylation, and in the presence or absence of NEAA and pyruvate to test the effects of nutrient deprivation. Unlike glucose, which can readily enter glycolysis to fuel cellular energy production, galactose requires conversion to glucose-6-phosphate before entering glycolysis, thus placing a greater demand on mitochondrial oxidative phosphorylation for ATP generation^{23,24}.

As expected in galactose medium, all cells exhibited a marked increase in the percentage of ATP generated through oxidative phosphorylation (as opposed to glycolysis) compared to glucose conditions, indicating an enhanced reliance on mitochondrial respiration (**Figure 4-4B**). Control cells showed a slight, but statistically significant, decrease in oxidative phosphorylation efficiency upon NEAA and pyruvate depletion ($p=2.13E-02$), highlighting the importance of these nutrients in supporting mitochondrial function.

Basal oxygen consumption rates (OCR) remained consistent across all conditions, averaging 5-6 pmol/min, significantly higher than rates observed in glucose conditions (**Figure 4-4C**). However, maximal OCR revealed a differential response; depletion of pyruvate ($p=6.21E-03$), depletion of NEAA ($p=4.25E-03$), and depletion of both pyruvate and NEAA

($p=1.45E-05$) drove significantly reduced capacities for upregulating respiratory activity in BOS cells, indicating compromised mitochondrial function and dependency on exogenous nutrients (**Figure 4-4D**). In comparison, control cells did not show any statistically significant difference in maximal OCR when depleted of NEAA, pyruvate, or both NEAA and pyruvate.

In contrast to glucose conditions, all galactose conditions showed significantly lower basal ECAR, with averages ranging from 0.4 to 0.6 mpH/min, as expected. This reduction confirms a decreased reliance on glycolysis for energy production in the absence of glucose and in the presence of galactose (**Figure 4-4E**). Interestingly, control cells exhibited a slight increase in ECAR upon NEAA and pyruvate depletion ($p=5.99E-03$), potentially indicating a normal metabolic adaptation to the depletion of these nutrients. BOS cells did not exhibit this compensatory response.

Mitochondrial mass per cell area, assessed using mitogreen tracker (MTG), significantly increased in conditions of pyruvate depletion for both control and BOS cells ($p=3.31E-18$ to $3.07E-15$), suggesting an acute compensatory mitochondrial biogenesis in response to metabolic stress (**Figure 4-4F**). Depletion of NEAA did not significantly affect control cells while BOS cells experienced a significant decrease in mitochondrial mass per cell area ($p=2.40E-03$).

BOS cells rely on increased glycolytic flux and de novo amino acid biosynthesis

Untargeted metabolomics revealed elevated pyruvate levels in BOS cells compared to controls ($p=1.92E-02$), which normalized to control-like levels upon NEAA depletion, indicating a complex interplay between pyruvate and NEAA levels in BOS pathology (**Figure 4-4G**).

Western blot analysis of mitochondrial pyruvate carrier proteins 1 and 2 (MPC1 and MPC2) showed significant reductions in BOS cells compared to controls, regardless of NEAA presence (**Figure 4-4H**). This suggests a mechanistic link between *ASXL1* mutations and a critical bottleneck in mitochondrial pyruvate transport.

In exploring the metabolic alterations induced by ASXL1 mutations in BOS, this study examined the glycolytic and TCA cycle dynamics, and *de novo* amino acid synthesis pathways through glucose tracing experiments (**Figure 4-5A**). The findings highlight an enhanced glycolytic rate alongside upregulated *de novo* amino acid synthesis, even under amino acid-rich conditions.

Despite comparable glucose uptake across both cell types (**Figure 4-5B**), BOS cells demonstrated a pronounced increased flux of glucose through the glycolytic pathway, evidenced by a significant accumulation of fructose-1,6-bisphosphate in both NEAA replete and deplete conditions (**Figure 4-5C**, $p=1.67E-02$). This enhanced glycolytic flux suggests an increased reliance on glycolysis. Moreover, the isotopic composition analysis revealed a higher percentage of M6-labeled pyruvate in BOS cells (**Figure 4-5D**).

Isotope tracing also revealed significant dysregulations in TCA cycle dynamics (**Figure 4-5E**, $p<0.0001$) and an increased biosynthesis of glutamate (**Figure 4-5F**, $p<0.0001$) and aspartate (**Figure 4-5G**, $p<0.0001$) from TCA cycle intermediates α -ketoglutarate and oxaloacetate respectively, suggestive of an increased demand for amino acid biosynthesis in BOS cells. Intriguingly, while BOS cells with NEAA exhibited significantly increased labelled carbon flux for glutamate and aspartate, the total peak area was not significantly different from control cells with NEAA ($p=0.36$ for glutamate and $p=0.78$ for aspartate). Intriguingly, while BOS cells without NEAA exhibited control-like levels of isotope composition, the total peak area was significantly decreased compared to controls ($p=1.66E-02$ for glutamate and $p=4.40E-03$ for aspartate) and to BOS cells with NEAA ($p=0.13$ for glutamate and $p<0.0001$ for aspartate). This suggests that while BOS cells are able to compensate through increased carbon flux and *de novo* amino acid synthesis in nutrient-rich conditions to maintain a control-like level of certain amino acids, depletion of exogenous nutrients leads to an inability to compensate in BOS cells.

NEAA depletion exacerbates mTOR signaling defects in BOS cells, which can be rescue by L-asparagine repletion

To explore the impact of truncating *ASXL1* mutations on mTOR signaling, a key nutrient sensing pathway, and cellular stress responses in BOS cells, we conducted western blot analyses on whole cell lysates. These lysates were collected after 16 hours, 48 hours, and 72 hours of NEAA depletion to assess changes in key protein levels, providing insight into mTORC1 activity and the cellular stress response.

In Glucose media (+NEAA), we identified no significant differences in levels of total S6 or total S6K proteins between BOS and control cells (0.94-fold change) (**Figure 4-6A, Figure 4-6B, Figure 4-6D**). However, we noted an elevated baseline ASNS level in BOS at baseline compared to controls (1.49-fold increase), and in active phosphorylated S6K; pS6K p70 (0.60-fold decrease) and pS6K p85 (0.54-fold decrease). This suggests baseline elevated stress response and impaired activation of mTOR signaling (**Figure 4-6A**).

Upon acute NEAA depletion (16 hours), BOS cells were marked by a pronounced increase in ASNS expression compared to at baseline (1.22-fold increase) and to control cells (t-test, $p=0.0168$), indicating an acute stress reaction likely mediated through disruptions in mTORC1 signaling and asparagine synthesis or utilization. Upon NEAA depletion, both BOS and control cells exhibited a slight increase in ATF4 (1.26-fold increase and 1.22-fold increase respectively) and pS6K p70 (1.26-fold increase and 1.33-fold increase respectively), and p85 (1.18-fold increase and 1.15-fold increase respectively), suggesting a standard early mTORC1 response to nutrient scarcity.

By 48 hours after NEAA depletion (**Figure 4-6B**), differences in ASNS levels between BOS and control cells normalized under both +NEAA and -NEAA conditions. However, significant reductions in both pS6K p70 (t-test, $p=4.41E-02$) and in pS6K p85 (t-test, $p=5.40E-03$) levels were identified compared to controls (**Figure 4-6E, Figure 4-6F**).

This reduction in S6K phosphorylation, a hallmark of dysregulated mTORC1 signaling, was reversed upon L-asparagine repletion, underscoring a critical metabolic bottleneck and potential therapeutic target in BOS cells. Intriguingly, mTORC1 activity in BOS responded positively to L-asparagine repletion (**Figure 4-6C**). Specifically, with repletion of L-asparagine alone, pS6K levels returned to comparable levels as in Glucose media (+NEAA). We identified partial rescue of the mTORC1 phenotype in BOS cells with an increase of pS6K p70 (1.49-fold increase) and pS6K p85 (1.18-fold increase) (**Figure 4-6E, Figure 4-6F**). Control cells did not exhibit this response, highlighting a unique metabolic response in BOS fibroblasts. No effect was observed on ASNS expression.

At 72 hours after NEAA depletion, protein levels were similar to those at 48 hours, with no significant differences in ASNS or ATF4 levels, and significantly reduced protein levels of pS6K p70 (0.79-fold decrease, t-test, $p=3.29E-02$) and pS6K p85 (0.73-fold decrease) (**Figure 4-6D, Figure 4-6E, Figure 4-6F**). This finding highlighted the specific dependency of BOS cells on asparagine for normal mTOR pathway activity which aligns with our early observations of increased ASNS levels upon acute depletion of NEAA. This may point towards an adaptive mechanism in BOS cells to synthesize asparagine internally when external supply is restricted.

Asparagine has been identified as a key modulator of mTOR signaling. When cellular levels of asparagine are depleted, it leads to the activation of the integrated stress response (ISR) pathway, a cellular stress response mechanism. This activation, in turn, induces the expression of *ATF4*, which regulates amino acid metabolism, stress responses, and mTORC1 activity in response to asparagine availability^{18,20}. At the transcriptomic level, BOS cells exhibited increased *ATF4* levels compared to controls (**Figure 4-7A**). From our previous multi-omics study³, we observed H3K4me3 enrichment at the *ATF4* promoter and gene body in BOS fibroblast cell lines, indicating derepression and corroborating the transcriptomic findings (**Figure 4-7B**).

BOS cells are dependent on NEAA for cell migration, and partial rescue can be achieved with repletion of L-asparagine

The scratch wound assay is a widely used *in vitro* technique to study the process of wound healing in vitro. We assessed the growth and migration capabilities of BOS fibroblasts under various nutritional conditions. This technique sheds light on cellular motility and the inherent mechanisms of wound healing ^{25,26}, and facilitates the high-throughput evaluation of therapeutic agents by modulating nutrient availability ^{27,28}.

Given the interplay between *ASXL1* mutations and mTOR pathway dysregulation¹⁷, which regulates epithelial cell proliferation, migration, cutaneous wound healing ²⁹, and negatively regulates autophagy³⁰, we investigated the impact of nutrient deprivation on cell migration in BOS. We hypothesized that the increased sensitivity of the mTOR pathway in BOS cells to nutrient availability might influence cell growth and migration.

In Glucose media (+NEAA), we identified no significant difference in wound healing between BOS patient fibroblasts and control fibroblasts (**Figure 4-8A**). However, the depletion of NEAAs significantly reduced the growth rate of BOS cells post-scratch wound, highlighting a stark contrast to control fibroblasts, which remained unaffected. Remarkably, the repletion of L-asparagine, mimicking the NEAA cocktail composition, substantially mitigated this growth deficit in BOS cells, improving growth rates by 30-40% at 24 hours and 40-50% at 48 hours post-supplementation, pinpointing a critical metabolic dependency on asparagine (**Figure 4-8B**).

Interestingly, the concurrent removal of pyruvate and supplementation with L-asparagine normalized the growth rates of BOS cells to levels observed under +NEAA conditions or in control fibroblasts (**Figure 4-8C**), suggesting a nuanced metabolic requirement in BOS cells that extends beyond simple nutrient availability and warrants further study.

Discussion

Our study examines the intricate link between truncating *ASXL1* mutations and mitochondrial dysfunction in Bohring-Opitz syndrome (BOS, OMIM#605309), revealing profound impacts on cellular metabolism and signaling pathways (**Figure 4-9**). The metabolic characterization of BOS patient-derived fibroblasts demonstrates a marked reduction in the expression levels of MPC1 and MPC2, which are crucial for mitochondrial pyruvate import (**Figure 4-1B, Figure 4-1C**). The MPC reduction is substantiated by a significant decrease in both basal and maximal OCR (**Figure 4-3F, Figure 4-3G**), indicating compromised mitochondrial function.

In humans, MPC is a hetero-oligomeric complex formed by the MPC1 and MPC2 paralogs that are both necessary to stabilize each other and form a functional MPC. MPC deficiency (OMIM#614741) due to pathogenic MPC1 variants is a rare autosomal recessive disease involving developmental delay, microcephaly, growth failure, and increased serum lactate and pyruvate³¹. Similarly, BOS patients present with developmental delay and growth failure, as well as increased metabolite levels of pyruvate in patient-derived fibroblasts (**Figure 4-3A, Figure 4-3E**). Intriguingly, MPC loss-of-function studies in *Drosophila* intestinal epithelial cells were identified to increase stem cell proliferation, while overexpression suppressed stem cell proliferation, suggesting that MPC loss supports stem cell expansion and prevents differentiation^{32,33}. Similar studies in human cells confirmed that MPC loss can increase stem-like cell proliferation and drive cells towards a glycolytic state³⁴.

The efficiency of mitochondria in utilizing pyruvate through the MPC is essential for linking glycolysis, the primary pathway of glucose metabolism in the cytosol, with mitochondrial energy production processes to generate ATP³⁵. Dysfunctional MPC, due to either reduced expression or altered activity of MPC1 and MPC2, can severely impair cellular metabolism. This dysfunction restricts the flow of pyruvate into the mitochondria, leading to a metabolic bottleneck that forces cells to rely more heavily on glycolysis for ATP production, even in the presence of

oxygen—also known as aerobic glycolysis or the Warburg effect ³⁵. The Warburg effect is typically observed in cancer cells, other rapidly proliferating cells, and under certain pathological conditions, where cells rely on glycolysis rather than oxidative phosphorylation even under oxygen-rich conditions.

Our metabolomic analyses revealed a clear metabolic reprogramming in BOS cells, characterized by the increased flux through glycolysis and altered pyruvate metabolism (**Figure 4-3A, Figure 4-3B, Figure 4-3D**). Such metabolic shifts are reminiscent of the Warburg effect or of an early pluripotent state, where aerobic glycolysis is required to stimulate acetylation of histones in chromatin and maintain cells in a stem-like state ^{36–39}. In BOS cells, the increased reliance on glycolysis, despite the presence of oxygen, could suggest a shift towards a similar metabolic profile, aiming to support rapid cell proliferation or adapt to metabolic stress.

These findings were supported by observations of increased ECAR in BOS cells, suggesting a higher glycolytic activity a direct consequence of a preference to generate increased biomass or to compensate for mitochondrial deficiencies (**Figure 4-3H, Figure 4-3I**). The latter may suggest impaired metabolic flexibility, where cells cannot efficiently switch between glycolysis and oxidative phosphorylation, leading to an overreliance on glycolysis. This metabolic reprogramming or altered set point could be an adaptive mechanism to mitigate mitochondrial-derived reactive oxygen species (ROS) production, as glycolysis is less prone to generating ROS compared to oxidative phosphorylation. Additionally, the accumulation of glycolytic intermediates might fuel anabolic pathways, supporting biomass accumulation essential for cell growth and division.

Sercel et al. (2024) suggest that pathophysiology of mitochondrial defects actually lies in maladaptive stress responses triggered by oxidative phosphorylation defects, leading to a state of hypermetabolism rather than mere ATP deficiency ⁴⁰. This hypermetabolic state imposes energetic constraints that are pivotal in causing the pathophysiology observed in mitochondrial diseases. This perspective aligns with our observations in BOS cells, where increased

glycolysis, despite reduced oxidative phosphorylation, could be part of a broader compensatory mechanism to meet the energetic demands imposed by cellular stress responses. Recent studies have also shown that *ASXL1* mutations regulate glucose metabolism ⁴¹. Together, our findings highlight a profound metabolic reprogramming characterized by an enhanced glycolytic rate and rerouted carbon flux into amino acid biosynthesis in BOS cells (**Figure 4-9**).

The hypermetabolic state induced by mitochondrial dysfunction, as described by Sercel et al., may be linked to accelerated ageing and physiological decline. This reorientation of energy expenditure could divert resources away from longevity-promoting processes, reducing fitness and lifespan ⁴⁰. Interestingly, DNA methylation studies in BOS patient blood samples show acceleration in DNA methylation age ⁴².

Recent findings suggest a broader network involving *ASXL1*'s binding partner, *BAP1*, in the regulation of cellular metabolism. Germline *BAP1* mutations have been linked to the induction of a Warburg effect, characterized by increased aerobic glycolysis and lactate secretion, sacrificing mitochondrial respiration and ATP production to facilitate uptake and incorporation of nutrients into the biomass even in the presence of oxygen ^{43,44}. This metabolic phenotype was observed in primary fibroblasts from carriers of heterozygous germline *BAP1* mutations (*BAP1*^{+/-}), highlighting the role of *BAP1* in modulating cellular metabolism. Further in vitro studies using siRNA-mediated reduction of *BAP1* confirmed these metabolic alterations ^{44,45}. Notably, these findings challenge the conventional notion that the Warburg effect solely arises as an adaptive response to carcinogenesis, suggesting its potential role as an early facilitator of malignancy ⁴⁴.

The response of BOS cells to nutrient stress, as observed under different growth conditions, highlights the maladaptive stress responses in BOS cells. The pronounced reduction in maximal OCR under nutrient-deprived conditions (**Figure 4-4D**) underscores the limited capacity of these cells to upregulate respiratory activity in response to metabolic stress. This is supported by our findings that BOS cells exhibit increased levels of glycolytic intermediates and

increased biosynthesis of key amino acids like glutamate and aspartate, under nutrient-rich conditions, but are unable to maintain this increased biosynthesis of key amino acids under nutrient-depleted conditions (**Figure 4-5F**, **Figure 4-5G**) which may drive the growth impairment phenotype observed in our functional assays.

The scratch wound assay revealed that BOS cells are highly dependent on external nutrient availability. Specifically, the depletion of NEAAs drastically impairs cell migration, an effect partially rescued by the addition of L-asparagine (**Figure 4-8B**). These findings indicate not only a dependency on specific amino acids but also suggests potential therapeutic targets to mitigate such growth deficits.

Our study also highlights significant disruptions in the mTOR signaling pathway, a central regulator of cell growth and metabolism. The reduction in pS6K upon NEAA depletion (**Figure 4-6B**, **Figure 4-6D**) points to mTORC1 dysregulation. Notably, the repletion of L-asparagine restores mTORC1 activity, suggesting a specific amino acid sensitivity and a potential therapeutic avenue to normalize mTOR signaling in BOS cells (**Figure 4-6C**).

These metabolic and signaling alterations are likely driven by transcriptional and epigenetic changes, as evidenced by our CUT&RUN and RNA-seq analyses. Increased H3K4me3 deposition at the *ATF4* promoter (**Figure 4-7B**) and elevated *ATF4* expression under nutrient stress (**Figure 4-7A**) suggest that stress response pathways are also engaged in an attempt to adapt to mitochondrial dysfunction and nutrient scarcity.

Our findings align with previous studies indicating that disruptions in mitochondrial pyruvate transport can lead to compensatory metabolic reconfigurations. For instance, isotopomeric flux analysis has shown that during MPC suppression, alternative carbon sources can replenish TCA cycle intermediates, a mechanism potentially sustained through glutamine oxidation^{46,47}, and which may be at play in the BOS patient population. This adaptation aligns with the upregulation of pathways involving glutamine transporters and asparagine synthetase

during metabolic stress, which we also observe in the context of ATF4-mediated transcriptional responses.

Our study is the first that we know of to identify metabolic and mitochondrial dysfunctions in BOS. However, several limitations are inherent to our study of rare diseases. One significant limitation is the small sample size, which can impact the generalizability of our findings and statistical power of the analyses. Additionally, the use of patient-derived fibroblasts, while informative, does not encompass the full spectrum of tissue types that may be affected by *ASXL1* mutations. Although fibroblasts are a reliable source for metabolic studies, they do not necessarily reflect the metabolic states of more ubiquitously affected tissues *in vivo*. Our study did extend to patient-derived iPSCs (**Figure 4-2A**) and NPCs (**Figure 4-2B**) which did not show the MPC defect observed in fibroblasts. This discrepancy could be attributed to cellular reprogramming, which may alter mitochondrial characteristics because mitochondrial function is known to be repressed in pluripotent cells compared to differentiated cells ⁴⁸, or could lend insight into BOS pathophysiology whereby BOS cells are unable to increase MPC levels appropriately as cells differentiate.

Future expansions of this investigation to other rare genetic neurodevelopmental disorders could validate whether the metabolic and mitochondrial disruptions observed in BOS are common across other similar disorders, especially those not typically associated with direct metabolic pathway mutations. Moreover, ongoing studies in patient-derived peripheral blood mononuclear cells are anticipated to provide additional insights into the systemic metabolic alterations in BOS, potentially offering a more accessible biomarker or therapeutic target. Further metabolic tracing studies, such as glutamine tracing and lactate tracing, will enhance our understanding of the role these metabolic intermediates play in the pathogenesis of BOS, providing a clearer picture of how metabolic fluxes support the survival and function of affected cells under metabolic stress. Such studies could pave the way for novel therapeutic strategies

aimed at modulating metabolic pathways to mitigate the effects of genetic mutations in BOS and potentially other similar disorders.

Overall, our findings indicate a marked reduction in MPC1 and MPC2 protein levels in BOS cells, suggesting that *ASXL1* mutations may interfere with pyruvate transport into mitochondria, thereby compromising cellular energetics. This mitochondrial pyruvate bottleneck could contribute to the metabolic inflexibility observed in BOS. Our comprehensive analysis not only underscores the profound metabolic and signaling disturbances in BOS but also provides a foundation for exploring metabolic reprogramming as a potential therapeutic strategy. By delineating the cellular adaptations to mitochondrial dysfunction and nutrient stress, we highlight the potential for targeted interventions that could alleviate the metabolic deficiencies observed in BOS.

Conclusion

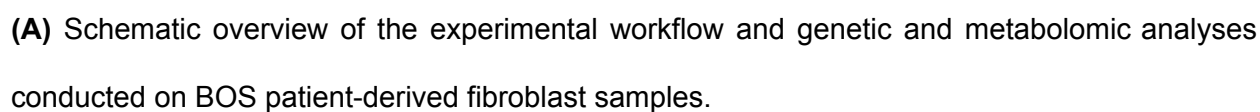
Our comprehensive investigation into the metabolic and mitochondrial implications of *ASXL1* mutations in BOS underscores a complex interplay between genetic aberrations and cellular metabolic responses. We have demonstrated significant disruptions in mitochondrial function, particularly a notable reduction in the levels of mitochondrial pyruvate carriers MPC1 and MPC2 in patient-derived fibroblasts. This finding aligns with decreased mitochondrial respiration capacities as evidenced by lower basal and maximal oxygen consumption rates. Furthermore, our metabolic profiling revealed a profound reprogramming towards increased glycolysis and amino acid biosynthesis, suggesting an adaptive response to mitochondrial inefficiency.

The cellular adaptations observed in BOS cells—such as increased glycolytic flux and altered amino acid metabolism—highlight a fundamental shift towards alternative metabolic pathways, likely as a compensatory mechanism for impaired mitochondrial function. These adaptations are consistent with the hypermetabolic state described in other mitochondrial

pathologies, which may contribute not only to the disease phenotype but also to broader implications such as accelerated cellular aging and altered physiological resilience.

In conclusion, our study not only provides crucial insights into the metabolic consequences of *ASXL1* mutations in BOS but also sets a precedent for the metabolic examination of other genetic neurodevelopmental disorders. Our findings highlight the importance of integrating metabolic perspectives into the study of rare diseases to ultimately improve patient outcomes and quality of life.

Figure 4-1: Overview of methods and identification of intrinsic metabolic defects in BOS



(C) Normalized boxplots quantifying protein expression of MPC1 ($p=8.32E-04$) and MPC2 ($p=1.70E-03$), indicate statistically significant reduction in BOS compared to control samples.

181

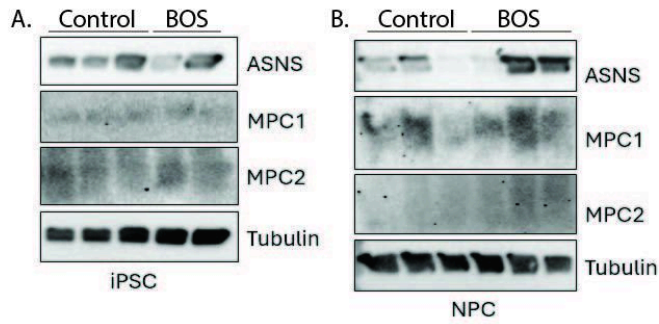


Figure 4-2: Conserved expression of MPC1 and MPC2 in pluripotent and neural progenitor cells from BOS patients

(A) Western blots showing the expression levels of MPC1 and MPC2 in control iPSC samples (n=3) and BOS iPSC samples (n=2).

(B) Western blots showing the expression levels of MPC1 and MPC2 in control NPC samples (n=3) and BOS NPC samples (n=3).

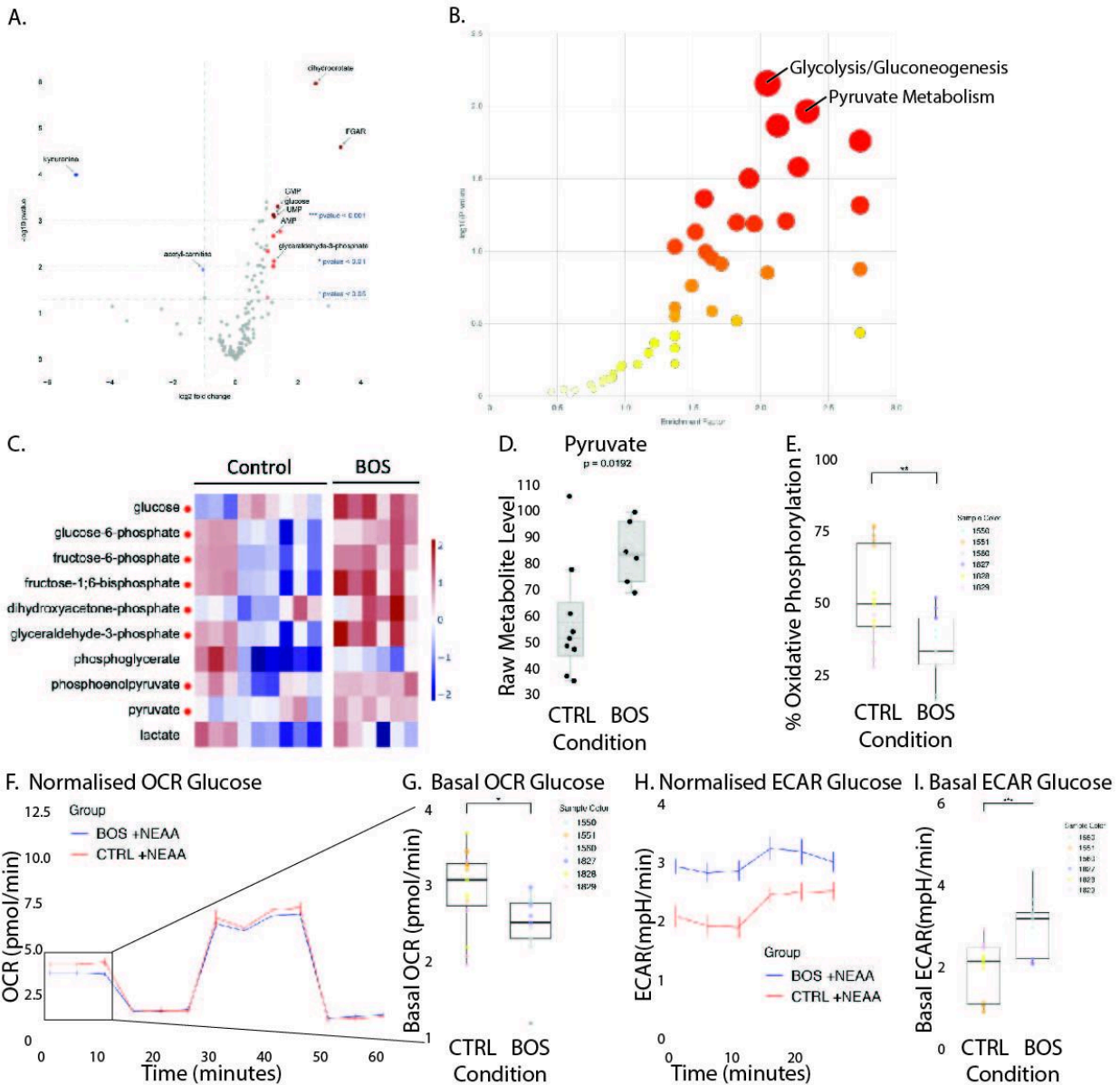


Figure 4-3: BOS patient-derived fibroblasts utilize increased glycolysis and decreased oxidative phosphorylation

(A) Volcano plot of untargeted metabolomics data, highlighting significant differential metabolites between BOS and control fibroblasts. P-value thresholds are marked with horizontal dotted lines

(B) Mummichog 2.0 pathway analysis show the primary affected metabolic pathways in BOS fibroblasts, with glycolysis/gluconeogenesis and pyruvate metabolism being most significantly disrupted.

(C) Heatmap of glycolytic pathway intermediates comparing metabolite levels in BOS fibroblasts (n=2 biological replicates, in technical triplicate) to control fibroblasts (n=3 biological replicates, in technical triplicate).

(D) Box plot depicting elevated pyruvate metabolite levels in BOS fibroblasts ($p=1.92E-02$) compared to controls, indicating disrupted pyruvate metabolism.

(E) Percentage of oxidative phosphorylation (% OxPhos) capacity is significantly reduced in BOS fibroblasts.

(F) Normalized OCR (pmol/min) data from Seahorse assay of BOS and control patient-derived fibroblasts.

(G) Basal OCR (pmol/min) comparison, indicating significantly lower rates in BOS fibroblasts.

(H) Normalized extracellular acidification rate (ECAR, pmol/min) data, a measure of glycolytic activity, is increased in BOS cells.

(I) Basal ECAR (mpH/min) is significantly increased in BOS cells, reflecting a compensatory increase in glycolytic flux.

Statistical significance is indicated by asterisks: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, and **** $p < 0.0001$.

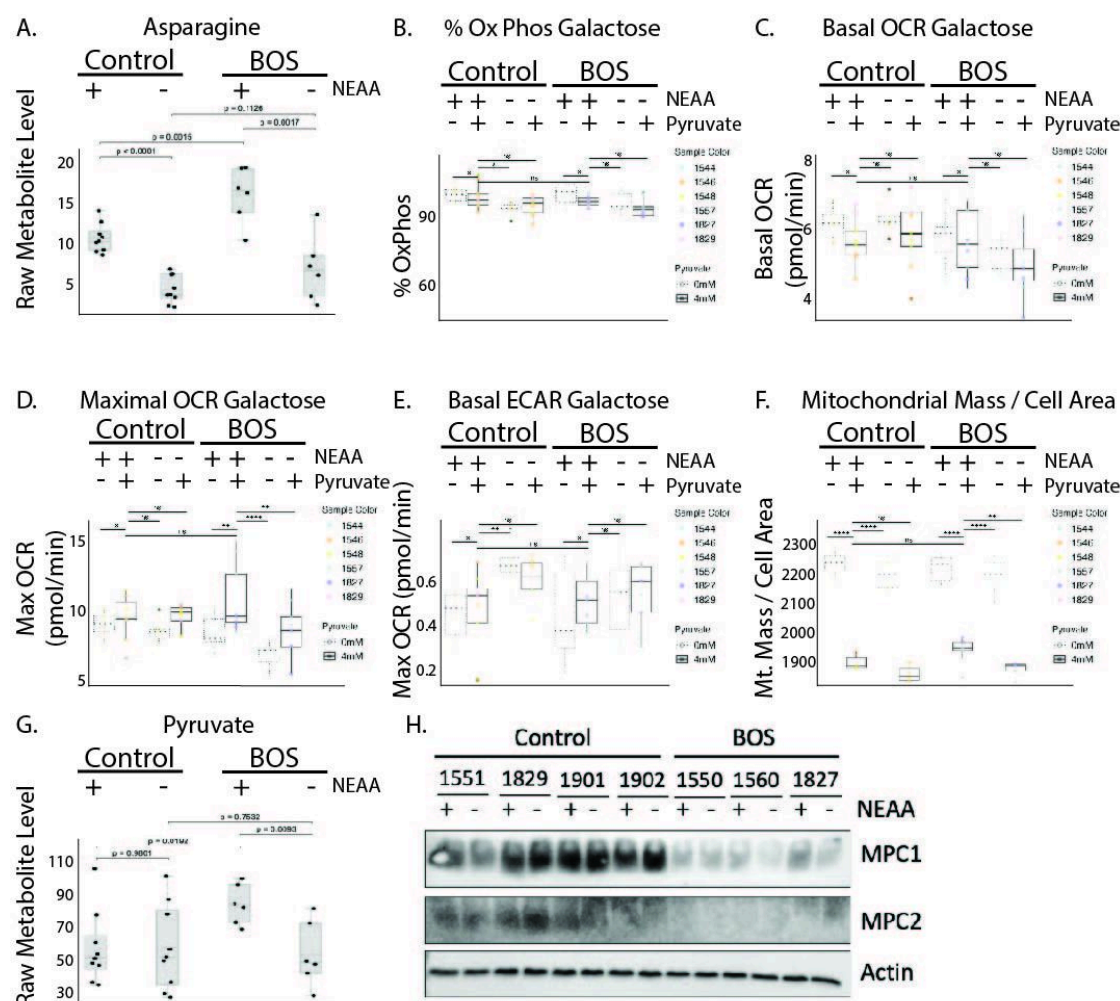


Figure 4-4: Depletion of nutrients exacerbate metabolic inflexibility in BOS cells

(A) Untargeted metabolomics showing asparagine levels in control and BOS cells under different nutrient conditions.

(B) Percentage of ATP production via oxidative phosphorylation in control and BOS fibroblasts under various nutrient conditions, showing a general increase in oxidative phosphorylation in galactose and a slight decrease in control cells upon NEAA and pyruvate depletion ($p=2.13E-2$).

(C) Basal oxygen consumption rate (OCR) in cells cultured in galactose. No significant differences were identified across nutrient conditions, but a notable increase was identified compared to glucose conditions.

(D) Maximal OCR capacity showcasing a reduced ability in BOS cells to increase respiratory activity without pyruvate ($p=6.21E-03$), without NEAA ($p=4.25E-03$), and without both ($p=1.45E-05$).

(E) Basal extracellular acidification rate (ECAR) in galactose conditions, highlighting a significant reduction in glycolytic reliance compared to glucose conditions, with a slight increase in control cells under nutrient depletion ($p=5.99E-03$).

(F) Mitochondrial mass per cell area increased significantly in all conditions lacking pyruvate, suggesting a compensatory mitochondrial biogenesis in response to nutrient stress (p values = $3.31E-18$ to $3.07E-15$).

(G) Untargeted metabolomics showing pyruvate levels in control and BOS cells under different NEAA conditions, revealing significant alterations in BOS cells that align with nutrient availability.

(H) Western blot analysis of MPC1 and MPC2 levels in control and BOS cells with or without NEAA, demonstrating significantly reduced levels in BOS cells across nutrient conditions.

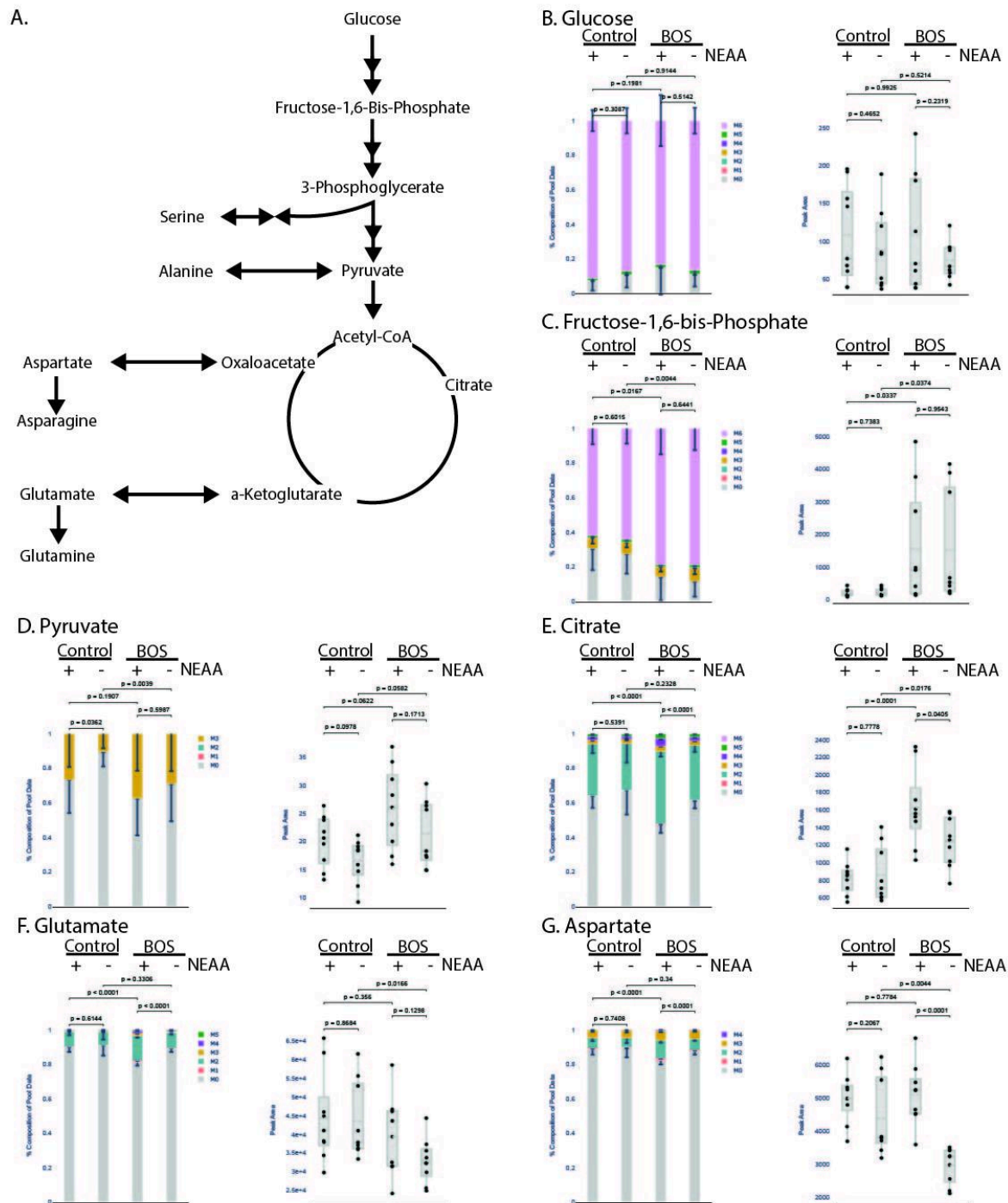


Figure 4-5: BOS cells rely on increased glycolytic flux and de novo amino acid biosynthesis

(A) Schematic representation of the glycolysis and TCA cycle pathways, highlighting the intermediates for *de novo* amino acid synthesis.

(B-G) Comparative glucose tracing analysis in control (n=3 biological replicates, in technical triplicates) versus BOS (n=3 biological replicates, in technical triplicates) cells, under conditions with and without non-essential amino acids (NEAA). Results illustrate isotopic enrichment (histograms) and quantification (pooled peak area boxplots) across key metabolites: **(B)** Glucose, **(C)** Fructose-1,6-bisphosphate, **(D)** Pyruvate, **(E)** Citrate, **(F)** Glutamate, and **(G)** Aspartate.

Statistical significance determined by Student's t-test is annotated for each metabolite comparison.

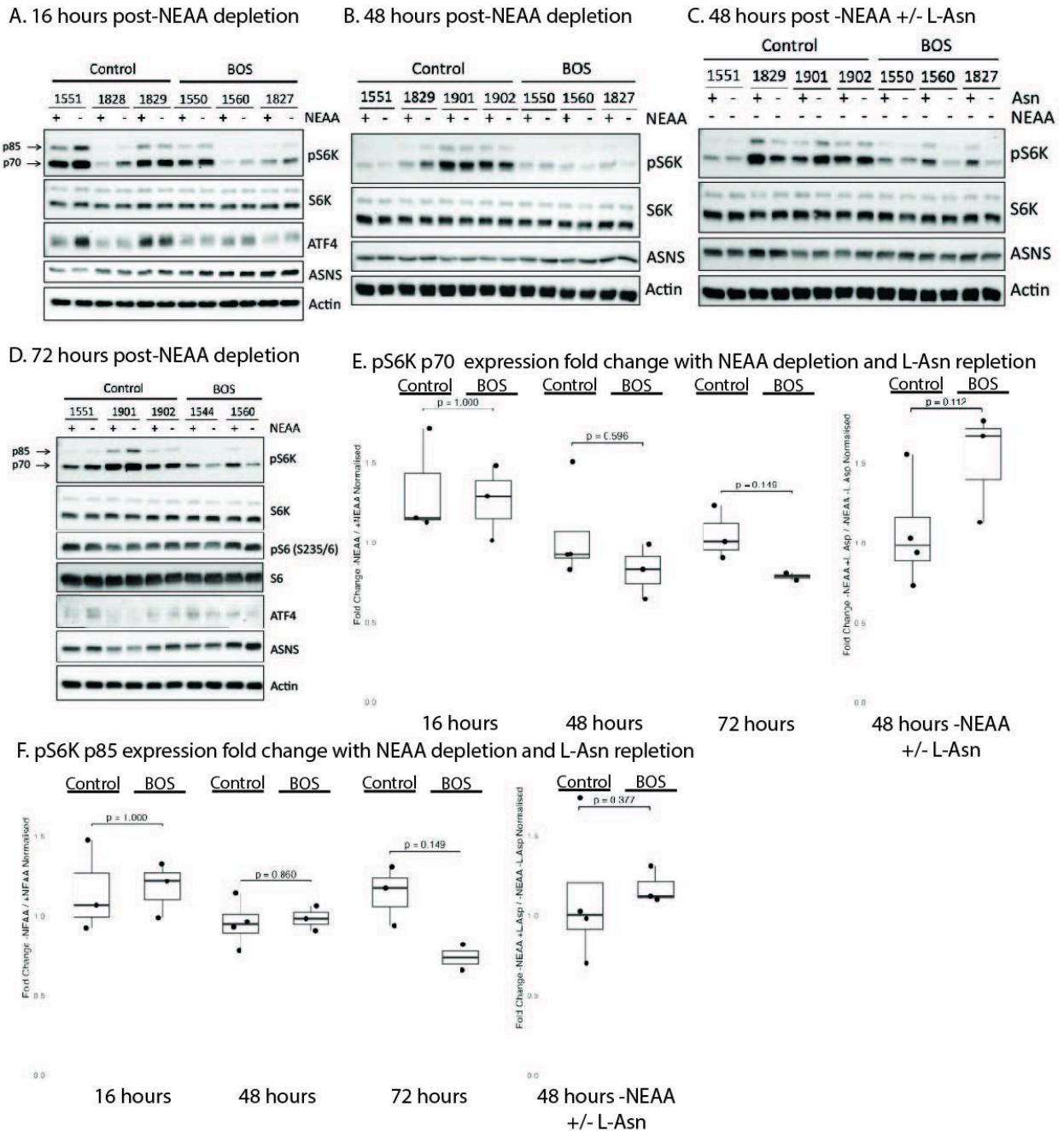


Figure 4-6: NEAA depletion exacerbates mTOR signaling defects in BOS cells, which can be rescue by L-asparagine repletion

Whole cell lysate (WCL) western blots highlight the effects of NEAA depletion on mTOR signaling pathways and the ATF4/ASNS stress response in BOS cells. '+' and '-' symbols

indicate the presence or absence of 1% NEAA, or of 10mM L-asparagine, respectively. Actin is a loading control across all conditions.

(A) WCL of control (n=3) and BOS (n=3) cells reveal early responses to acute NEAA depletion at 16 hours. The blots highlight elevated ASNS levels in BOS cells, both at baseline and upon NEAA depletion.

(B) WCL of control (n=4) and BOS (n=3) cells at 48 hours post-NEAA depletion show a baseline reduction in phosphorylated S6K (pS6K) forms upon NEAA depletion in BOS cells. Controls showed no significant changes with or without NEAA.

(C) WCL of control (n=4) and BOS (n=3) cells, under NEAA-depleted conditions, with and without 10mM L-asparagine repletion after 48 hours. Repletion of L-asparagine rescued pS6K levels in BOS cells to control-like levels. Control samples exhibited no marked differences upon L-asparagine repletion.

(D) WCL of control (n=3) and BOS (n=2) cells at 72 hours post-NEAA depletion highlight the continued reduced levels of pS6K.

(E) Quantitative box plots of the fold change in pS6K p70 expression over time with NEAA depletion, including under L-asparagine repletion.

(F) Quantitative box plots of the fold change in pS6K p85 expression over time with NEAA depletion, including under L-asparagine repletion.

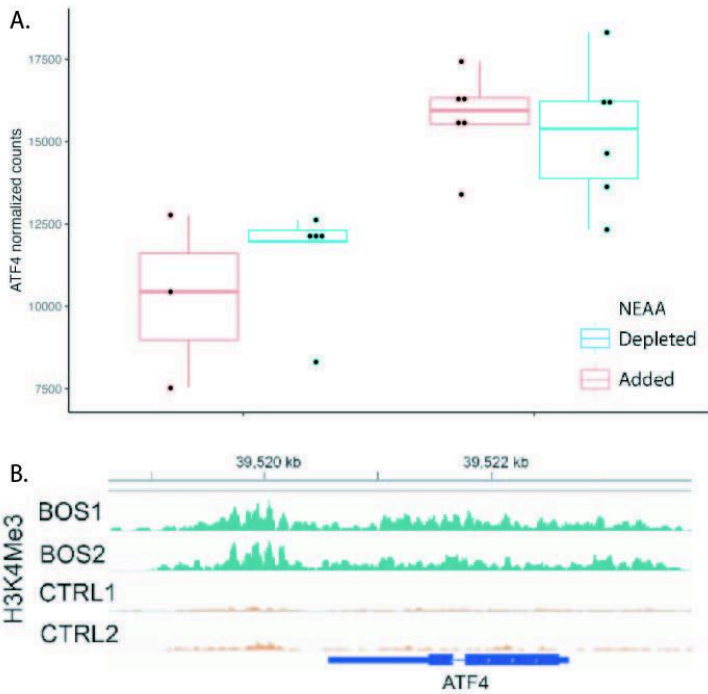


Figure 4-7: Transcriptional upregulation and epigenetic activation of ATF4 in BOS fibroblasts

(A) RNA sequencing normalized expression levels of *ATF4* in BOS and control fibroblasts, both with and without non-essential amino acids (NEAA). The data reveal significantly elevated levels of *ATF4* in BOS cells under both nutrient conditions, compared to controls.

(B) CUT&RUN (Cleavage Under Targets and Release Using Nuclease) data illustrating increased deposition of the histone mark H3K4me3 at the promoter and transcription start site (TSS) of *ATF4* in BOS fibroblasts.

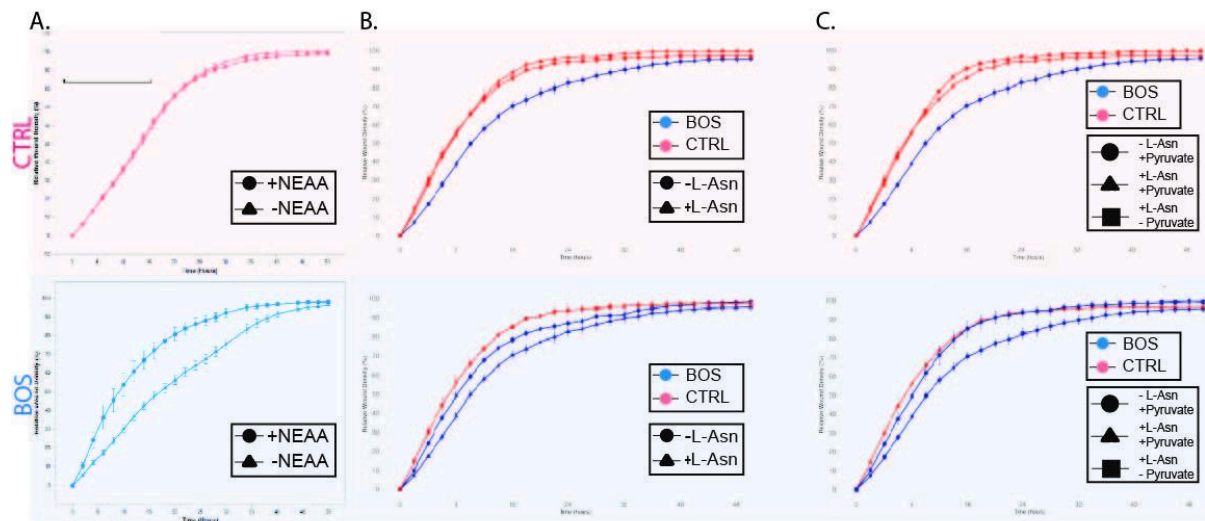


Figure 4-8: BOS cells are dependent on NEAA for cell migration, and partial rescue can be achieved with repletion of L-asparagine

(A) Relative wound density plots show scratch wound healing in control patient fibroblasts (above, red) and BOS fibroblasts (below, blue) under Glucose media (+NEAA, circle points) vs Glucose media (-NEAA, triangle points).

(B) Comparative growth rates post-scratch wound under NEAA depletion (circle points) with 10mM L-asparagine supplementation (triangle points) show no change in growth rate in controls (above), and partial rescue in BOS (below).

(C) Comparative growth rates post-scratch wound under NEAA depletion (circle points) with the removal of pyruvate and addition of 10mM L-asparagine (square point) demonstrate no growth difference in control cells (above) and rescue of BOS cell growth rates (below).

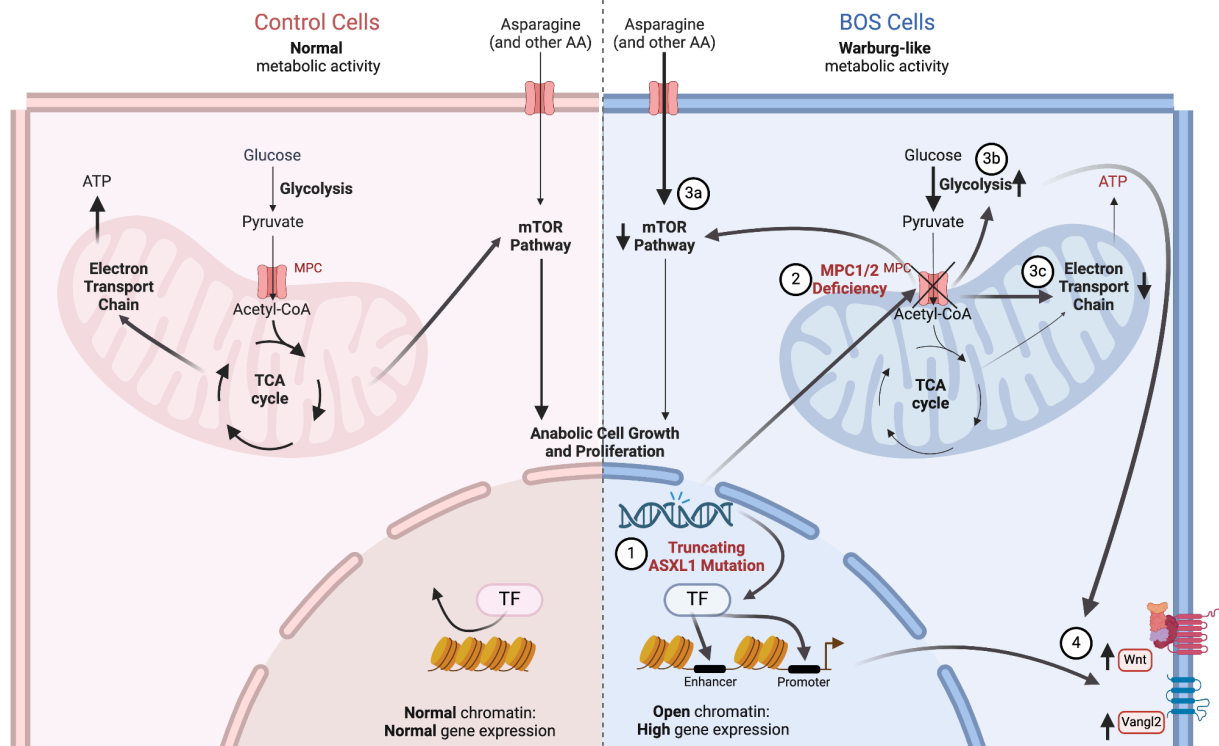


Figure 4-9: ASXL1 mutations drive metabolic dysregulation and mitochondrial dysfunction in Bohring-Opitz syndrome

Multi-Omics Paper Pt#	Metabolism Paper Pt#	Condition	Sex	Age	ASXL1 Mutation		Severity
					cDNA Change NM_015338.6	Protein Change Q8IXJ9	
Pt2	Pt1	BOS	F	9	c.1210C>T	p.R404*	Moderate
Pt3	Pt2	BOS	M	24	c.1910_1922del	p.E635Rfs*15	Moderate
Pt4	Pt3	BOS	M	8	c.1924G>T	p.G642*	Moderate
Pt6	Pt4	BOS	F	15	c.2013_2014delAG	p.C672Wfs*4	Moderate
Pt9	Pt5	BOS	F	3	c.2477delG	p.G826Efs*12	Unknown
Pt14	Pt6	BOS	F	5	c.2981delC	p.P994Lfs*30	Severe
Pt15	Pt7	BOS	F	7	c.3077delG	p.G1026Dfs*21	Severe
Ctrl1	Ctrl1	Control	F	37			
Ctrl2	Ctrl2	Control	M	38			
Ctrl3	Ctrl3	Control	F	40			
Ctrl5	Ctrl4	Control	F	41			
Ctrl50	Ctrl5	Control	F	56			
Ctrl51	Ctrl6	Control	M	55			
Ctrl52	Ctrl7	Control	F	45			
Ctrl53	Ctrl8	Control	M	44			

Table 4-1: Bohring-Opitz syndrome (BOS) patient and control demographics

Bohring-Opitz syndrome (BOS) patients and controls included in our assays are identified with their patient IDs and prior IDs from our previous multi-omics paper ³. *ASXL1* mutation, age at collection, sex, and severity of BOS are shown.

Media Name	Base Media	Sugar Source	Serum	NEAA	Pen-Strep	Glutamine	Sodium Pyruvate	Other Reagents
Glucose media (+NEAA)	DMEM (#11995065)	25mM Glucose (base media)	10% FBS	1X (10mM)	1%	4mM (base media)	1mM (base media)	None
Glucose media (-NEAA)	DMEM (#11995065)	25mM Glucose (base media)	10% FBS	None	1%	4mM (base media)	1mM (base media)	None
Glucose media (+L-Asn)	DMEM (#11995065)	25mM Glucose (base media)	10% FBS	None	1%	4mM (base media)	1mM (base media)	2mM L-Asparagine
Galactose media (+NEAA)	DMEM (#11966025)	10mM D-(+)-Galactose	10% FBS	1X (10mM)	1%	4mM (base media)	2mM	None
Galactose media (-NEAA)	DMEM (#11966025)	10mM D-(+)-Galactose	10% FBS	None	1%	4mM (base media)	2mM	None
Glucose-labelled media (+NEAA)	DMEM (#A1443001)	15mM U-13C6 D-Glucose	10% dialysed FBS	1X (10mM)	1%	1X GlutaMAX (2mM)	None	None
Glucose-labelled media (-NEAA)	DMEM (#A1443001)	15mM U-13C6 D-Glucose	10% dialysed FBS	None	1%	1X GlutaMAX (2mM)	None	None
Gln-labelled media (+NEAA)	DMEM (#A1443001)	15mM D-(+)-Glucose	10% dialysed FBS	1X (10mM)	1%	1mM U-13C5 Glutamine	None	None
Gln-labelled media (-NEAA)	DMEM (#A1443001)	15mM D-(+)-Glucose	10% dialysed FBS	None	1%	1mM U-13C5 Glutamine	None	None

Table 4-2: Media compositions

Component	Supplier	Catalogue	Lot
Gibco™ DMEM, high glucose, pyruvate	ThermoFisher	11995065	
Gibco™ DMEM, no glucose	ThermoFisher	11966025	
Gibco™ DMEM, no glucose, no glutamine, no phenol red	ThermoFisher	A1443001	
Fetal Bovine Serum	Avantor	97068-085	167K22
Gibco™ MEM NEAA Solution (100X)	ThermoFisher	11140050	
Gibco™ Pen-Strep (10,000 U/mL)	ThermoFisher	15140122	
Sodium Pyruvate	ThermoFisher	11360070	
U-13C6 D-Glucose	Cambridge Isotopes	CLM-1396-2	PR-32747A
U-13C5 Glutamine	Cambridge Isotopes	CLM-1822-H-PK	
GlutaMAX (100X)	ThermoFisher	35-050-061	
L-Asparagine	ThermoFisher	371601000	
D-(+)-Glucose	Sigma	G8270	
D-(+)-Galactose	Sigma	G0750	

Table 4-3: Reagents used in media compositions

Acknowledgements:

Scratch wound assays were performed at the UCLA Broad Stem Cell Research Center Microscopy Core Facility.

Seahorse assays were performed at UCLA Mitochondria and Metabolism Core.

References

1. Chen, C., Wang, Z. & Qin, Y. Connections between metabolism and epigenetics: mechanisms and novel anti-cancer strategy. *Front. Pharmacol.* **13**, 935536 (2022).
2. Keating, S. T. & El-Osta, A. Epigenetics and metabolism. *Circ. Res.* **116**, 715–736 (2015).
3. Lin, I. *et al.* Multi-omics of Bohring-Opitz syndrome truncating ASXL1 mutations identify canonical and non-canonical Wnt signaling dysregulation. *JCI Insight* (2023) doi:10.1172/jci.insight.167744.
4. Zhang, P., Xu, M. & Yang, F.-C. The Role of ASXL1/2 and Their Associated Proteins in Malignant Hematopoiesis. *Current Stem Cell Reports* **6**, 6–15 (2020).
5. Russell, B., Tan, W.-H. & Graham, J. M., Jr. Bohring-Opitz syndrome. in *GeneReviews®* (eds. Adam, M. P. *et al.*) (University of Washington, Seattle, Seattle (WA), 2018).
6. Russell, B. E. *et al.* Clinical findings in 39 individuals with Bohring-Opitz syndrome from a global patient-driven registry with implications for tumor surveillance and recurrence risk. *Am. J. Med. Genet. A* **191**, 1050–1058 (2023).
7. Bohring, A. *et al.* Severe end of Opitz trigonocephaly (C) syndrome or new syndrome? *Am. J. Med. Genet.* **85**, 438–446 (1999).
8. Cheong, S.-S. *et al.* The Planar Polarity Component VANGL2 Is a Key Regulator of Mechanosignaling. *Front Cell Dev Biol* **8**, 577201 (2020).
9. Torban, E., Wang, H.-J., Groulx, N. & Gros, P. Independent mutations in mouse Vangl2 that cause neural tube defects in looptail mice impair interaction with members of the Dishevelled family. *J. Biol. Chem.* **279**, 52703–52713 (2004).
10. Torban, E. *et al.* Tissue, cellular and sub-cellular localization of the Vangl2 protein during embryonic development: effect of the Lp mutation. *Gene Expr. Patterns* **7**, 346–354 (2007).
11. Lei, Y.-P. *et al.* VANGL2 mutations in human cranial neural-tube defects. *The New England journal of medicine* vol. 362 2232–2235 (2010).

12. Matheus, F. *et al.* Pathological ASXL1 Mutations and Protein Variants Impair Neural Crest Development. *Stem Cell Reports* **12**, 861–868 (2019).
13. Fan, H. *et al.* Critical role of mTOR in regulating aerobic glycolysis in carcinogenesis (Review). *Int. J. Oncol.* **58**, 9–19 (2021).
14. Linke, M., Fritsch, S. D., Sukhbaatar, N., Hengstschläger, M. & Weichhart, T. mTORC1 and mTORC2 as regulators of cell metabolism in immunity. *FEBS Lett.* **591**, 3089–3103 (2017).
15. Jhanwar-Uniyal, M. *et al.* Diverse signaling mechanisms of mTOR complexes: mTORC1 and mTORC2 in forming a formidable relationship. *Adv. Biol. Regul.* **72**, 51–62 (2019).
16. Youn, H. S. *et al.* Asxl1 deficiency in embryonic fibroblasts leads to cellular senescence via impairment of the AKT-E2F pathway and Ezh2 inactivation. *Sci. Rep.* **7**, 5198 (2017).
17. Fujino, T. *et al.* Mutant ASXL1 induces age-related expansion of phenotypic hematopoietic stem cells through activation of Akt/mTOR pathway. *Nat. Commun.* **12**, 1826 (2021).
18. Krall, A. S. *et al.* Asparagine couples mitochondrial respiration to ATF4 activity and tumor growth. *Cell Metab.* **33**, 1013–1026.e6 (2021).
19. Krall, A. S., Xu, S., Graeber, T. G., Braas, D. & Christofk, H. R. Asparagine promotes cancer cell proliferation through use as an amino acid exchange factor. *Nat. Commun.* **7**, 11457 (2016).
20. Pakos-Zebrucka, K. *et al.* The integrated stress response. *EMBO Rep.* **17**, 1374–1395 (2016).
21. Gong, S. S., Guerrini, L. & Basilico, C. Regulation of asparagine synthetase gene expression by amino acid starvation. *Mol. Cell. Biol.* **11**, 6059–6066 (1991).
22. Meers, M. P., Bryson, T. D., Henikoff, J. G. & Henikoff, S. Improved CUT&RUN chromatin profiling tools. *Elife* **8**, (2019).
23. Aguer, C. *et al.* Galactose enhances oxidative metabolism and reveals mitochondrial dysfunction in human primary muscle cells. *PLoS One* **6**, e28536 (2011).
24. Frey, P. A. The Leloir pathway: a mechanistic imperative for three enzymes to change the

- stereochemical configuration of a single carbon in galactose. *FASEB J.* **10**, 461–470 (1996).
25. Cory, G. Scratch-wound assay. *Methods Mol. Biol.* **769**, 25–30 (2011).
 26. Liang, C.-C., Park, A. Y. & Guan, J.-L. In vitro scratch assay: a convenient and inexpensive method for analysis of cell migration in vitro. *Nat. Protoc.* **2**, 329–333 (2007).
 27. Main, K. A., Mikelis, C. M. & Doçi, C. L. In Vitro Wound Healing Assays to Investigate Epidermal Migration. *Methods Mol. Biol.* **2109**, 147–154 (2020).
 28. Jonkman, J. E. N. *et al.* An introduction to the wound healing assay using live-cell microscopy. *Cell Adh. Migr.* **8**, 440–451 (2014).
 29. Hu, X., Zhang, H., Li, X., Li, Y. & Chen, Z. Activation of mTORC1 in fibroblasts accelerates wound healing and induces fibrosis in mice. *Wound Repair Regen.* **28**, 6–15 (2020).
 30. Scepanovic, G. & Fernandez-Gonzalez, R. mTOR inhibits autophagy to facilitate cell swelling and rapid wound repair. *bioRxiv* 2023.10.10.561758 (2023) doi:10.1101/2023.10.10.561758.
 31. Jiang, H. *et al.* Identification and characterization of novel MPC1 gene variants causing mitochondrial pyruvate carrier deficiency. *J. Inherit. Metab. Dis.* **45**, 264–277 (2022).
 32. Chernyavskij, D. A., Galkin, I. I., Pavlyuchenkova, A. N., Fedorov, A. V. & Chelombitko, M. A. Role of Mitochondria in Intestinal Epithelial Barrier Dysfunction in Inflammatory Bowel Disease. *Mol. Biol.* **57**, 1024–1037 (2023).
 33. Schell, J. C. *et al.* Control of intestinal stem cell function and proliferation by mitochondrial pyruvate metabolism. *Nat. Cell Biol.* **19**, 1027–1036 (2017).
 34. Buchanan, J. L. & Taylor, E. B. Mitochondrial Pyruvate Carrier Function in Health and Disease across the Lifespan. *Biomolecules* **10**, (2020).
 35. Zangari, J., Petrelli, F., Maillot, B. & Martinou, J.-C. The Multifaceted Pyruvate Metabolism: Role of the Mitochondrial Pyruvate Carrier. *Biomolecules* **10**, (2020).
 36. Kondoh, H. *et al.* A high glycolytic flux supports the proliferative potential of murine

- embryonic stem cells. *Antioxid. Redox Signal.* **9**, 293–299 (2007).
37. Folmes, C. D. L. *et al.* Somatic oxidative bioenergetics transitions into pluripotency-dependent glycolysis to facilitate nuclear reprogramming. *Cell Metab.* **14**, 264–271 (2011).
 38. Takubo, K. *et al.* Regulation of glycolysis by Pdk functions as a metabolic checkpoint for cell cycle quiescence in hematopoietic stem cells. *Cell Stem Cell* **12**, 49–61 (2013).
 39. Moussaieff, A. *et al.* Glycolysis-mediated changes in acetyl-CoA and histone acetylation control the early differentiation of embryonic stem cells. *Cell Metab.* **21**, 392–402 (2015).
 40. Sercel, A. J. *et al.* Hypermetabolism and energetic constraints in mitochondrial disorders. *Nat Metab* **6**, 192–195 (2024).
 41. Xia, Y.-K. *et al.* Tumor-derived neomorphic mutations in ASXL1 impairs the BAP1-ASXL1-FOXK1/K2 transcription network. *Protein Cell* **12**, 557–577 (2021).
 42. Awamleh, Z. *et al.* DNA methylation signature associated with Bohring-Opitz syndrome: a new tool for functional classification of variants in ASXL genes. *Eur. J. Hum. Genet.* (2022) doi:10.1038/s41431-022-01083-0.
 43. Vander Heiden, M. G., Cantley, L. C. & Thompson, C. B. Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science* **324**, 1029–1033 (2009).
 44. Bononi, A. *et al.* Germline BAP1 mutations induce a Warburg effect. *Cell Death Differ.* **24**, 1694–1704 (2017).
 45. Szczepanski, A. P. & Wang, L. Emerging multifaceted roles of BAP1 complexes in biological processes. *Cell Death Discov* **7**, 20 (2021).
 46. Bader, D. A. *et al.* Mitochondrial pyruvate import is a metabolic vulnerability in androgen receptor-driven prostate cancer. *Nat Metab* **1**, 70–85 (2019).
 47. Yang, C. *et al.* Glutamine oxidation maintains the TCA cycle and cell survival during impaired mitochondrial pyruvate transport. *Mol. Cell* **56**, 414–424 (2014).
 48. Petrelli, F. *et al.* Mitochondrial pyruvate metabolism regulates the activation of quiescent

adult neural stem cells. *Sci Adv* **9**, eadd5220 (2023).

CHAPTER 5

Expanding the *ASXL1* Mutation Spectrum: Transcriptomic and Metabolic Dysregulation

Underlying Overgrowth Phenotype in Inherited *ASXL1* Missense Mutations

Abstract

Mutations in genes associated with chromatin remodeling underlie a spectrum of congenital abnormalities and neurodevelopmental disorders. Additional sex combs-like 1 (*ASXL1*) is a key chromatin regulator, and *de novo* truncating mutations in *ASXL1* are known to cause the rare genetic disorder Bohring–Opitz syndrome (BOS). BOS is characterized by distinctive neurodevelopmental and physical abnormalities.

In this study, we present a 14-year-old female with overgrowth syndrome, macrocephaly, and normal neurodevelopment, features that diverge from the typical BOS phenotype. This proband was identified to be carrying a paternally-inherited *ASXL1* missense mutation (c.4562C>T p.Ala1521Val) in the plant homeodomain and an additional *KIF22* missense mutation, both variants of unknown significance.

Through genetic, transcriptomic, and metabolomic analyses, we uncovered 104 differentially expressed genes dysregulated in *ASXL1* missense mutation samples across both blood and fibroblast. Notably, the mutation significantly affects gene expression of core components of the mitochondrial respiratory chain. Moreover, metabolomics analysis confirmed significant metabolic alterations consistent with the transcriptomic findings. These findings highlight the distinct metabolic phenotype driven by the *ASXL1* missense mutation in this family, distinct from classical BOS. This study contributes to the literature around inherited *ASXL* mutations, and expands the phenotypic spectrum associated with *ASXL* mutations.

Introduction

Introduction to chromatin remodeling and genetic disorders

Chromatin remodeling is a fundamental process that regulates gene expression and genomic integrity, pivotal to the proper development and functioning of an organism. This dynamic process involves the modification of chromatin architecture, allowing access to genetic information stored within DNA for replication, repair, and transcription. The complexity of chromatin remodeling is underscored by its involvement in a wide array of genetic disorders. Mutations in genes encoding chromatin remodeling proteins can disrupt these essential processes, leading to a spectrum of congenital abnormalities and neurodevelopmental disorders. These conditions manifest a wide range of severity and complexity, often reflecting the diverse roles of chromatin remodeling components in gene regulation. Understanding the link between chromatin remodeling defects and genetic disorders is crucial for examining the molecular mechanisms underlying these conditions and for the development of targeted therapeutic strategies.

ASXL1: role in chromatin remodeling

The Additional sex combs-like 1 (*ASXL1*) gene plays a critical role in chromatin remodeling, acting as a regulator of gene expression through its involvement in histone modification and epigenetic signaling pathways ^{1,2}. Truncating mutations in *ASXL1* cause the rare genetic condition Bohring-Opitz syndrome (BOS; OMIM #605039), which is characterized by a distinct set of physical and neurodevelopmental features ³. Features of BOS include intrauterine growth restriction, severe developmental delay, craniofacial abnormalities, and a unique posture known as the "BOS posture," among other clinical manifestations ⁴. This syndrome reflects the essential role of *ASXL1* in the orchestration of gene expression critical for normal growth and neurological development.

ASXL genes, including *ASXL1*, play a critical role in epigenetic and transcriptional regulation by influencing histone modifications. These modifications occur through interactions with polycomb repressive complexes (PRC) and BRCA-1-associated protein 1 (BAP1), which is essential for modulating histone deubiquitination and methylation⁵⁻⁷. The plant homeodomain (PHD) within *ASXL1* is crucial for its function, mediating these interactions and thereby maintaining chromatin integrity. Truncating mutations in *ASXL1*, leading to BOS, result in the heterozygous loss of this PHD, which is pivotal in the gene's role in chromatin remodeling⁸. The loss of the PHD disrupts the *ASXL1*-BAP1 complex's ability to counteract PRC1-mediated monoubiquitination of histone H2A lysine 119 (H2AK119), a pivotal epigenetic marker. Consequently, truncated forms of *ASXL1* enhance the catalytic activity of BAP1 disproportionately, causing a significant reduction in H2AK119 ubiquitination levels and thus perturb the normal chromatin state^{8,9}. This loss of interaction with PRC1 in truncated *ASXL1* leads to diminished PRC1 activity and unchecked PR-DUB activity, culminating in the widespread de-repression of genes across various tissues¹⁰. This effect underscores the hypothesized dominant negative or gain-of-function impact of truncated *ASXL1* mutations¹¹.

Missense mutations in *ASXL* genes: literature review

Literature on *ASXL* missense mutations paints a different picture of phenotypic outcomes, indicating a broader and less understood aspect of *ASXL* gene mutations. While truncating mutations often lead to well-documented syndromes like BOS, the effects of missense mutations in the *ASXL* genes remain less predictable. For instance, several patients with *ASXL3* missense variants exhibited phenotypes vastly different from those associated with loss-of-function mutations, often inherited from clinically unaffected parents, suggesting a nuanced impact of missense mutations on gene function and disease phenotype¹². Moreover, compound heterozygous mutations in *ASXL3* have been linked to phenotypes overlapping with reported cases of Bainbridge-Ropers syndrome, indicating potential loss-of-function effects¹³.

In this study, we aim to delve into the unique clinical and molecular landscapes shaped by a novel *ASXL1* missense mutation identified in a family. Unlike the truncating mutations associated with BOS, which significantly alter chromatin remodeling and gene expression, missense mutations in *ASXL1* offer a less understood pathogenic pathway. This case study presents a paternally inherited *ASXL1* missense mutation, identified in a patient with an overgrowth syndrome, yet exhibiting normal neurodevelopmental progress. As far as we know, this is the first known germline case of *ASXL1* missense mutations, and the first known case of inherited *ASXL1* missense mutations.

Methods

Cell culture

Primary fibroblast cultures were established and maintained following the protocols outlined by Lin et al. (2023) ¹⁰.

DNA/RNA extraction

Genomic DNA was extracted from cultured fibroblasts using the Quick-DNA Miniprep Kit (Zymo, Cat#D3024) and total RNA was extracted using the RNeasy Plus Mini Kit (Qiagen, #74136), adhering to manufacturer protocol. The RNA integrity and concentration were assessed with TapeStation and Qubit respectively before proceeding with RNA sequencing.

Genetic and transcriptomic analysis

Transcriptomic profiling was conducted through RNA sequencing (RNA-seq), as per methodology detailed by Lin et al. (2023) ¹⁰. The RNA sequencing data was analyzed according to the established pipeline outlined by Lin et al. (2023) ¹⁰, with adjustment for patient sex and technical replicates where applicable. Technical replicates prepared by different individuals clustered together and were collapsed in downstream analyses.

Sanger sequencing was performed through Genewiz and visualized using SnapGene software.

Metabolomic analysis

Fibroblast cells were seeded in six-well plates, with or without 1% non-essential amino acids (NEAA) and metabolites were extracted at 70–80% confluence per established protocol ¹⁴. Dried metabolites were stored at -80°C prior to mass spectrometry analysis. Analysis was conducted using the Christofk Lab Grapha platform.

ESM1b

ESM1b pathogenicity scores for missense mutations were identified using the publicly available score predictions generated by Brandes et al. (2023) ¹⁵.

AlphaMissense

AlphaMissense pathogenicity scores for missense mutations were generated using the model developed by Cheng et al. (2023) ¹⁶.

Results

This case presentation focuses on a 14-year-old female proband diagnosed with a distinct phenotype from the traditional phenotypic spectrum associated with *ASXL1* mutations. The proband's clinical picture is characterized by an overgrowth syndrome, macrocephaly, and normal neurodevelopment, distinctly lacking the severe neurodevelopmental delay typically observed in BOS (**Figure 5-1A**). Genotyping revealed the presence of a paternally inherited *ASXL1* missense mutation (c.4562C>T p.Ala1521Val) located in the PHD (**Figure 5-1A, Figure 5-1B**), and a kinesin family member 22 (*KIF22*) missense mutation (c.130G>A p.Val44Ile), both classified as variants of unknown significance (VUS). Exome studies confirmed the mutations in the proband and father, and not in the mother.

Clinical history and physical findings

The proband's medical history reveals a history of proportionate overgrowth, first noted at 20 months of age. At 14 years of age, the proband had a height of 5'7" (93rd percentile, +1.4SD, **Figure 5-1C**) and weighed 290 lbs (100th percentile, +4.28SD, **Figure 5-1D**), with a BMI of 46.42 (100th percentile, +3.88SD, **Figure 5-2A**). Her growth metrics significantly surpassed standard deviations for her age, pointing to a pronounced overgrowth syndrome. Other clinical features of note include obesity (100th percentile), macrocephaly (100th percentile, +6.39SD, **Figure 5-1E**), increased bone age, premature adrenarche, and dysfunctional uterine bleeding attributed to ovarian androgen hypersecretion.

The *ASXL1* missense mutation was paternally inherited, and may co-segregate with increased height. The proband's father has a height of 5'11" while the mother has a height of 5'3" (**Figure 5-1D**). On the paternal side of the family, there are also multiple cousins, aunts and uncles over 6' tall, whom have not been genetically evaluated (**Figure 5-1B**). However, the proband's 9-year-old sister's growth parameters (99th percentile for weight and 75th percentile for height) without inheritance of the *ASXL1* missense mutation and without pubertal development signs suggest a potential genetic predisposition towards overgrowth within the family, albeit with varying manifestations.

Family history

The proband's paternal lineage exhibits a notable history of overgrowth, hemihypertrophy, early facial hair development in males, and other symptoms mirroring the proband's condition, suggesting a hereditary pattern of inheritance (**Figure 5-1B**). Exome, sanger sequencing, and RNA-seq confirmed that the proband and father both have the *ASXL1* missense mutation. The mother, sister, and paternal grandmother are confirmed to have neither *ASXL1* nor *KIF22* mutation (**Figure 5-2B**).

The maternal side reports significant health issues, including precancerous conditions and lung cancer, though no direct linkage to the proband's condition is evident (**Figure 5-1B**).

Variant assessment

Genetic analysis confirms the presence of an *ASXL1* missense mutation affecting the PHD domain, an essential component for its role in chromatin remodeling, as well as a *KIF22* missense mutation, where *KIF22* has been shown to play a role in bone growth. Clinically, the proband displays a wide array of symptoms including macrocephaly, accelerated skeletal maturation, polycystic ovaries, and metabolic dysregulation, among others. These symptoms underscore a multisystem impact, pointing towards *ASXL1* missense mutations driving this dysregulation. *KIF22* missense mutation adds a confounding factor and could be acting in a compounding manner, and may explain the increased bone age phenotype.

Using *ESM1b*, a protein language model for predicting the effects of coding variants, the *ASXL1* missense mutation had a predicted pathogenicity score of -10.792 while the *KIF22* missense mutation had a predicted pathogenicity score of -6.62. *ESM1b* defines likely pathogenic missense mutations as having a score less than -7.5.

Similarly, with *AlphaMissense* annotations, the *ASXL1* missense mutation was predicted as likely pathogenic, with a pathogenicity score of 0.9541 while the *KIF22* missense mutation was predicted as likely benign, with a pathogenicity score of 0.2687.

Clinical management and follow-up:

The proband's management has involved multidisciplinary care, including pediatric endocrinology for her growth and metabolic concerns, gynecological oversight for menstrual irregularities, and continued monitoring for potential cardiovascular complications suggested by an enlarged aortic valve. The intricate interplay of genetic mutations presents a clinical challenge, requiring ongoing evaluation to understand the full scope of phenotypic impact and to tailor her management accordingly.

The atypical phenotypic spectrum warrants further investigation.

Transcriptomic profiling of *ASXL1* missense mutation identifies signature distinct from Bohring-Opitz syndrome

The transcriptomic analysis was designed to dissect the genetic expression patterns resulting from the *ASXL1* missense mutation in comparison to both the typical BOS signature, caused by truncating *ASXL1* mutations, and unaffected controls. Utilizing both blood and fibroblast RNA-seq, we aimed to capture a comprehensive view of the transcriptomic alterations across different cell types, offering insights into systemic and tissue-specific effects of the mutations.

Blood samples from the proband, father, and mother, and control individuals underwent RNA-seq to evaluate the expression patterns associated with the missense *ASXL1* mutation. Notably, the analysis of allele ratios in the father highlighted an intriguing skew towards the wild-type allele for the father (77% wild type, 23% mutation), suggesting a compensatory mechanism or selective expression favoring the wild type *ASXL1* allele in the father. The proband exhibited an equal distribution of allele ratios at the mutation site (50% wild type, 50% mutation). This pattern was distinct from what would typically be observed in germline mutations, indicating a nuanced interaction between the mutated and wild-type alleles in response to the *ASXL1* missense mutation.

The principal component analysis (PCA) identified that PC3, accounting for 12% of the variance, distinctly separated the *ASXL1* missense mutation carriers from both BOS samples and healthy controls (**Figure 5-3A**). Unsupervised clustering identified strong familial effect, and distinct clustering separate from BOS samples and control samples (**Figure 5-3B**). DESeq2 analysis identified 4468 differentially expressed genes (DEGs), of which 2060 (46.11%) DEGs were upregulated and 2408 (53.89%) DEGs were downregulated in blood samples (**Figure 5-3C, Table 5-1**). 2723/4468 (60.94%) of these DEGs were protein-coding, with 1646 (60.44%) DEGs being upregulated and 1077 (39.56%) DEGs being downregulated. Intriguingly, 8 of the 30 most strongly dysregulated DEGs encoded oxidative phosphorylation subunits. We identified

significant enrichment in gene ontologies related to ATP metabolic processes and histone modifications among other biological pathways (**Figure 5-3D, Table 5-2**). Unlike BOS samples¹⁰, there was no significant dysregulation in the Wnt pathway, underscoring the distinct molecular impacts of the *ASXL1* missense mutations (**Figure 5-3D**).

The expression levels of *ASXL1* in missense mutation carriers were not significantly different from controls, indicating an absence of nonsense-mediated decay (NMD) - a feature also observed in BOS samples (**Figure 5-3E**). Similarly, *Vang-like 2* (*VANGL2*) expression did not show significant changes in *ASXL1* missense samples, in contrast to its strong upregulation in BOS patient blood compared to control¹⁰, further illustrating the divergent transcriptional impacts of the *ASXL1* mutations (**Figure 5-3F**).

Patient-derived fibroblasts were cultured from the proband, father, mother, sister, and paternal grandmother. RNA-seq analysis, as well as Sanger sequencing (**Figure 5-2B**), confirmed expression of the *ASXL1* missense mutation in the proband and father. The mutation was absent in the mother, sister, and paternal grandmother.

The PCA of fibroblast RNA-seq identified that PC1, accounting for 30% of the variance, clustered samples with *ASXL1* missense mutations closer to that of controls than BOS samples (**Figure 5-4A**). Unsupervised clustering using our previously derived transcriptomic signature for BOS showed distinct clustering separate from BOS samples. Instead, *ASXL1* missense samples clustered with control samples (**Figure 5-5A**). Similarly, *ASXL1* missense fibroblasts showed stronger clustering with control samples with unsupervised clustering using the *ASXL1* missense DEGs (**Figure 5-3B**).

DESeq2 analysis was conducted with (**Figure 5-5B**) and without (**Figure 5-4C**) technical replicate collapse. We identified 1565 DEGs driven by missense *ASXL1* mutations. Of these, 1356/1565 (86.65%) of DEGs downregulated ($\log_2$ fold change ($\log_2FC$) < 0), and 209/1565 (13.45%) of DEGs upregulated ($\log_2FC$ > 0) (**Figure 5-5B**). This negative skew of dysregulation is in stark contrast to the de-repression of a key subset of genes identified in BOS fibroblast

RNA-seq ¹⁰. After filtering for protein-coding genes, there remained 577/738 (78.18%) downregulated DEGs and 161/738 (21.82%) upregulated DEGs (**Table 5-3**). Technical replicates prepared by different individuals clustered together and were collapsed in downstream analyses.

One of the most significantly dysregulated DEGs in patient-derived fibroblasts was *thyrotropin releasing hormone* (*TRH*, log2FC = 5.87) which controls the secretion of thyroid-stimulating hormone (TSH) and is involved in thyroid hormone synthesis regulation. Other significantly dysregulation genes of interest included upregulation of multiple *nuclear receptor subfamily 4 members, retinol binding protein 4* (*RBP4*) and *activating transcription factor 3* (*ATF3*) which all act in the nucleus to regulate transcription.

We identified significant enrichment in gene ontologies related to mitotic cell cycle, response to oxidative stress, ossification, and Wnt signaling (**Figure 5-4D, Table 5-4**).

Again, in fibroblasts, the expression levels of *ASXL1* in missense mutation carriers were not significantly different from controls, indicating an absence of NMD (**Figure 5-4E**). Similarly, *VANGL2* expression did not show significant changes in *ASXL1* missense samples, in contrast to its strong upregulation in BOS patient fibroblasts ¹⁰ (**Figure 5-4F**).

Furthermore, direct comparison of the *ASXL1* missense fibroblast to BOS fibroblast samples revealed distinct separation of the samples along PC1, accounting for 53% of the variance (**Figure 5-6A**). Differential gene expression analysis identified 6497 DEGs, with 3741 (57.58%) downregulated and 2756 (42.42%) upregulated (**Figure 5-6B, Table 5-5**). Of the most significantly dysregulated genes, was *RAS like estrogen regulated growth inhibitor* (*RERG*, log2FC = -10.15) a member of the RAS superfamily of GTPases, which acts to inhibit cell proliferation and tumor formation. We also found significant downregulation *Ras-Related GTP-Binding Protein D* (*RRAGD*, log2FC = -9.08) which is a G protein that regulates the mammalian target of rapamycin complex 1 (mTORC1) signaling cascade ^{18–20}, and *Sp9 transcription factor* (*SP9*, log2FC = -8.99), which is a transcription factor involved in limb

development, and predicted to be involved in regulation of transcription by RNA polymerase II, and chromatin functions. *ASXL1* is known to interact with RNA polymerase II ¹⁷ so this could provide insight into the mechanistic dysregulation driven by missense mutations in the *ASXL1* PHD.

Gene ontology analysis of these DEGs identified dysregulation of key pathways including cell growth, ossification, hemostasis and coagulation, and response to oxygen levels (**Figure 5-6C, Table 5-6**).

Transcriptomic profiling of *ASXL1* missense mutation across tissues identifies significant dysregulation of oxidative phosphorylation subunits

Following the analyses of blood and fibroblast RNA-seq to explore the transcriptomic landscape associated with the *ASXL1* missense mutation, a comprehensive integration of the DEGs was conducted on both datasets. This integrative approach aimed to identify common pathways and gene expression patterns altered in both systemic (blood) and localized (fibroblast) cellular environments.

From the initial sets of 4468 DEGs identified in blood and 1565 DEGs in fibroblast RNA-seq, a subset of 104 common DEGs was identified to be significantly dysregulated (**Figure 5-5C**). A significant finding from this integrated analysis was the identification of 11 genes related to the oxidative phosphorylation pathway that were dysregulated in both datasets (**Table 5-7**). Notably, six of these genes are core subunits of the NADH:ubiquinone oxidoreductase complex (Complex I of the mitochondrial respiratory chain), three of these genes encoded cytochrome b and cytochrome c oxidase I and cytochrome C oxidase III, and one of these genes encoded a subunit of the ATP synthase (**Figure 5-7**). Complex I reduces ubiquinone to ubiquinol, and the cytochrome bc₁ complex then reoxidises ubiquinol to reduce oxygen to water at Complex IV ²¹⁻²³. This discovery points to a substantial impact of the *ASXL1* missense mutation on mitochondrial function and energy metabolism.

Intriguingly, in blood samples, all identified mitochondrially encoded oxidative phosphorylation genes were significantly downregulated, with log2FC ranging from -6.68 to -1.13. This suggests a profound suppression of mitochondrial oxidative phosphorylation activity in the systemic environment. Conversely, in fibroblast samples, these genes were significantly upregulated, with log2FC values ranging from 1.06 to 1.64. Notably, these encompass the small percentage of upregulated genes (13.4%) identified in fibroblast RNA-seq. This upregulation in fibroblasts highlights a contrasting cellular response to the mutation, potentially informing the patient phenotype.

Metabolomic profiling in *ASXL1* missense mutation identifies reduction in glycolysis and generation of energy carriers

Metabolomic analysis of fibroblasts from the proband unveiled significant alterations in metabolic pathways, including a notable downregulation of key glycolytic intermediates ($p < 0.001$) including glycerol-3-phosphate, phosphoglycerate, and phosphoenolpyruvate when compared to control fibroblasts (**Figure 5-8A**). This contrasts sharply with BOS-derived fibroblasts, which exhibit an upregulation of glycolytic intermediates (**Figure 5-8B**), suggesting a distinct metabolic phenotype driven by the *ASXL1* missense mutation.

Intriguing, depletion of exogenous NEAA results in downregulation of key energy carriers NADPH ($p < 0.001$) and NAD⁺ ($p < 0.01$), and further depletion of glycolytic intermediates and the tricarboxylic acid (TCA) cycle (**Figure 5-8C**, **Figure 5-8D**). While no significant change was identified in the NEAA depleted state for control and BOS cells for NAD⁺ (**Figure 5-8E**) or NADPH (**Figure 5-8F**), suggesting a distinct response to NEAA depletion in *ASXL1* missense fibroblasts. Other significantly dysregulated metabolic intermediates include downregulation of the amino acids glycine and proline, which are components of the exogenous NEAA, and downregulation of phosphoglycerate and alpha-ketoglutarate, which are intermediates of glycolysis and the TCA cycle respectively.

Gene ontology pathway analysis identified significant pathway enrichment in beta-Alanine metabolism (p (Fisher) = 0.009), glycolysis / gluconeogenesis (p (Fisher) = 0.025), among other biological pathways (**Table 5-8**). These findings illuminate the profound impact of this mutation on cellular metabolism, offering new insights into the role of *ASXL1* missense mutations in disease pathogenesis.

Discussion

This study describes a novel phenotype linked to a paternally inherited *ASXL1* missense mutation, characterized by an overgrowth syndrome with macrocephaly and normal neurodevelopment. This presentation starkly contrasts with classical BOS, which is caused by *de novo* truncating mutations in *ASXL1*²⁴. The divergence in clinical manifestations between the *ASXL1* missense family and classical BOS cases underscores the mutation's unique impact on *ASXL1* function, and prompted our investigation into the role of chromatin remodeling gene mutations and their intersection with metabolic processes.

Truncating mutations causing BOS occur 5' of the *ASXL1* PHD and lead to the hyperstabilization and hyperactivation of the truncated *ASXL1*-BAP1 complex⁸. This results in subsequent dysregulation of epigenetic markers such as H2AK119ub. In contrast, the mutation within the PHD domain as observed in our case likely alters the normal interaction mechanisms of *ASXL1* with its molecular partners. This alteration could lead to a distinctly aberrant functional interaction, driving the unique clinical presentation observed.

Genetic interplay and clinical implications of *ASXL1* and *KIF22* missense mutations

The clinical manifestations observed in the proband, including macrocephaly, accelerated skeletal maturation, polycystic ovaries, and metabolic dysregulation, are largely attributed to the *ASXL1* missense mutation impacting the PHD domain, crucial for chromatin remodeling. This mutation is predicted by *ESM1b* and *AlphaMissense* as likely pathogenic.

These findings align with prior research, such as the case study by Efthymiou et al. (2019), which documented a case of segmental overgrowth associated with an *ASXL1* mutation, albeit with a different mutation and a more restricted phenotype, with overgrowth restricted to the craniofacial region ²⁸. This suggests that different mutations within the same gene can lead to diverse phenotypic spectrums.

However, the broader implication of missense *ASXL1* mutations in the PHD, include associations with distinct cancer types recorded in the Catalogue of Somatic Mutations in Cancer (COSMIC), indicating a potential oncogenic role. COSMIC identified 28 samples with mutations (18 missense, 6 nonsense, 3 synonymous, 1 frameshift truncating) in the PHD region of *ASXL1* in their cohort of samples with *ASXL1* mutations (n=4714). In fact, this exact mutation (p.Ala1521Val) was identified in a nasopharyngeal tumour (#P-04-27987-Tumor-SM-7CKBE) and a laryngeal cancer (#TIS54078F1D1L1) ^{25,26}. This dual role in developmental disorders and cancer progression emphasizes the need for deeper investigation into the *ASXL1* pathways and their mechanistic underpinnings in disease causation.

Additionally, the presence of a *KIF22* missense mutation in the proband introduces a complex layer to the clinical presentation. Although this mutation received a lower pathogenicity score, suggesting it might be benign, its established role in skeletal development could contribute to the proband's accelerated bone age. This compound effect, where the *KIF22* mutation possibly exacerbates certain phenotypes driven by *ASXL1* mutations, illustrates the intricate genetic interplay affecting the proband's condition. This scenario is less typical than isolated *ASXL1* effects and indicates the necessity of considering multiple genetic factors when diagnosing and treating complex multisystem disorders.

Differential expression analysis identifies systemic oxidative phosphorylation dysregulation in *ASXL1* missense patients

The differential expression analysis conducted on RNA-seq data from blood and fibroblast samples of subjects with the *ASXL1* missense mutation highlights a profound and complex alteration in mitochondrial functions, particularly in the oxidative phosphorylation pathway. This analysis identified significant dysregulation in critical components, such as the NADH:ubiquinone oxidoreductase complex (Complex I) and cytochrome components of the mitochondrial respiratory chain. Notably, the contrasting cellular responses—downregulation in blood and upregulation in fibroblasts—suggest distinct, tissue-specific impacts of the *ASXL1* mutation on mitochondrial dynamics.

The involvement of mitochondrial dysfunction across different cell types provides a nuanced insight into the cellular energetics alterations induced by nuclear gene mutations. In blood, the significant downregulation of mitochondrial genes may represent a systemic response to reduce energy production, which could potentially serve as a protective mechanism against excessive ROS production and its associated risks, such as oxidative damage and oncogenesis ²⁷. On the other hand, the upregulation observed in fibroblasts might be an adaptive response aimed at meeting increased metabolic demands typically associated with cell growth and repair processes. This upregulation could drive cellular processes that contribute to the phenotypic characteristics observed in the proband, such as accelerated cell proliferation and growth, possibly explaining the overgrowth phenotypes.

Although mitochondrial genetics is predominantly maternally inherited, the interplay between nuclear and mitochondrial DNA complicates inheritance patterns, particularly for diseases involving chromatin remodeling genes like *ASXL1* ^{29–31}.

Integration of Differential Gene Expression and Metabolomic Profiling Identify Systemic Oxidative Phosphorylation Dysregulation in *ASXL1* Missense Patients

Our metabolomic and transcriptomic analyses revealed a complex interplay between metabolic and mitochondrial pathways affected by the *ASXL1* missense mutation, which

appears to have distinct, tissue-specific effects on cellular metabolism and mitochondrial function. The observed downregulation of glycolytic intermediates in the fibroblasts from the proband, including glycerol-3-phosphate, phosphoglycerate, and phosphoenolpyruvate, in contrast to their upregulation in BOS-derived fibroblasts, suggests a significant rerouting of glycolytic flux in response to this mutation. This differential regulation of glycolysis might be compensatory to mitochondrial dysfunction observed in the oxidative phosphorylation pathway as evidenced by RNA-seq data.

In fibroblasts, the upregulation of components of the oxidative phosphorylation pathway, despite a decrease in glycolytic intermediates, might indicate an adaptive increase in mitochondrial metabolic activity to counterbalance the energy deficit caused by impaired glycolysis. This is supported by the depletion of exogenous non-essential amino acids leading to further downregulation of glycolytic intermediates and key energy carriers like NADPH and NAD⁺. The specific downregulation of these metabolites in *ASXL1* missense fibroblasts, but not in controls or BOS cells, underscores the unique metabolic vulnerabilities driven by the *ASXL1* missense mutation.

These findings raise important questions about the metabolic reprogramming in cells with the *ASXL1* mutation such as whether the upregulation of oxidative phosphorylation in patient fibroblasts represent a failed compensatory mechanism or an effective adaptive strategy. Future studies should aim to dissect these complex metabolic shifts using targeted metabolic flux analysis and explore the possibility of leveraging these vulnerabilities for therapeutic interventions. Additionally, examining these pathways in a model organism or through in vitro systems could provide deeper insights into the mechanistic underpinnings of these changes and help in the development of more precise treatment strategies that target specific metabolic pathways altered by *ASXL1* mutations.

Our findings contribute to the growing body of evidence linking chromatin remodeling disturbances to mitochondrial dysfunction and altered cellular metabolism. Previous studies

have indicated that mutations in complex I subunits can significantly impact ROS levels and cellular signaling pathways, including those mediated by TOR signaling, which plays a pivotal role in responding to mitochondrial respiratory dysfunction^{32,33}. Our observations suggest that the *ASXL1* missense mutation may drive similar complex signaling cascades, impacting growth and development through altered mitochondrial function. Future research should prioritize functional studies to dissect the precise mechanisms by which *ASXL1* missense mutations, particularly mutations in the PHD, influence chromatin regulation and gene expression.

Conclusion

In conclusion, our study presents, to our knowledge, the first case of *ASXL1* missense mutations, and the first case of inherited *ASXL1* missense mutations. The phenotypic and functional characterisation of this family highlights the interplay between chromatin remodeling, mitochondrial function, and cellular metabolism. The distinct clinical presentation and the underlying molecular profiles observed in this family provide a basis for further exploration of targeted therapies and interventions.

the protein structure. Germline truncating mutations associated with Bohring-Opitz syndrome are indicated in grey.

(B) Family pedigree illustrating the inheritance pattern of the *ASXL1* mutation and overgrowth phenotype likely associated with this mutation. The proband is indicated with an arrow. Affected or likely affected individuals are shaded in yellow, denoting overgrowth. Individuals with a history of cancer are indicated in blue.

WHO Growth Charts for Canada for girls aged 2 to 20 years were used to show:

(C) Height trajectory of the proband across four different time points, consistently registering above the 90th percentile and **(D)** weight for age of the proband across four different time points consistently at the 100th percentile.

(E) Head circumference of the proband charted on the United States Head Circumference Growth Reference Charts for girls aged 2 to 20 years. Two different time points are represented by triangles, both at the 100th percentile, reflecting macrocephaly.

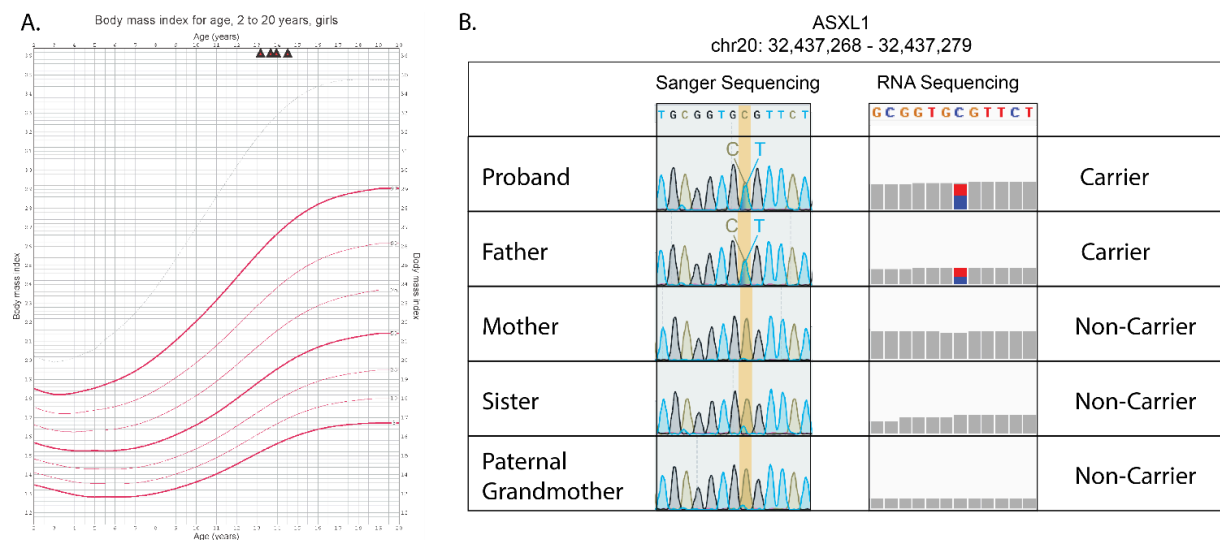


Figure 5-2: Molecular and phenotypic characterisation of inherited *ASXL1* missense mutation (Extended)

(A) Body mass index for age, 2 to 20 years, girls, based on the WHO Growth Charts for Canada, showing four different time points for the proband consistently at the 100th percentile.

(B) Sanger sequencing traces visualised using *SnapGene* (left) and RNA sequencing visualised using *Integrative Genomics Viewer* (right) for patient-derived fibroblasts from the proband, father, mother, sister, and paternal grandmother at the *ASXL1* mutation locus. Heterozygosity was identified in both the proband and father.

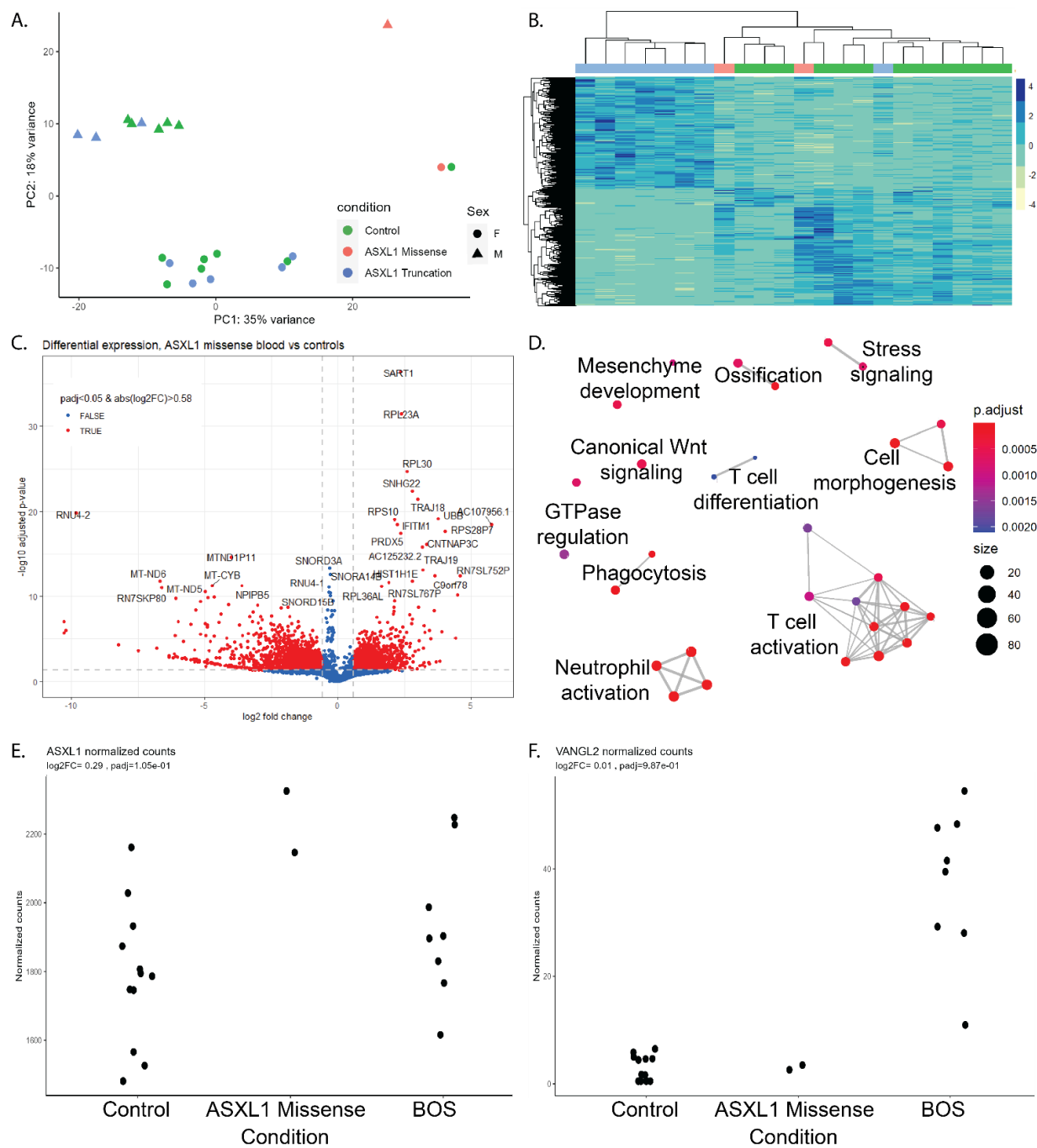


Figure 5-3: Blood transcriptomic profiles and gene expression in *ASXL1* missense samples

(A) Principal component analysis (PCA) of blood RNA sequencing (RNA-seq) data illustrates distinct separation between samples from patients with *ASXL1* missense mutations, *ASXL1* truncating mutations which cause Bohring-Opitz syndrome (BOS), and control samples.

(B) Heatmap of blood RNA-seq differentially expressed genes (DEGs) for *ASXL1* missense samples compared to controls are depicted, along with the BOS samples.

(C) Volcano plot of *ASXL1* missense blood samples compared to control blood samples RNA-seq DEGs, with log₂ fold change on the x-axis and -log₁₀ adjusted p-value on the y-axis.

(D) Gene ontology of *ASXL1* missense blood DEGs identifies dysregulation of key pathways including ribonucleoprotein complex organisation, ATP metabolic processes and histone modification.

(E) *ASXL1* normalised counts showing control, BOS, *ASXL1* missense. *ASXL1* missense compared to control is a log₂FC = 0.29, padj 0.105.

(F) *VANGL2* normalised counts showing control, BOS, *ASXL1* missense. *ASXL1* missense compared to control is a log₂FC = 0.01, padj 0.987.

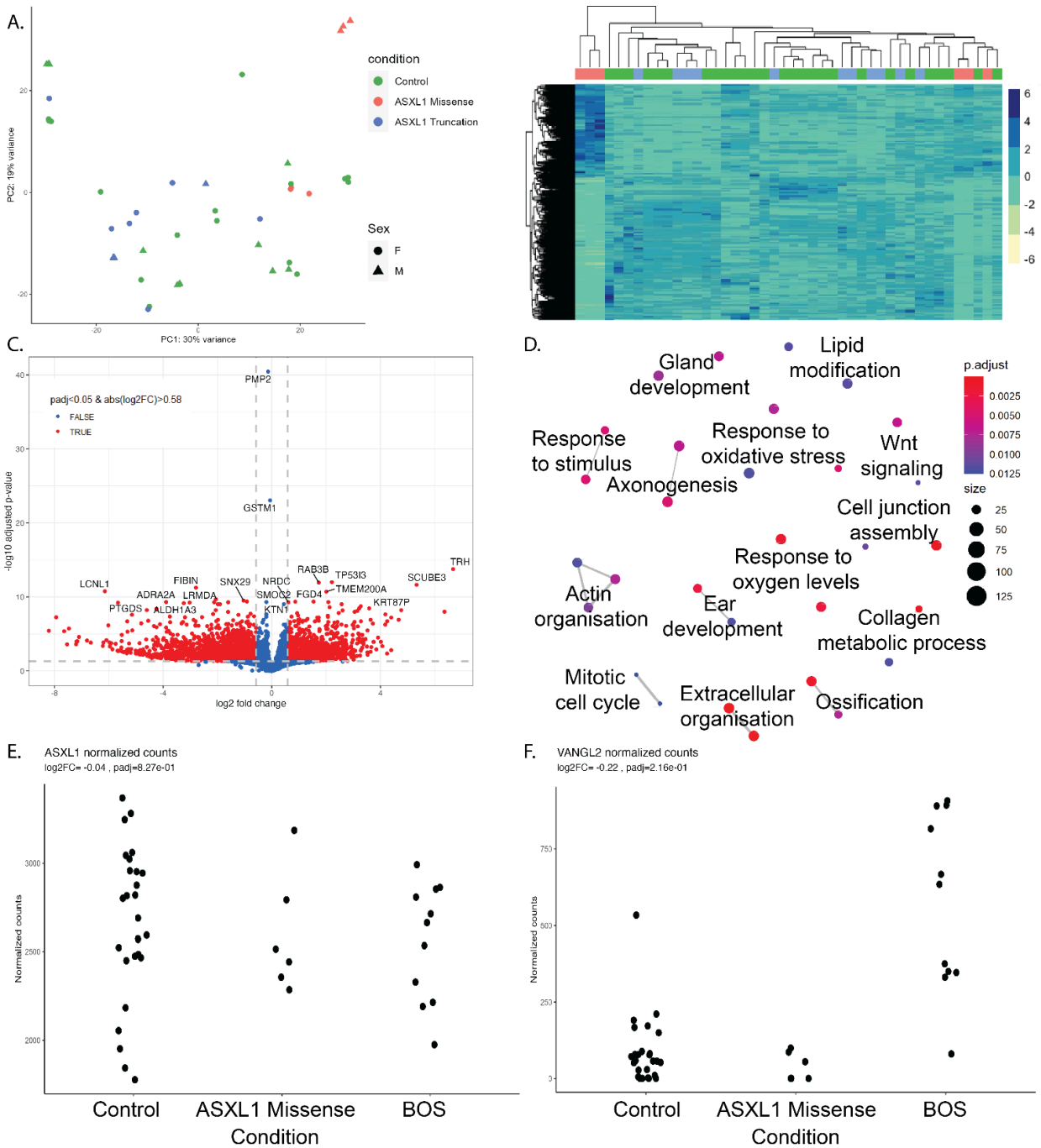


Figure 5-4: Patient-derived fibroblast transcriptomic profiles and gene expression in ASXL1 missense samples

(A) Principal component analysis (PCA) of patient-derived fibroblast RNA sequencing (RNA-seq) data for samples from patients with *ASXL1* missense mutations, *ASXL1* truncating

mutations which cause Bohring-Opitz syndrome (BOS), and control samples. Independently generated technical replicates cluster together.

(B) Heatmap of patient-derived fibroblasts RNA-seq for *ASXL1* missense samples, control samples, and BOS samples, using previously identified BOS transcriptomic signature.

(C) Volcano plot of *ASXL1* missense patient-derived fibroblast samples compared to control fibroblast samples RNA-seq differentially expressed genes (DEGs), with log2 fold change on the x-axis and -log10 adjusted p-value on the y-axis.

(D) Gene ontology of *ASXL1* missense blood DEGs identifies dysregulation of key pathways including response to oxidative stress, mitotic cell cycle, and Wnt signaling.

(E) *ASXL1* normalised counts showing control, BOS, *ASXL1* missense. *ASXL1* missense compared to control is a log2FC = -0.04, padj 0.827.

(F) *VANGL2* normalised counts showing control, BOS, *ASXL1* missense. *ASXL1* missense compared to control is a log2FC = -0.22, padj 0.216.

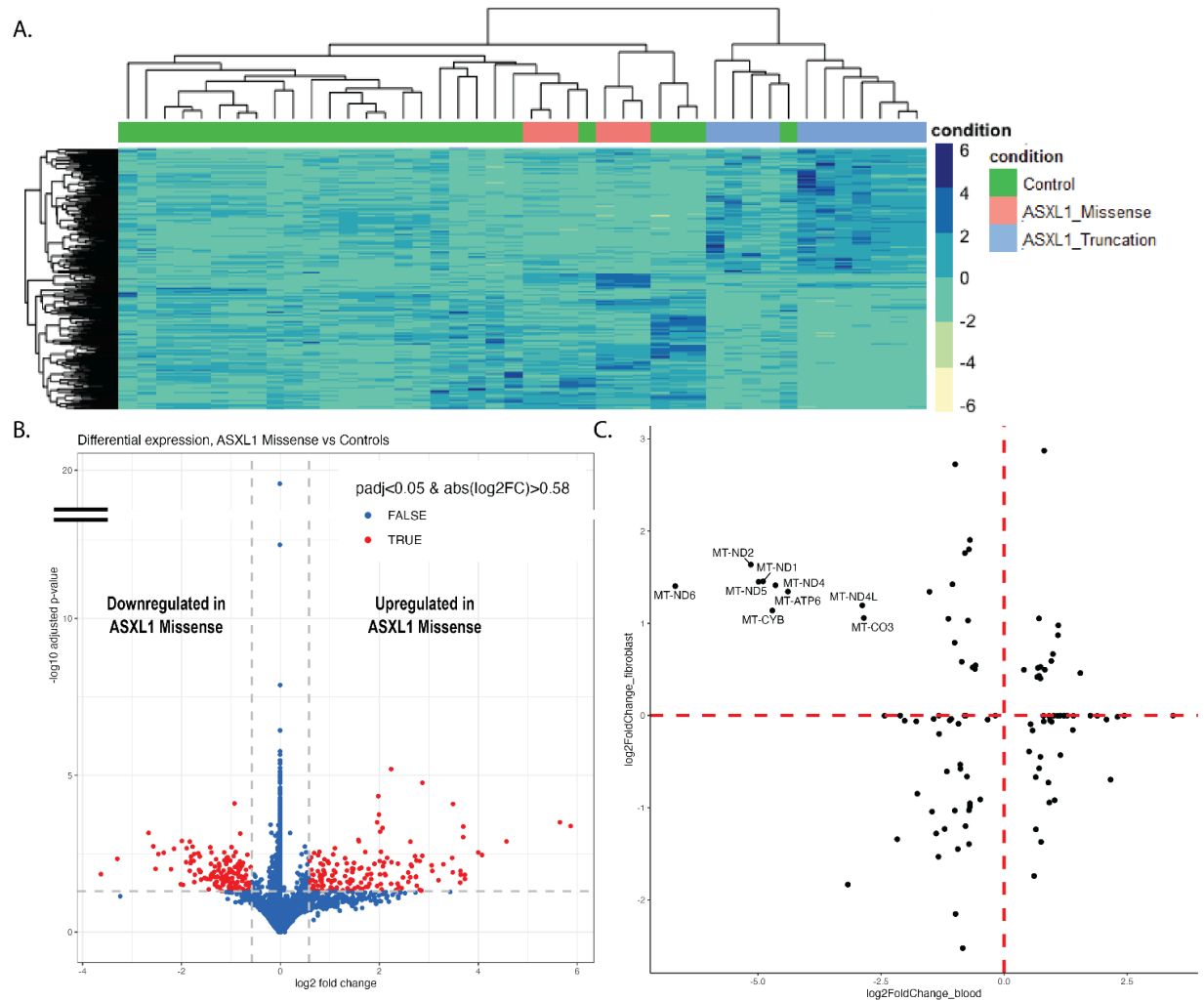


Figure 5-5: Transcriptomic profiling and differential expression analysis in *ASXL1* missense mutation identifies oxidative phosphorylation dysregulation

(A) Heatmap of patient-derived fibroblast RNA-seq for *ASXL1* missense, Bohring-Opitz syndrome (BOS, *ASXL1* truncation) and control individuals using the BOS transcriptomic signature. All technical replicates are displayed. **(B)** Volcano plot of differentially expressed genes (DEGs) from *ASXL1* missense fibroblast. RNA-seq identified 1565 DEGs, of which 1356 (86.6%) are downregulated (log2 fold change < 0), and 209 (13.4%) are upregulated (log2 fold change > 0). **(C)** Integration of DEGs identified in both blood (x-axis) and fibroblast (y-axis) RNA-seq analysis of *ASXL1* missense mutation samples pinpoints 104 common DEGs across

different tissues. The $\log_2FC$ for blood RNA-seq is shown on the x-axis and the $\log_2FC$ for fibroblast RNA-seq on the y-axis.

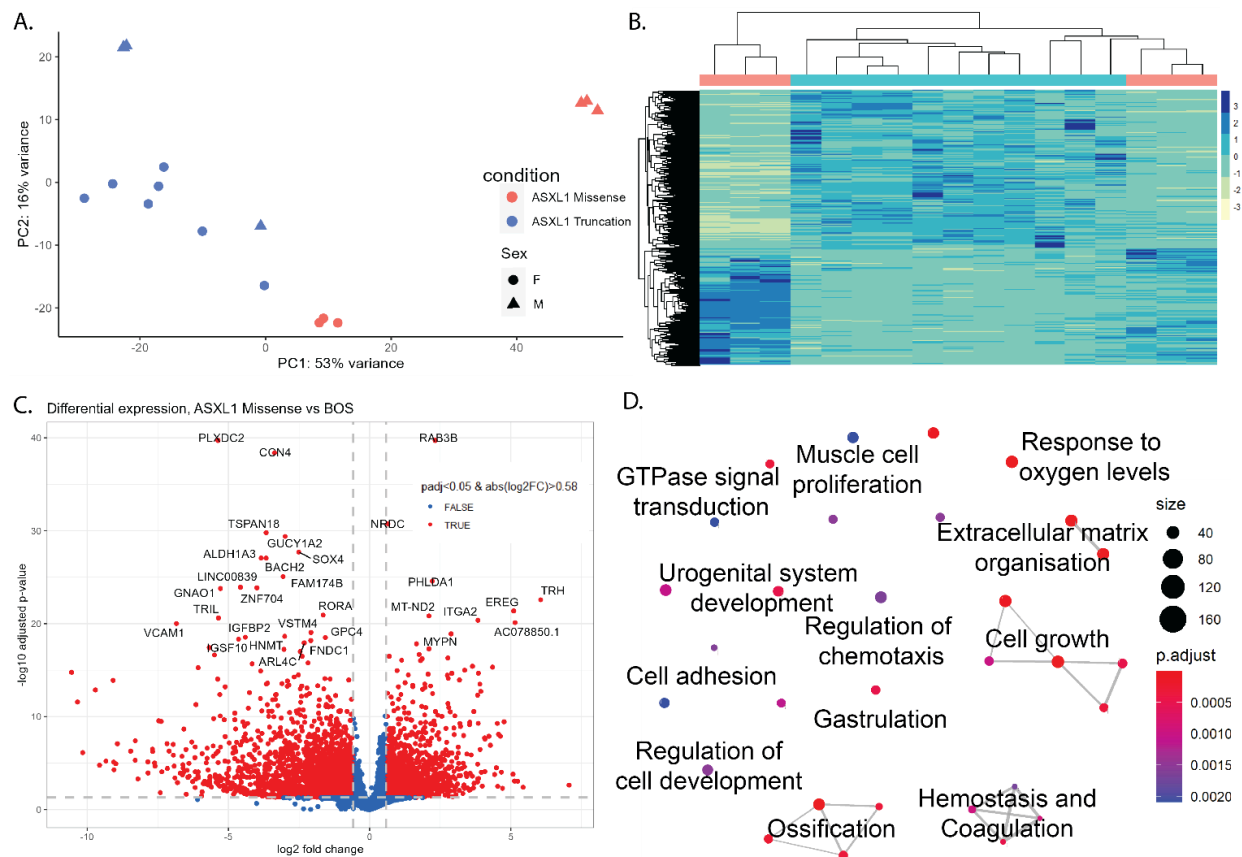


Figure 5-6: Comparative analysis of fibroblast RNA-seq data in *ASXL1* missense and truncating mutations

(A) Principal component analysis (PCA) of patient-derived fibroblast RNA sequencing (RNA-seq) data for samples from patients with *ASXL1* missense mutations compared to *ASXL1* truncating mutations which cause Bohring-Opitz syndrome (BOS). Independently generated technical replicates cluster together.

(B) Volcano plot of *ASXL1* missense patient-derived fibroblast samples compared to BOS fibroblast samples RNA-seq differentially expressed genes (DEGs), with log2 fold change on the x-axis and $-\log_{10}$ adjusted p-value on the y-axis.

(C) Gene ontology of DEGs identified from comparison of *ASXL1* missense blood to BOS blood samples identifies dysregulation of key pathways including cell growth, ossification, hemostasis, and response to oxygen levels.

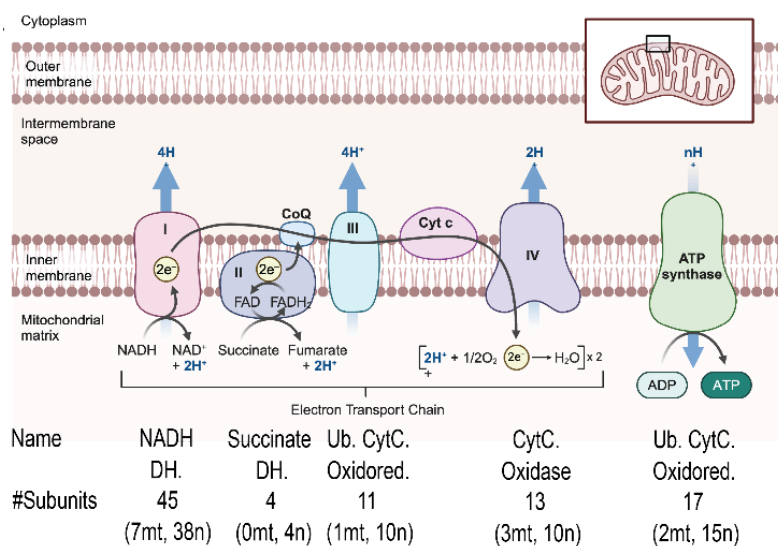


Figure 5-7: Oxidative phosphorylation (OXPHOS) complexes and subunits

(A) Schematic representation of the oxidative phosphorylation (OXPHOS) chain illustrating the complexes and key reactions. Complex names and number of subunits, subdivided into mitochondrially encoded (mt) and nuclearly encoded (n) are listed below each complex. Created with *BioRender*.

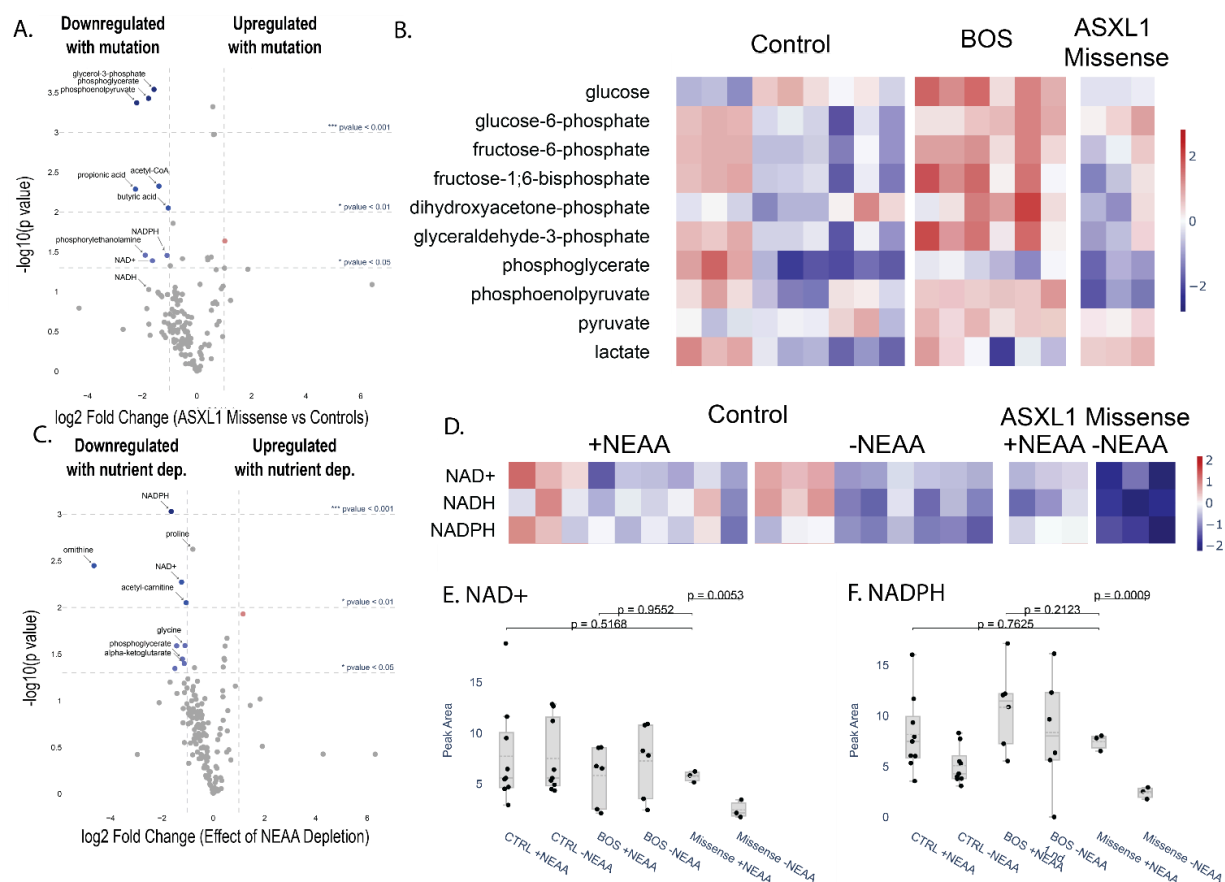


Figure 5-8: Metabolomic insights into *ASXL1* missense mutation effects and energy metabolism dysregulation

(A) Volcano plot of untargeted metabolomic analysis comparing *ASXL1* missense fibroblast and control cells identified pronounced downregulation of glycolytic intermediates, indicated a disrupted glycolytic pathway and altered energy production mechanisms in the presence of *ASXL1* missense mutations. **(B)** Heatmap of untargeted metabolomics glycolytic metabolites with control samples (n=3, technical triplicates), BOS (n=2, technical triplicates), *ASXL1* missense proband (technical triplicates).

(C) Volcano plot of untargeted metabolomic analysis illustrating the metabolic response of *ASXL1* missense mutation fibroblast to depletion of non-essential amino acids (NEAA). Key dysregulated metabolites include pronounced downregulation of energy carriers NADPH and NAD⁺, indicating vulnerability to amino acid availability and impaired metabolic flexibility. **(D)**

Heatmap of untargeted metabolomics energy carrier metabolites with control samples (n=3, technical triplicates), BOS (n=2, technical triplicates), *ASXL1* missense proband (technical triplicates).

(E) NAD⁺ metabolite levels in control, BOS, and *ASXL1* missense samples, with or without NEAA.

(F) NADPH metabolite levels in control, BOS, and *ASXL1* missense samples, with or without NEAA.

Gene ID	log2FC	pvalue	padj	Gene Name	Approved Name
ENSG00000198695.2	-6.68	1.17E-15	1.62E-12	MT-ND6	mitochondrially encoded NADH:ubiquinone oxidoreductase core subunit 6
ENSG00000286106.1	-6.18	4.79E-05	1.48E-03	AC253572.1	NA
ENSG00000166143.9	-5.41	8.37E-05	2.18E-03	PPP1R14D	protein phosphatase 1 regulatory inhibitor subunit 14D
ENSG00000231989.4	-5.36	9.79E-05	2.44E-03	PPP1R2B	PPP1R2 family member B
ENSG00000198918.8	-5.32	5.66E-12	3.74E-09	RPL39	ribosomal protein L39
ENSG00000198763.3	-5.14	4.78E-13	3.72E-10	MT-ND2	mitochondrially encoded NADH:ubiquinone oxidoreductase core subunit 2
ENSG00000198786.2	-4.99	2.72E-14	2.85E-11	MT-ND5	mitochondrially encoded NADH:ubiquinone oxidoreductase core subunit 5
ENSG00000213937.4	-4.98	4.48E-10	1.72E-07	CLDN9	claudin 9
ENSG00000198888.2	-4.89	1.49E-13	1.30E-10	MT-ND1	mitochondrially encoded NADH:ubiquinone oxidoreductase core subunit 1
ENSG00000198727.2	-4.71	4.18E-15	5.29E-12	MT-CYB	mitochondrially encoded cytochrome b
ENSG00000198886.2	-4.64	1.17E-13	1.04E-10	MT-ND4	mitochondrially encoded NADH:ubiquinone oxidoreductase core subunit 4
ENSG00000270276.2	-4.57	4.83E-04	7.69E-03	HIST2H4B	H4 clustered histone 15
ENSG00000228253.1	-4.47	2.93E-10	1.20E-07	MT-ATP8	mitochondrially encoded ATP synthase membrane subunit 8
ENSG00000163431.13	4.43	4.92E-08	7.81E-06	LMOD1	leiomodin 1
ENSG00000198899.2	-4.39	1.95E-11	1.10E-08	MT-ATP6	mitochondrially encoded ATP synthase membrane subunit 6
ENSG00000188729.6	-3.97	1.54E-03	1.72E-02	OSTN	osteocrin
ENSG00000150281.6	-3.94	4.86E-06	2.83E-04	CTF1	cardiotrophin 1
ENSG00000221882.3	-3.89	1.50E-03	1.70E-02	AC087498.2	NA
ENSG00000170315.13	3.76	1.73E-23	7.49E-20	UBB	ubiquitin B
ENSG00000239732.3	-3.71	5.83E-10	2.13E-07	TLR9	toll like receptor 9
ENSG00000136819.15	3.63	2.67E-16	4.06E-13	C9orf78	chromosome 9 open reading frame 78
ENSG00000243716.10	-3.60	4.75E-15	5.77E-12	NPIP5	nuclear pore complex interacting protein family member B5
ENSG00000258653.3	-3.55	1.52E-04	3.38E-03	AC005520.1	NA
ENSG00000169894.18	3.49	5.95E-08	8.91E-06	MUC3A	mucin 3A, cell surface associated
ENSG00000160472.5	-3.41	2.54E-03	2.40E-02	TMEM190	transmembrane protein 190
ENSG00000283378.1	3.35	2.99E-20	6.99E-17	CNTNAP3C	contactin associated protein like 3C
ENSG00000011028.14	3.28	6.54E-09	1.58E-06	AC080038.1	NA
ENSG00000076864.19	3.25	2.84E-04	5.28E-03	RAP1GAP	RAP1 GTPase activating protein
ENSG00000254788.7	3.24	8.10E-10	2.83E-07	CKLF-CMTM1	CKLF-CMTM1 readthrough
ENSG00000104892.17	3.24	7.43E-10	2.63E-07	KLC3	kinesin light chain 3

Table 5-1: Differentially expressed genes in blood samples from *ASXL1* missense mutation patients compared to controls

The 30 most strongly dysregulated protein-coding differentially expressed genes (DEGs) identified in RNA sequencing of blood samples collected from a family with an *ASXL1* missense mutation compared to controls.

GO ID	Description	GeneRatio	BgRatio	pvalue	p.adjust
GO:0008380	RNA splicing	152/2773	483/18654	7.33E-21	4.41E-17
GO:0042773	ATP synthesis coupled electron transport	52/2773	97/18654	5.59E-19	1.12E-15
GO:0006402	mRNA catabolic process	121/2773	376/18654	1.09E-17	7.28E-15
GO:0022613	ribonucleoprotein complex biogenesis	143/2773	478/18654	1.60E-17	9.62E-15
GO:0046034	ATP metabolic process	105/2773	308/18654	1.76E-17	9.62E-15
GO:0070125	mitochondrial translational elongation	46/2773	88/18654	2.18E-16	1.01E-13
GO:0071103	DNA conformation change	112/2773	351/18654	3.58E-16	1.35E-13
GO:0065004	protein-DNA complex assembly	86/2773	242/18654	8.92E-16	2.98E-13
GO:0097549	chromatin organization, negative regulation of transcription	60/2773	141/18654	1.73E-15	5.48E-13
GO:0071824	protein-DNA complex subunit organization	94/2773	281/18654	3.32E-15	9.56E-13
GO:0060968	regulation of gene silencing	58/2773	138/18654	9.91E-15	2.56E-12
GO:0033108	mitochondrial respiratory chain complex assembly	48/2773	102/18654	1.02E-14	2.56E-12
GO:0031497	chromatin assembly	72/2773	193/18654	1.13E-14	2.72E-12
GO:0045814	negative regulation of gene expression, epigenetic	54/2773	125/18654	2.05E-14	4.75E-12
GO:0140053	mitochondrial gene expression	64/2773	166/18654	5.67E-14	1.26E-11
GO:1903311	regulation of mRNA metabolic process	105/2773	344/18654	7.94E-14	1.65E-11
GO:0006333	chromatin assembly or disassembly	76/2773	218/18654	1.32E-13	2.41E-11
GO:0019083	viral transcription	66/2773	178/18654	2.01E-13	3.36E-11
GO:0045652	regulation of megakaryocyte differentiation	40/2773	83/18654	6.35E-13	9.79E-11
GO:0006403	RNA localization	78/2773	235/18654	1.18E-12	1.57E-10

Table 5-2: Gene ontology enrichment in blood samples from *ASXL1* missense mutation patients compared to controls

Top 20 gene ontology (GO) terms identified in RNA sequencing of blood samples collected from a family with an *ASXL1* missense mutation compared to controls. Dysregulated pathways include ATP synthesis coupled electron transport and ATP metabolic processes, as well as protein-DNA complex assembly and chromatin regulation of gene expression.

Gene ID	log2FC	pvalue	padj	Gene Name	Approved Name
ENSG00000170893.4	5.87	1.21E-06	4.15E-04	TRH	thyrotropin releasing hormone
ENSG00000186847.6	5.65	7.92E-07	3.14E-04	KRT14	keratin 14
ENSG00000146197.9	4.57	7.38E-06	1.28E-03	SCUBE3	signal peptide, CUB domain and EGF like domain containing 3
ENSG00000111057.11	4.08	3.43E-05	3.48E-03	KRT18	keratin 18
ENSG00000257335.8	3.73	3.11E-04	1.49E-02	MGAM	maltase-glucoamylase
ENSG00000171345.13	3.64	3.01E-04	1.47E-02	KRT19	keratin 19
ENSG00000164161.10	3.63	7.29E-04	2.64E-02	HHIP	hedgehog interacting protein
ENSG00000139973.16	3.47	4.92E-05	4.32E-03	SYT16	synaptotagmin 16
ENSG00000115008.5	3.19	5.73E-04	2.21E-02	IL1A	interleukin 1 alpha
ENSG00000122862.5	3.07	2.02E-04	1.12E-02	SRGN	serglycin
ENSG00000138207.14	2.85	1.63E-03	4.74E-02	RBP4	retinol binding protein 4
ENSG00000125538.11	2.82	1.48E-03	4.48E-02	IL1B	interleukin 1 beta
ENSG00000153162.9	2.81	7.40E-05	5.43E-03	BMP6	bone morphogenetic protein 6
ENSG00000124882.4	2.74	3.77E-04	1.70E-02	EREG	epiregulin
ENSG00000163347.6	2.74	4.06E-04	1.75E-02	CLDN1	claudin 1
ENSG00000073756.12	2.72	3.92E-05	3.78E-03	PTGS2	prostaglandin-endoperoxide synthase 2
ENSG00000125246.15	-2.66	2.63E-06	6.89E-04	CLYBL	citrate lyase beta like
ENSG00000181195.11	2.65	1.24E-04	7.89E-03	PENK	proenkephalin
ENSG00000123358.20	2.64	1.08E-04	7.10E-03	NR4A1	nuclear receptor subfamily 4 group A member 1
ENSG00000153234.14	2.63	7.89E-06	1.31E-03	NR4A2	nuclear receptor subfamily 4 group A member 2
ENSG00000120156.21	2.62	4.17E-04	1.78E-02	TEK	TEK receptor tyrosine kinase
ENSG00000134594.5	-2.47	2.94E-05	3.23E-03	RAB33A	RAB33A, member RAS oncogene family
ENSG00000168491.10	2.45	1.24E-03	3.96E-02	CCDC110	coiled-coil domain containing 110
ENSG00000164171.11	2.24	8.36E-05	5.89E-03	ITGA2	integrin subunit alpha 2
ENSG00000115129.14	2.24	2.17E-09	6.35E-06	TP53I3	tumor protein p53 inducible protein 3
ENSG00000173391.9	2.12	3.25E-04	1.54E-02	OLR1	oxidized low density lipoprotein receptor 1
ENSG00000164220.7	2.09	8.65E-05	6.06E-03	F2RL2	coagulation factor II thrombin receptor like 2
ENSG00000106484.15	2.07	1.27E-04	7.99E-03	MEST	mesoderm specific transcript
ENSG00000162772.17	2.03	2.24E-05	2.75E-03	ATF3	activating transcription factor 3
ENSG00000168497.5	2.03	8.36E-04	2.92E-02	CAVIN2	caveolae associated protein 2

Table 5-3: Differentially expressed genes in fibroblast samples from *ASXL1* missense mutation patients compared to controls

The 30 most strongly dysregulated protein-coding differentially expressed genes (DEGs) identified in RNA sequencing of patient-derived fibroblast samples collected from a family with an *ASXL1* missense mutation compared to controls.

GO ID	Description	GeneRatio	BgRatio	pvalue	p.adjust
GO:0045229	external encapsulating structure organization	104/3488	320/18669	1.83E-09	1.17E-05
GO:0032963	collagen metabolic process	44/3488	107/18669	5.94E-08	9.47E-05
GO:0016055	Wnt signaling pathway	129/3488	455/18669	2.43E-07	3.10E-04
GO:0198738	cell-cell signaling by wnt	129/3488	457/18669	3.16E-07	3.36E-04
GO:0019216	regulation of lipid metabolic process	103/3488	347/18669	3.78E-07	3.45E-04
GO:0034329	cell junction assembly	121/3488	431/18669	9.88E-07	7.33E-04
GO:0001503	ossification	120/3488	427/18669	1.03E-06	7.33E-04
GO:0007409	axonogenesis	122/3488	439/18669	1.57E-06	1.00E-03
GO:0016049	cell growth	132/3488	491/18669	3.99E-06	2.12E-03
GO:0071496	cellular response to external stimulus	94/3488	328/18669	6.17E-06	3.03E-03
GO:0045785	positive regulation of cell adhesion	128/3488	481/18669	9.55E-06	4.36E-03
GO:0070482	response to oxygen levels	95/3488	337/18669	1.16E-05	4.92E-03
GO:0062012	regulation of small molecule metabolic process	95/3488	339/18669	1.50E-05	5.09E-03
GO:0032535	regulation of cellular component size	102/3488	370/18669	1.58E-05	5.09E-03
GO:0072594	establishment of protein localization to organelle	119/3488	446/18669	1.70E-05	5.09E-03
GO:0050920	regulation of chemotaxis	69/3488	229/18669	1.74E-05	5.09E-03
GO:0050921	positive regulation of chemotaxis	48/3488	144/18669	1.90E-05	5.09E-03
GO:0009612	response to mechanical stimulus	65/3488	213/18669	1.93E-05	5.09E-03
GO:0002478	antigen processing and presentation	19/3488	39/18669	1.99E-05	5.09E-03
GO:0060389	pathway-restricted SMAD protein phosphorylation	27/3488	66/18669	2.30E-05	5.32E-03

Table 5-4: Gene ontology analysis of patient-derived fibroblasts with *ASXL1* missense mutation.

Top 20 gene ontology (GO) terms identified in RNA sequencing of fibroblasts derived from a family with an *ASXL1* missense mutation compared to controls. Dysregulated pathways include Wnt signaling pathway, axonogenesis, and metabolic processes.

Gene ID	log2FC	pvalue	padj	Gene Name	Approved Name
ENSG00000002933.9	-10.55	3.03E-18	1.67E-15	TMEM176A	transmembrane protein 176A
ENSG00000162706.13	-10.34	1.05E-14	2.68E-12	CADM3	cell adhesion molecule 3
ENSG00000134533.6	-10.15	1.42E-08	7.54E-07	RERG	RAS like estrogen regulated growth inhibitor
ENSG00000106565.18	-9.71	4.15E-16	1.35E-13	TMEM176B	transmembrane protein 176B
ENSG00000162624.14	-9.56	5.52E-07	1.67E-05	LHX8	LIM homeobox 8
ENSG00000147588.7	-9.33	1.65E-07	6.10E-06	PMP2	peripheral myelin protein 2
ENSG00000025039.14	-9.08	2.85E-17	1.24E-14	RRAGD	Ras related GTP binding D
ENSG00000187140.5	-9.06	2.16E-07	7.63E-06	FOXD3	forkhead box D3
ENSG00000198650.11	-9.05	2.19E-06	5.24E-05	TAT	tyrosine aminotransferase
ENSG00000217236.2	-8.99	3.71E-07	1.19E-05	SP9	Sp9 transcription factor
ENSG00000116329.11	-8.93	1.12E-09	8.27E-08	OPRD1	opioid receptor delta 1
ENSG00000206579.8	-8.46	9.20E-09	5.18E-07	XKR4	XK related 4
ENSG00000151834.15	-8.30	1.32E-07	5.03E-06	GABRA2	gamma-aminobutyric acid type A receptor alpha2 subunit
ENSG00000100146.17	-8.30	4.93E-10	3.99E-08	SOX10	SRY-box transcription factor 10
ENSG00000145708.11	-8.20	3.32E-05	0.005074	CRHBP	corticotropin releasing hormone binding protein
ENSG00000180828.2	-7.85	2.06E-07	7.37E-06	BHLHE22	basic helix-loop-helix family member e22
ENSG00000155886.11	-7.84	5.58E-06	0.0001149	SLC24A2	solute carrier family 24 member 2
ENSG00000054938.15	-7.68	1.02E-08	5.63E-07	CHRD12	chordin like 2
ENSG00000111262.6	-7.61	2.86E-07	9.61E-06	KCNA1	potassium voltage-gated channel subfamily A member 1
ENSG00000000005.6	-7.60	6.99E-07	2.05E-05	TNMD	tenomodulin
ENSG00000150275.18	-7.47	1.81E-12	2.84E-10	PCDH15	protocadherin related 15
ENSG00000204655.12	-7.41	0.0023010	0.0150384	MOG	myelin oligodendrocyte glycoprotein
ENSG00000277893.2	-7.38	2.15E-12	3.33E-10	SRD5A2	steroid 5 alpha-reductase 2
ENSG00000145536.15	-7.36	3.98E-06	8.65E-05	ADAMTS16	ADAM metalloproteinase with thrombospondin type 1 motif 16
ENSG00000147257.14	-7.17	1.10E-09	8.15E-08	GPC3	glypican 3
ENSG00000034971.17	-7.10	9.11E-09	5.14E-07	MYOC	myocilin
ENSG00000077274.9	-6.95	1.34E-09	9.82E-08	CAPN6	calpain 6
ENSG00000133937.4	-6.95	2.95E-08	1.39E-06	GSC	goosecoid homeobox
ENSG00000172000.7	-6.92	7.36E-07	2.12E-05	ZNF556	zinc finger protein 556
ENSG00000162692.12	-6.84	7.50E-24	9.76E-21	VCAM1	vascular cell adhesion molecule 1

Table 5-5: Differentially expressed genes in blood samples from *ASXL1* missense mutation patients compared to *ASXL1* truncation patients.

The 30 most strongly dysregulated protein-coding differentially expressed genes (DEGs) identified in RNA sequencing of blood samples collected from a family with an *ASXL1* missense mutation compared to Bohring-Opitz syndrome (BOS) individuals with *ASXL1* truncation mutations.

GO ID	Description	GeneRatio	BgRatio	pvalue	p.adjust
GO:0043062	extracellular structure organization	155/4310	396/18654	4.03E-13	2.26E-09
GO:0030198	extracellular matrix organization	154/4310	395/18654	6.97E-13	2.26E-09
GO:0034329	cell junction assembly	161/4310	434/18654	2.28E-11	4.91E-08
GO:0001503	ossification	148/4310	410/18654	1.33E-09	2.15E-06
GO:0016049	cell growth	169/4310	488/18654	3.26E-09	4.21E-06
GO:0007409	axonogenesis	163/4310	482/18654	3.94E-08	4.24E-05
GO:0070482	response to oxygen levels	138/4310	395/18654	4.88E-08	4.51E-05
GO:0060348	bone development	79/4310	203/18654	2.92E-07	2.36E-04
GO:0007369	gastrulation	74/4310	189/18654	5.19E-07	3.38E-04
GO:0030282	bone mineralization	51/4310	116/18654	5.22E-07	3.38E-04
GO:0030308	negative regulation of cell growth	74/4310	190/18654	6.59E-07	3.69E-04
GO:0030278	regulation of ossification	80/4310	210/18654	6.85E-07	3.69E-04
GO:0001655	urogenital system development	114/4310	330/18654	1.31E-06	5.04E-04
GO:0007160	cell-matrix adhesion	85/4310	230/18654	1.37E-06	5.04E-04
GO:0045926	negative regulation of growth	92/4310	254/18654	1.41E-06	5.04E-04
GO:1900046	regulation of hemostasis	39/4310	83/18654	1.48E-06	5.04E-04
GO:0030193	regulation of blood coagulation	38/4310	82/18654	3.06E-06	9.00E-04
GO:0061041	regulation of wound healing	60/4310	151/18654	3.42E-06	9.61E-04
GO:0060560	developmental growth involved in morphogenesis	85/4310	235/18654	3.72E-06	1.00E-03
GO:0045785	positive regulation of cell adhesion	139/4310	427/18654	4.10E-06	1.06E-03

Table 5-6: Gene ontology analysis of patient-derived fibroblasts with *ASXL1* missense mutation compared to *ASXL1* truncation mutations.

Top 20 gene ontology (GO) terms identified in RNA sequencing of fibroblasts derived from a family with an *ASXL1* missense mutation compared to Bohring-Opitz syndrome (BOS) patients, characterized by *ASXL1* truncation mutations. Dysregulated pathways include developmental growth, response to oxygen levels, and hemostasis.

Gene ID	Gene Name	Approved Name	log2FC_blood	padj_blood	log2FC_fibro	padj_fibro
ENSG00000198695.2	MT-ND6	mitochondrially encoded NADH:ubiquinone oxidoreductase core subunit 6	-6.68	1.62E-12	1.40	0.03
ENSG00000198727.2	MT-CYB	mitochondrially encoded cytochrome b	-4.71	5.29E-12	1.14	0.04
ENSG00000198763.3	MT-ND2	mitochondrially encoded NADH:ubiquinone oxidoreductase core subunit 2	-5.14	3.72E-10	1.64	0.01
ENSG00000198786.2	MT-ND5	mitochondrially encoded NADH:ubiquinone oxidoreductase core subunit 5	-4.99	2.85E-11	1.45	0.01
ENSG00000198804.2	MT-CO1	mitochondrially encoded cytochrome c oxidase I	-1.13	1.00E-02	1.05	0.02
ENSG00000198886.2	MT-ND4	mitochondrially encoded NADH:ubiquinone oxidoreductase core subunit 4	-4.64	1.04E-10	1.41	0.02
ENSG00000198888.2	MT-ND1	mitochondrially encoded NADH:ubiquinone oxidoreductase core subunit 1	-4.89	1.30E-10	1.46	0.01
ENSG00000198899.2	MT-ATP6	mitochondrially encoded ATP synthase membrane subunit 6	-4.39	1.10E-08	1.34	0.01
ENSG00000198938.2	MT-CO3	mitochondrially encoded cytochrome c oxidase III	-2.85	5.68E-08	1.06	0.04
ENSG00000205138.3	SDHAF1	succinate dehydrogenase complex assembly factor 1	1.03	6.08E-03	-0.92	0.01
ENSG00000212907.2	MT-ND4L	mitochondrially encoded NADH:ubiquinone oxidoreductase core subunit 4L	-2.88	5.32E-07	1.20	0.03

Table 5-7: Differential expression of significantly dysregulated genes involved in oxidative phosphorylation across tissues in *ASXL1* missense samples.

This table lists the log2 fold changes and adjusted p-values (padj, < 0.05) for 11 genes related to the oxidative phosphorylation chain. These genes were significantly dysregulated in both blood and fibroblast samples from a family with an *ASXL1* missense mutation, compared to control samples.

KEGG Pathway	Pathway total	Hits.total	Hits.sig	Expected	P(Fisher)
beta-Alanine metabolism	21	4	3	1.03	0.01
Pentose and glucuronate interconversions	19	9	7	1.20	0.02
Glycolysis / Gluconeogenesis	26	11	7	1.88	0.03
Fructose and mannose metabolism	20	14	9	1.88	0.03
Glycerolipid metabolism	16	2	2	0.34	0.03
Terpenoid backbone biosynthesis	18	4	3	0.86	0.04
Glycine, serine and threonine metabolism	33	4	4	1.71	0.07
Caffeine metabolism	10	3	3	0.51	0.08
N-Glycan biosynthesis	41	2	1	0.51	0.08
Nitrogen metabolism	6	2	2	0.51	0.08
Pyruvate metabolism	23	6	3	1.20	0.10
Drug metabolism - cytochrome P450	55	17	6	3.60	0.13
Propanoate metabolism	22	2	1	0.69	0.14
Pentose phosphate pathway	23	7	4	1.37	0.14
D-Amino acid metabolism	15	8	4	1.37	0.14

Table 5-8: Enriched metabolic pathways identified in fibroblasts with *ASXL1* missense mutation.

Top 15 enriched metabolic pathways identified in untargeted metabolomics analysis using Mummichog from a family with an *ASXL1* missense mutation compared to controls. Dysregulated pathways include glycolysis / gluconeogenesis, and glycine metabolism.

References

1. Mills, A. A. Throwing the cancer switch: reciprocal roles of polycomb and trithorax proteins. *Nat. Rev. Cancer* **10**, 669–682 (2010).
2. Poynter, S. T. & Kadoch, C. Polycomb and trithorax opposition in development and disease. *Wiley Interdiscip. Rev. Dev. Biol.* **5**, 659–688 (2016).
3. Bohring, A. *et al.* Severe end of Opitz trigonocephaly (C) syndrome or new syndrome? *Am. J. Med. Genet.* **85**, 438–446 (1999).
4. Russell, B. E. *et al.* Clinical findings in 39 individuals with Bohring-Opitz syndrome from a global patient-driven registry with implications for tumor surveillance and recurrence risk. *Am. J. Med. Genet. A* **191**, 1050–1058 (2023).
5. Katoh, M. Functional and cancer genomics of ASXL family members. *Br. J. Cancer* **109**, 299–306 (2013).
6. Abdel-Wahab, O. *et al.* ASXL1 mutations promote myeloid transformation through loss of PRC2-mediated gene repression. *Cancer Cell* **22**, 180–193 (2012).
7. Lai, H.-L. & Wang, Q. T. Additional sex combs-like 2 is required for polycomb repressive complex 2 binding at select targets. *PLoS One* **8**, e73983 (2013).
8. Balasubramani, A. *et al.* Cancer-associated ASXL1 mutations may act as gain-of-function mutations of the ASXL1–BAP1 complex. *Nat. Commun.* **6**, 1–15 (2015).
9. Asada, S. *et al.* Mutant ASXL1 cooperates with BAP1 to promote myeloid leukaemogenesis. *Nature Communications* vol. 9 Preprint at <https://doi.org/10.1038/s41467-018-05085-9> (2018).
10. Lin, I. *et al.* Multi-omics of Bohring-Opitz syndrome truncating ASXL1 mutations identify canonical and non-canonical Wnt signaling dysregulation. *JCI Insight* (2023) doi:10.1172/jci.insight.167744.
11. Yang, H. *et al.* Gain of function of ASXL1 truncating protein in the pathogenesis of myeloid

- malignancies. *Blood* **131**, 328–341 (2018).
12. Balasubramanian, M. *et al.* Delineating the phenotypic spectrum of Bainbridge-Ropers syndrome: 12 new patients with de novo, heterozygous, loss-of-function mutations in ASXL3 and review of published literature. *J. Med. Genet.* **54**, 537–543 (2017).
 13. Giri, D. *et al.* Novel compound heterozygous ASXL3 mutation causing Bainbridge-ropers like syndrome and primary IGF1 deficiency. *Int. J. Pediatr. Endocrinol.* **2017**, 8 (2017).
 14. Krall, A. S. *et al.* Asparagine couples mitochondrial respiration to ATF4 activity and tumor growth. *Cell Metab.* **33**, 1013–1026.e6 (2021).
 15. Brandes, N., Goldman, G., Wang, C. H., Ye, C. J. & Ntranos, V. Genome-wide prediction of disease variant effects with a deep protein language model. *Nat. Genet.* **55**, 1512–1522 (2023).
 16. Cheng, J. *et al.* Accurate proteome-wide missense variant effect prediction with AlphaMissense. *Science* **381**, eadg7492 (2023).
 17. Zhang, P. *et al.* Loss of ASXL1 in the bone marrow niche dysregulates hematopoietic stem and progenitor cell fates. *Cell Discov* **4**, 4 (2018).
 18. Sancak, Y. *et al.* Ragulator-Rag complex targets mTORC1 to the lysosomal surface and is necessary for its activation by amino acids. *Cell* **141**, 290–303 (2010).
 19. Tsun, Z.-Y. *et al.* The folliculin tumor suppressor is a GAP for the RagC/D GTPases that signal amino acid levels to mTORC1. *Mol. Cell* **52**, 495–505 (2013).
 20. Schlingmann, K. P. *et al.* mTOR-Activating Mutations in RRAGD Are Causative for Kidney Tubulopathy and Cardiomyopathy. *J. Am. Soc. Nephrol.* **32**, 2885–2899 (2021).
 21. Sousa, J. S., D’Imprima, E. & Vonck, J. Mitochondrial Respiratory Chain Complexes. *Subcell. Biochem.* **87**, 167–227 (2018).
 22. Sharma, L. K., Lu, J. & Bai, Y. Mitochondrial respiratory complex I: structure, function and implication in human diseases. *Curr. Med. Chem.* **16**, 1266–1277 (2009).
 23. Kühlbrandt, W. Structure and function of mitochondrial membrane protein complexes. *BMC*

- Biol.* **13**, 89 (2015).
24. Bohring, A., Oudesluijs, G. G., Grange, D. K., Zampino, G. & Thierry, P. New cases of Bohring-Opitz syndrome, update, and critical review of the literature. *Am. J. Med. Genet. A* **140**, 1257–1263 (2006).
 25. Li, Y. Y. *et al.* Exome and genome sequencing of nasopharynx cancer identifies NF- κ B pathway activating mutations. *Nat. Commun.* **8**, 14121 (2017).
 26. Gong, X.-Y. *et al.* NOTCH1 mutation associates with impaired immune response and decreased relapse-free survival in patients with resected T1-2N0 laryngeal cancer. *Front. Immunol.* **13**, 920253 (2022).
 27. Liberti, M. V. & Locasale, J. W. The Warburg Effect: How Does it Benefit Cancer Cells? *Trends Biochem. Sci.* **41**, 211–218 (2016).
 28. Efthymiou, S. *et al.* A de novo truncating mutation in ASXL1 associated with segmental overgrowth. *J. Genet.* **98**, (2019).
 29. Schwartz Marianne & Vissing John. Paternal Inheritance of Mitochondrial DNA. *N. Engl. J. Med.* **347**, 576–580.
 30. Luo, S. *et al.* Biparental Inheritance of Mitochondrial DNA in Humans. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 13039–13044 (2018).
 31. Wei, W. & Chinnery, P. F. Inheritance of mitochondrial DNA in humans: implications for rare and common diseases. *J. Intern. Med.* **287**, 634–644 (2020).
 32. Fato, R., Bergamini, C., Leoni, S., Strocchi, P. & Lenaz, G. Generation of reactive oxygen species by mitochondrial complex I: implications in neurodegeneration. *Neurochem. Res.* **33**, 2487–2501 (2008).
 33. Perez-Gomez, R., Magnin, V., Mihajlovic, Z., Slaninova, V. & Krejci, A. Downregulation of respiratory complex I mediates major signalling changes triggered by TOR activation. *Sci. Rep.* **10**, 4401 (2020).

CHAPTER 6

Comparative Analysis of *ASXL1* Mutations in Bohring-Opitz syndrome and Acute Myeloid Leukemia

Abstract

This study delves into the shared epigenetic and transcriptomic impacts of additional sex combs-like 1 (*ASXL1*) mutations in Bohring-Opitz syndrome (BOS) and acute myeloid leukemia (AML) patients with *ASXL1* mutations (AML-*ASXL1*). *ASXL1* mutations result in two distinct disorders: AML, arising from somatic mutations, and BOS (OMIM #605039), a rare neurodevelopmental disorder caused by germline mutations, marked by multiorgan anomalies and severe intellectual disability.

Through the analysis of DNA methylation and RNA sequencing data from BOS, AML-*ASXL1*, and their respective controls, we identified common molecular signatures between these two distinct diseases. Our findings highlight key biomarkers including *VANGL2*, *GRIK5* and *GREM2* that are dysregulated across samples with *ASXL1* mutations, regardless of disease type. Notably, our data also reveal significant depression of posterior homeobox A (*HOXA*) genes and upregulation of Wnt signaling and hematopoietic regulators such as *HOXB4*, which may drive pathogenesis.

The study of differential exon usage (DEU) and isoform expression through DEXSeq analysis also revealed commonly dysregulated genes. Notably, we identified DEU of multiple exon bins for *RUNX3* in both datasets, although the dominant isoform used in BOS was *RUNX3/p46* transcript and in AML-*ASXL1* was *RUNX3/p44*. *RUNX3* encodes a transcription factor that regulates cellular development and myeloid and B-cell differentiation, and acts as a tumor suppressor in various cancers.

This study is the first to directly compare the effects of *ASXL1* mutations in AML and BOS, unveiling shared molecular pathways and suggesting potential targets for therapeutic intervention. While patterns of gene expression and epigenetic regulation may align, we find differences in isoform usage, which may inform pathophysiology of truncating *ASXL1* mutations. These findings demonstrate the significant, cross-disease influence of *ASXL1* mutations, and advocate for a gene-centric approach in the research and treatment of genetic disorders.

Introduction

Bohring-Opitz syndrome (BOS, OMIM#605309) ¹ and acute myeloid leukemia (AML) represent two conditions with markedly distinct clinical presentations and outcomes. BOS, a rare neurodevelopmental disorder characterized by severe intellectual disability, distinctive facial features, and a spectrum of physical anomalies affecting multiple organ systems, is primarily caused by germline mutations in the additional sex combs-like 1 (*ASXL1*) gene ^{2,3}. Notably, individuals with BOS exhibit an elevated risk of developing Wilms tumor, an embryonic kidney tumor that occurs in children, yet there are no reported cases of myelodysplastic syndrome or myeloid leukemias, even in the third decade of life ^{2,3}.

In contrast, AML—a cancer that affects the blood and bone marrow—exhibits a diverse genetic landscape, including somatic mutations in the *ASXL1* gene. Somatic *ASXL1* mutations are also implicated in a range of hematologic malignancies, including chronic myelomonocytic leukemia (CMML), myelodysplastic syndromes (MDS), myeloproliferative neoplasms, and both secondary and *de novo* AML ^{4–7}. Further studies have correlated low-frequency somatic mutations in *ASXL1* with mutagenic processes and increasing age, suggesting a pre-malignancy state termed clonal hematopoiesis of indeterminate potential (CHIP) ^{8,9}. The presence of the *ASXL1* mutations is a strong adverse prognostic factor in an AML, being associated with poor overall survival and therapeutic outcome ⁴. Despite the similarity in the disease pathology of BOS and AML-*ASXL1*, with some shared mutations observed in both disorders, the extent to which molecular mechanisms of dysregulation may be shared are unclear.

Despite the significant overlap between genes causing syndromic developmental delay and established cancer driver genes ¹⁰, few studies have established the presence of common, molecular mechanisms and pathway of these mutations across disease-types in patient samples. The hypothesis connecting development and cancer as an "error of development,"

proposed by Dr. Beatrice Mintz over fifty years ago ^{11,12}, posits that genetic anomalies in stem cells can lead to cancer by inducing a reversion to an undifferentiated state ¹³. Genes that regulate the epigenome, termed epigenes, play a pivotal role in controlling cell-specific RNA expression and complex biological signaling during human development ¹⁴. Mutations in epigenes serve as primary drivers in cancer, activating aberrant developmental programs that transform cells back to a malignant stem-cell state ¹⁵. Although the dual role of epigenes in both human development and cancer has been documented across numerous studies ¹⁶, the specific pathogenic mechanisms driven by lesions in the same gene across distinct diseases remain largely unexplored.

ASXL1 is a gene integral to chromatin remodeling and gene repression through polycomb-mediated mechanisms, with mutations in *ASXL1* implicated in BOS, AML, and a subset of other malignancies, highlighting its broad impact on human health ^{17,18}. It is known to regulate the balance between the proliferation and differentiation of stem progenitor cell populations, with mutations in *ASXL1* disrupting this equilibrium and favoring stem-cell identity over differentiation in both BOS and myeloid leukemias ^{19,20}.

Changes to chromatin accessibility and structure can affect alternative splicing events, which is a biological mechanism that follows mRNA transcription where non-coding introns are removed and coding exons are kept ²¹. Different combinations of exons may be included in the final mRNA transcript, resulting in multiple mRNA variants and therefore various protein isoforms, all from a single gene. Isoforms differ in their protein-coding sequences, and therefore in their protein structure and regulatory properties. Alternative splicing is a method by which cells regulate development. For example, runt-related transcription factor 3 (*RUNX3*) is a gene that has a role in the hematopoietic system and helps regulate cell differentiation, growth and tissue development.

The epigenetic and transcriptomic dysregulations driven by *ASXL1* mutations are cell-type specific, however across multiple cell types, key intermediates of Wnt-signaling

pathway intermediates are disrupted¹⁸. Some of the genetic mutations in *ASXL1* are shared between BOS and AML-*ASXL1*, and therefore we asked whether the same molecular and cellular pathways are perturbed in both distinct disorders. If these mutations drive common pathways, this highlights the potential for leveraging similar therapeutic targets against these distinct diseases.

Gene-centric approach informing molecular mechanisms

This study delves into this underexplored area of genetic research and proposes that a deeper understanding of a gene's role could reveal more about the disease mechanism than the clinical phenotype alone. A gene-specific investigative lens offers the potential to not only improve outcomes for patients with BOS and AML-*ASXL1* but also to establish a replicable model for examining other genetic disorders. This type of approach opens the door to identifying shared pathways, ultimately leading to improved diagnosis, treatment, and prevention strategies for affected individuals.

Integrative analyses of RNA-seq and DNA methylation data are important for a more comprehensive understanding of gene regulation^{22–24}. This helps pinpoint key molecular events driven by *ASXL1* mutations and how they drive distinct clinical manifestations. As an example, our recent study across distinct patient-derived tissue types in the BOS patient cohort identified *ASXL1* mutation-driven molecular dysregulations, such as the upregulation of *VANGL2*¹⁸. With this study, we aim to identify whether these findings are generalisable across distinct diseases too. The association between truncating *ASXL1* mutations and poorer prognosis in diseases could be directly linked to the disruption of critical developmental and cellular pathways such as the Wnt pathway, which was identified to be strongly upregulated in BOS samples¹⁸. Furthermore, differential exon usage and, by proxy, isoform usage of genes that play a key role in development and differentiation, such as *RUNX3*, may lend insight into the pathophysiology

of these disorders. These types of insights could have profound implications for disease severity, progression, and patient outcomes, and inform development of targeted therapies.

This study delves into the molecular landscape of *ASXL1* mutations driving two distinct diseases - BOS and AML. By employing a comprehensive approach integrating RNA-seq, DNA methylation, and exon usage analyses, we aim to unravel common epigenetic and transcriptomic effects driven by protein truncating mutations in *ASXL1*.

Methods

Selection and characterization of BOS and AML samples

Our study comprised patient cohorts from Bohring-Opitz syndrome (BOS) and Acute Myeloid Leukemia (AML). Specifically, we leverage samples from patients with AML harboring *ASXL1* mutations (AML-*ASXL1*, **Table 1**). Informed consent was obtained from all research participants according to the protocol approved by the Hospital for Sick Children (REB#1000038847) and UCLA (IRB#11-001087). Illumina 450K DNAm data was acquired for six AML samples from The Cancer Genome Atlas (TCGA) program²⁵, available on the Genomic Data Commons (GDC) repository²⁶. Among these, three samples had somatic *ASXL1* mutations, serving as the AML-*ASXL1* cohort, and three had somatic mutations in other genes, serving as AML controls. Transcriptomic data for AML-*ASXL1* bone marrow samples (n=26) were sourced from the Beat AML cohort²⁷ and AML-*ASXL1* blood samples (n=6) from TCGA, with non-AML blood controls (n=60) from the Genotype-Tissue Expression (GTEx) Portal, and bone marrow controls (n=8) from publicly available dataset GSE120444²⁸ (**Table 1 and Table 2**). Additionally, blood samples for RNA sequencing (RNAseq) and DNA methylation (DNAm) data from BOS patients (n=8 RNAseq, n=8 DNAm) and healthy blood controls from REACH biobank (n=11 RNAseq, n=26 DNAm) were collected from our previous studies and are publicly available datasets deposited at GSE230685 and GSE230686^{18,29}. A subset of patient-derived

blood (n=8 BOS, n=10 healthy controls, n=4 AML-ASXL1, n=6 AML controls) RNA-seq data was used to conduct differential exon usage (DEU) and isoform expression analysis.

RNA-seq and DNA methylation profiling of BOS and AML samples:

Sample collection, processing, and sequencing:

For DNA and RNA extraction, peripheral blood samples were processed using standardized protocols ^{18,29}, with EDTA tubes for DNA and PAXgene Blood Tubes (BDBiosciences, 762165) for RNA. DNAm profiles were generated following established methods ²⁹. RNAseq libraries were prepared using TruSeq Stranded TotalRNA Library Prep Gold (Illumina, #20020599) with QiaSelect rRNA and globin depletion (Qiagen, #334376 and #334386) following established methods ¹⁸. Pooled libraries were sequenced to 40 million reads per sample on a NovaSeq6000.

Preprocessing and quality control of RNA-seq and DNA methylation data:

RNA-seq data was processed using our established pipeline ¹⁸. Briefly, reads were mapped to hg38 using *STAR 2.7.0e* ³⁰, gene counts were generated using *featureCounts 1.6.5* ³¹, and differential expression adjusted p value (padj) and log2 fold change (log2FC) were quantified using *DESeq2 v1.24.0* ³², correcting for sex and tissue (**Table 1 and Table 2**).

DNAm data underwent processing through our established pipeline ²⁹. Briefly, highly differentially methylated sites ($|\Delta\beta| > 5\%$) were filtered, and significant CpG sites were identified with FDR < 0.05. Principal component analysis (PCA) and hierarchical clustering utilized the BOS DNAm episignature sites (413 CpG sites) ²⁹.

Integration of RNA-seq and DNA methylation data:

Integration of RNA-seq and DNAm data utilized beta values for DNAm ³³ and transcript per million (TPM) values for RNA-seq ³⁴, enabling the identification of shared and distinct molecular signatures across the conditions studied.

DEXSeq analysis of bulk RNAseq datasets:

The Bioconductor R package *DEXSeq* v1.50.0³⁵ was used to quantify DEU from RNAseq data. The reads mapping to a single exon bin were normalized against those mapping to all exon bins within the same gene, where an exon bin is a whole exon or part of an exon that arises when an exonic region occurs in different transcripts with varying boundaries. These were then compared across conditions to test for DEU. Preprocessing was done using two Python scripts built into the *DEXSeq* package, and the Python package *HTSeq* is required to access these scripts. *Samtools*/1.20 was used to convert BAM files to SAM files. P-values were adjusted for multiple testing by *DEXSeq* using the Benjamini-Hochberg method. Significant exon bins with DEU were identified with $\text{padj} < 0.05$.

Cell type deconvolution using CIBERSORTx:

CIBERSORTx is an online bioinformatics tool that assesses cell type-specific gene expression profiles and cellular composition from single cell RNA (sc-RNA) and bulk RNA-seq data³⁶. The software uses sc-RNA-seq or bulk sorted RNA-seq, or microarrays to create a deconvolution signature matrix containing marker gene profiles. The signature matrix is then used to impute cell type proportions and cell expression profiles, the latter of which has 2 modes: group-mode (gives a transcriptomic profile) and high-resolution mode (gives a sample-level profile)³⁷.

Results

Study design and genetic landscape of ASXL1 mutations

To compare the epigenetic and transcriptomics effects of ASXL1 mutations, we collected data from patients with BOS and AML with ASXL1 mutations. For BOS patients and matched controls, our integrative workflow collects patient blood samples and performs RNA-sequencing (RNA-seq) and DNA methylation (DNAm) analyses (**Figure 6-1A**). For comparison across disease-states, we identified RNA-seq and DNAm data from patients diagnosed with AML

harboring *ASXL1* mutations (AML-*ASXL1*) and controls from through dbGAP repositories ²⁷. We re-processed all samples through a shared RNA-seq and DNAm pipeline to minimize analytical batch effects (**Methods**).

The *ASXL1* gene is made up of 13 exons and encodes a protein that spans 1541 amino acids and the majority of pathogenic mutations in this study disrupt the latter half of the protein encoding region encoded by the last two exons of *ASXL1*. A diagrammatic representation of the *ASXL1* gene (**Figure 6-1B**), illustrates the BOS germline mutations (depicted below the gene) and somatic mutations driving AML-*ASXL1* in our study (depicted above the gene). Common mutation sites between the two disorders are highlighted in bold. Alternate allele read, allele frequency and mutation details for the AML-*ASXL1* bone marrow samples in this study are provided (**Table 6-3**).

Transcriptomic alterations driven by *ASXL1* mutations across blood and bone marrow

We next asked whether the pathogenic *ASXL1* mutations were the primary source of differential gene expression across disease types. Differential gene expression analysis was conducted for BOS blood (n=8), AML-*ASXL1* bone marrow (n=26), and AML-*ASXL1* blood (n=6) and controls. Since our samples spanned blood or bone marrow tissue, which are related by distinct tissue subsets, we also incorporated control data sets. These included healthy blood controls generated at UCLA (n=11), GTEX blood controls (n=60), and bone marrow controls (n=8).

To assess the role of tissue specificity on our gene expression, we processed the AML-*ASXL1* data with the blood and bone marrow combined as well as independently with the matched controls. Our principal components analysis showed that tissue type explains 17% of the variance along PC1, while *ASXL1* explains 11% along PC2 (**Figure 6-2A**). In the subsets analyzed within the same tissue types, *ASXL1* mutation status drove 34% variance in blood and 64% variance in bone marrow samples.

DESeq2 analysis was conducted to identify differentially expressed genes (DEGs) and these were visualized using unsupervised clustering (**Figure 6-2B**). This showed stronger clustering of samples by condition (AML-ASXL1) than by tissue type (blood and bone marrow). *DESeq2* analysis identified 9527 DEGs, adjusting for tissue type. Of these DEGs, 5696/9527 were upregulated (59.79%, $\log_2FC < 0$) and 3831/9527 were downregulated (40.21%, $\log_2FC > 0$) in the AML-ASXL1 samples (**Figure 6-2C**). 9140/9527 (95.94%) of these DEGs were protein coding. The 30 DEGs with the largest absolute $\log_2FC$ are shown in **Table 6-4**.

Using *clusterProfiler*, we conducted gene ontology analyses to identify dysregulated biological processes. We identified consistent dysregulation of epigenetic functions (p_{adj} 6.51E-06), including histone modification (p_{adj} 1.40E-05), chromatin modification (p_{adj} 2.35E-05), DNA replication (p_{adj} 1.30E-05) and conformational change (p_{adj} 7.63E-05), and immune activation (p_{adj} 1.30E-03) (**Figure 6-2D**).

Transcriptomic alterations driven by ASXL1 mutations in AML-ASXL1 and BOS highlight Wnt-signaling pathway

To identify DEGs that are common across different disease-types driven by *ASXL1* mutations, we integrated these AML-ASXL1 DEGs (AML-ASXL1 compared to tissue-matched controls) with the 2118 significant DEGs identified for BOS blood samples¹⁸ (BOS compared to healthy controls) to determine whether *ASXL1* mutations, regardless of whether the mutation occurred in the germline or somatic state, drove common transcriptional changes.

We identified 843 common DEGs between the AML-ASXL1 and BOS datasets, of which 566/843 (67.14%) DEGs were dysregulated in the same direction - either upregulated in both datasets, or downregulated in both datasets (**Figure 6-3A**). We plotted fold change for the transcriptomic analyses with AML-ASXL1 compared to their respective controls on the x-axis, and BOS compared to their respective controls on the y-axis. This highlighted 388/843 (46.03%) DEGs with a large effect size, with an absolute fold change greater than 1.5 in both BOS and

AML-ASXL1. Of these DEGs, 246/388 (63.40%) were dysregulated in the same direction. The 50 DEGs with the largest effect size in BOS are listed in **Table 6-5**, with respective log2FC and padj values.

Gene ontology analyses demonstrated enrichment in biological processes such as epigenetic gene regulation, histone modification, and chromatin regulation, underscoring the shared pathogenic pathways linked to the chromatin-modifying function of ASXL1 (**Figure 6-3B**, **Table 6-6**). This suggests that these dysregulated genes contribute to the disease mechanism in both BOS and AML-ASXL1 through similar molecular pathways.

Further analyses of these gene ontologies driven by truncating *ASXL1* mutations identified dysregulations in the biological processes anterior/posterior patterning (GO:0009952, padj 5.61E-03), pattern specification (GO:0007389, padj 7.51E-03), which are known functions of *ASXL* genes, were driven by DEGs upregulated in both BOS and AML-ASXL1. DEGs downregulated in both datasets drove immune response-activation pathways (GO:0002429, padj 5.61E-003). On the other hand, DEGs that were upregulated in BOS but downregulated in AML-ASXL1, were specific to T cell activation (GO:0042110, padj 4.35E-06) and T cell differentiation (GO:0045580, padj 2.71E-02) among other T cell functions, which aligns with the disease pathologies.

Wnt-signaling pathways are dysregulated across BOS and AML-ASXL1 samples.

Of note, the beta-catenin-TCF complex assembly, a key component of the canonical Wnt signaling pathway significantly dysregulated (padj 2.71E-02) in the DEGs across both datasets. The canonical Wnt signaling pathway was recently identified to be aberrantly upregulated in BOS patient-derived samples across blood and fibroblast tissues ¹⁸. Specifically, this dysregulation of Wnt signaling in the BOS and AML-ASXL1 datasets was driven by dysregulation in the blood samples (GO:0060070, padj 2.57E-03) and not the bone marrow samples suggesting dysregulations specific to the state of differentiation.

We further analyzed this dysregulation of Wnt signaling through comprehensive gene ontology network analysis ^{38,39}. We first depicted the consensus of DEGs against the KEGG pathway hsa05200 which represents the kernel regulatory factors that contribute to the initiation and progression of pan-cancer. This identified the KEGG pathway hsa04310, representing Wnt signaling pathway, as one of the most significantly dysregulated of the pan-cancer pathways (**Figure 6-4**).

The data highlight key dysregulated genes across different cell types in BOS ¹⁸, that are also some of the most highly dysregulated genes in the AML-ASXL1 dataset, revealing potential biomarkers and therapeutic targets. These genes include *Vang-like 2 (VANGL2)*, a member of the planar cell polarity pathway⁴⁰, *glutamate ionotropic receptor kainate type subunit 5 (GRIK5)*, a pre- and post-synaptic receptor for glutamate, a crucial excitatory neurotransmitter of the central nervous system ^{18,41}, and *gremlin 2 (GREM2)*, a bone morphogenetic protein antagonist involved in developmental processes and tissue differentiation ⁴², as well as the transmembrane low-density lipoprotein receptor-related proteins 5 and 6 (*LRP5* and *LRP6*), which are key components of the Wnt signaling pathway. In particular, we identified significant upregulation of *VANGL2* (BOS log2FC 3.80, AML-ASXL1 log2FC 4.24) (**Figure 6-3C**) and *GRIK5* (BOS log2FC 3.83, AML-ASXL1 log2FC 4.67) (**Figure 6-3D**) in both BOS and AML-ASXL1 and significant downregulation of *GREM2* in both conditions (BOS log2FC -2.48, AML-ASXL1 log2FC -1.67) (**Figure 6-3E, Table 6-5**). Similarly, we identified significant upregulation of *LRP5* (BOS log2FC 1.64, AML-ASXL1 log2FC 2.65) (**Figure 6-3F**) and *LRP6* (BOS log2FC 1.63, AML-ASXL1 log2FC 2.21) (**Figure 6-3G**) in both BOS and AML-ASXL1 samples.

These genes were already previously established in our study across different tissues in BOS as key biomarkers ¹⁸. A large scale study comparing AML-ASXL1 to AML also identified *VANGL2*, *LRP5* and *LRP6* as three of the most significantly upregulated genes using a limited microarray probe-set ⁴³, and were also identified across different tissues in BOS patient samples ¹⁸, supporting aberrant Wnt signaling in the context of *ASXL1* mutations. These expression

changes suggest roles for these genes in the pathophysiology of disorders associated with *ASXL1* mutations and highlight their potential as biomarkers for *ASXL1* mutations.

Interestingly, we identified significant upregulation of *ASXL1* expression in AML-*ASXL1* samples compared to tissue-matched controls (log2FC 0.46, padj 1.11E-02) but no significant dysregulation of *ASXL1* expression in BOS (**Figure 6-3H**). This may lend insight to disease-specific mechanisms.

ASXL1 mutations in BOS and AML-*ASXL1* identified DNA methylation driven de-repression of *HOX* genes

Our integrated analysis of DNAm leverages our previously published BOS-specific DNAm episignature ²⁹ to distinguish pathogenic *ASXL1* mutations from normotypic matched controls and variants of uncertain significance (VUS) in *ASXL1* ²⁹. To assess whether AML-*ASXL1* samples and BOS samples shared epigenetic mechanisms, we obtained Illumina 450K DNAm data for AML samples (n=6) from TCGA on the GBC repository ^{26,44}. This comprised individuals harboring somatic mutations in *ASXL1* (n=3, AML-*ASXL1*) or somatic mutations in other genes (n=3, AML).

We compared DNAm episignature profiles of blood samples from typically developing controls (n=26), BOS (n=8), AML-*ASXL1* (n=3), and AML controls (n=3). PCA based on the previously established BOS DNAm episignature sites (413 CpG sites) ²⁹ revealed significant differences in the epigenetic landscapes between the samples. AML-*ASXL1* samples clustered distinctly from other AML samples without *ASXL1* mutations and, instead, clustered closer to BOS samples, reflecting a shared epigenetic landscape driven by *ASXL1* mutations (**Figure 6-5A**).

Furthermore, unsupervised clustering of CpG methylation across the 413 BOS episignature sites confirmed the clustering pattern of AML-*ASXL1* alongside BOS samples,

distinct from AML controls (**Figure 6-5B**). This suggests that *ASXL1*-driven epigenetic alterations transcend disease type in defining DNAm patterns.

We identified 90 CpG sites corresponding to transcriptional start sites (TSS), 200bp upstream (TSS200) or 1500bp upstream (TSS1500) and the mean β values for each group are shown in **Table 6-7**.

We focused on the methylation status of posterior homeobox A (*HOXA*) genes, which are known to be regulated by *ASXL1*, and *HOXB3* and *HOXB4* (**Table 6-8**). The latter two genes are expressed in hematopoietic stem cells (HSCs) and progenitors as “master genes in early hematopoiesis,” and exhibit lineage and differentiation stage-restricted expression^{45–47}.

The *HOXB4* 3'UTR was hypermethylated in BOS and AML-*ASXL1* samples compared to respective controls (**Figure 6-5C**). Consistent with findings that 3'UTR methylation correlates positively with gene expression⁴⁸, RNA-seq data showed significant upregulation of *HOXB4* (BOS log2FC 1.41, AML-*ASXL1* log2FC 1.89) (**Figure 6-5D**). Significant upregulation was also observed in *HOXB3* (BOS log2FC 1.34, AML-*ASXL1* log2FC 2.48) (**Figure 6-6A**). While overexpression of *HOXB4* has been shown to drive enhanced hematopoietic stem cell regeneration, deficiency of *Hoxb3* or *Hoxb4* leads to defects in proliferative ability of hematopoietic progenitors⁴⁸. Furthermore, one of the key targets of *HOXB4* is Wnt signaling⁵⁰. Our investigation of Wnt signaling revealed significant upregulation of Wnt signaling coreceptors *LRP5* and *LRP6* in both BOS and AML-*ASXL1* samples (**Figure 6-3F, Figure 6-3G**).

DNAm analysis further identified hypomethylation at specific CpG sites upstream of the *HOXA11* transcriptional start site (cg00705992 and cg16038003) in BOS and AML-*ASXL1* samples, supporting literature that truncating *ASXL1* mutations lead to derepression of posterior *HOXA* genes (**Figure 6-5E, Figure 6-7A**). Interestingly, while *ASXL1* mutations drove significant upregulation of *HOXA11* transcript expression (log2FC 3.87, padj 1.33E-06) in AML-*ASXL1*, there was no significant dysregulation of *HOXA11* in BOS. Similar trends were identified in other

posterior *HOXA* genes including *HOXA5* (**Figure 6-6B**, **Figure 6-7B**) and *HOXA9* (**Figure 6-6C**); there was no significant dysregulation in BOS samples.

These findings underscore the pervasive influence of *ASXL1* mutations in modulating the epigenetic and transcriptomic landscapes across diseases, promoting abnormal gene expression and signaling pathways crucial for disease pathology.

Differential *RUNX3* isoform expression in BOS and AML-*ASXL1*

Differential exon usage (DEU) analysis was conducted for BOS blood (n=8), blood controls (n=10), AML-*ASXL1* blood (n=4), and AML blood controls (n=6). We profiled the transcriptome to determine whether *ASXL1* mutations drive downstream effects on exon usage and isoform expression in other genes (**Figure 6-8A**).

Exon counts were calculated for a total of 597,773 exonic regions in 56,940 genes and analyzed DEU. A total of 88 significant exonic regions from 63 affected genes were identified in BOS blood ($p_{adj} < 0.05$). Some of the most significant exons from BOS blood were in the following genes: runt-related transcription factor 3 (*RUNX3*), plasmolipin (*PLLP*), ST6 beta-galactoside alpha-2,6-sialyltransferase 1 (*ST6GAL1*), serpin family F member 1 (*SERPINF1*), and pleckstrin homology and RhoGEF domain containing G6 (*PLEKHG6*). Many of these genes were also identified as DEGs in the multi-omics study by Lin et al. (2023) and are involved in embryonic/brain development and neuronal differentiation^{18,51,52}. In AML-*ASXL1* blood, a total of 11624 significant exon bins from 4521 affected genes were identified ($p_{adj} < 0.05$). Many of the most significant exons in AML-*ASXL1* blood had immune- or cancer-related function, such as TBL1X/Y related 1 (*TBL1XR1*), DEAD-box helicase 42 (*DDX42*), Rho associated coiled-coil containing protein kinase 1 (*ROCK1*), neurobeachin like 2 (*NBEAL2*), and UDP-glucose pyrophosphorylase 2 (*UGP2*).

Notably, 5 exon bins from *RUNX3* were among the most significant DEUs identified in BOS blood (**Figure 6-8B**). All 5 exon bins were significantly upregulated in BOS samples,

corresponding to higher usage of exon 1 of *RUNX3* (**Figure 6-8C**). Heatmap clustering revealed that this DEU was correlated with disease status and not tissue type, when analyzing the BOS blood and fibroblast samples from the multi-omics study on BOS (**Figure 6-12A**)¹⁸. Comparison of exon base means for each *RUNX3* exonic bin in BOS showed that *RUNX3* is more highly expressed in blood than fibroblast (**Figure 6-12E**), suggesting important functionality in blood cells.

Similarly, we observed DEU of 6 exon bins in *RUNX3* for AML-ASXL1 blood samples (**Figure 6-8D**). However, based on the exon usage plot, we identified overall upregulation of the first 10 exon bins and *down*regulation of the last exon, which is the opposite effect observed in BOS blood, where we saw *up*regulation of the first exon (**Figure 6-8E**). *RUNX3* is known to have two main transcripts expressed in blood cells— p46 which includes the first exon of the gene, and p44 which does not⁵³. Thus, our results suggest that the p44 transcript is more highly expressed in AML-ASXL1 blood, while the p46 transcript is more highly expressed in BOS blood.

ASXL1 exon usage displays tissue-specific patterns in BOS and AML-ASXL1

We did not observe significant DEU within *ASXL1* itself for BOS blood or fibroblast (**Figure 6-11A-B**). Heatmap clustering of *ASXL1* exon usage in BOS blood and fibroblast compared to controls showed grouping by tissue type, indicating that tissue specificity, not disease status, drives exon usage for *ASXL1* in BOS (**Figure 6-11D**). In AML-ASXL1 blood, however, we observed DEU for multiple exons bins of *ASXL1*, but this did not correspond to differential isoform expression (**Figure 6-11C**).

Discussion

This study provides the first detailed exploration into the epigenetic and transcriptomic landscapes associated with *ASXL1* mutations in BOS and AML-ASXL1. BOS and AML present with vastly different clinical manifestations—BOS as a congenital disorder characterized by

developmental delays and multiple malformations, and AML as a malignancy of the blood and bone marrow. Yet, both are linked by heterozygous truncating mutations in *ASXL1* (**Figure 6-1B**) that strike at the respective stem cell. By employing RNA-seq and DNAm analyses, we have uncovered significant cross-disease epigenetic signatures and gene expression alterations that underlie the profound influence of *ASXL1* mutations across distinct yet pathophysiologically connected diseases.

Identification of biomarkers and role of Wnt signaling in *ASXL1* mutation-driven diseases

Our analysis demonstrates a strong influence of *ASXL1* mutations on gene expression profiles, with significant numbers of DEGs identified in blood and bone marrow samples of AML patients as well as in BOS patients. The identification of 843 common DEGs between AML-*ASXL1* and BOS datasets, with a substantial proportion dysregulated in the same direction, underscores shared molecular dysfunctions that underlie both conditions. This overlap suggests that, beyond the specific symptoms or affected tissues, there may be a common molecular foundation to the disorders driven by mutations in the *ASXL1* gene. To further investigate the extent of this shared dysregulation in the transcriptome, we analyzed differential exon usage and isoform expression in BOS and AML-*ASXL1* and identified differential isoform expression of key genes such as *RUNX3* that may play a key role in the distinct pathogenesis of these diseases.

This shared dysregulation is particularly evident in genes involved in epigenetic regulation, chromatin modification, and the canonical Wnt signaling pathway, which are critical for cell fate determination and proliferation. These pathways suggest a common thread in the disease mechanism, likely through the chromatin-modifying function of *ASXL1*. Notably, the identification of the beta-catenin-TCF complex assembly and its significant dysregulation highlights the crucial role of the Wnt signaling pathway in the pathology of both diseases. Our

findings suggest that this pathway could serve as a therapeutic target, given its involvement in cell differentiation and proliferation.

The dysregulation of key biomarkers such as *VANGL2*, *GRIK5*, *GREM2*, *LRP5*, and *LRP6* across both BOS and AML suggests roles for these genes in the pathophysiology of disorders associated with *ASXL1* mutations. The significant upregulation of these genes across different tissue types in both diseases highlights their potential as diagnostic and therapeutic targets. In our study, the *ASXL1* mutation-driven upregulation of *VANGL2*, *GRIK5*, *LRP5*, and *LRP6*, and downregulation of *GREM2* presents unique opportunities to redefine our understanding of BOS and AML-*ASXL1*.

VANGL2, identified as a key player in the planar cell polarity pathway, showed significant upregulation in both BOS and AML-*ASXL1* (**Figure 6-3A, Figure 6-3C**). This upregulation suggests a broader role for *VANGL2* in diseases involving *ASXL1* mutations, potentially influencing cell morphology and migration patterns that are pivotal in both oncogenic transformations and developmental anomalies. The consistent upregulation and epigenetic dysregulation across different datasets and studies ¹⁸, including ours and those reported by Metzeler et al. (2011) ⁴³, underscores its potential as a robust biomarker, insights into mechanisms of disease progression in these disorders, and suggest avenues for therapeutic intervention in conditions characterized by *ASXL1* mutations.

GRIK5, a receptor for glutamate, was also notably upregulated in our study (**Figure 6-3A, Figure 6-3D**). Considering the critical role of glutamate in neuronal and possibly non-neuronal excitatory signaling, the dysregulation of *GRIK5* might influence cellular communication and proliferation in leukemia.

Conversely, *GREM2*, a known antagonist in the BMP pathway, showed significant downregulation (**Figure 6-3A, Figure 6-3E**). This alteration could disrupt normal developmental and differentiation processes, potentially contributing to the pathogenesis of BOS and

AML-*ASXL1*, and offering a novel angle through modulation of BMP signaling for therapeutic development.

Similarly, the Wnt signaling co-receptors *LRP5* and *LRP6* are upregulated (**Figure 6-3A**, **Figure 6-3F**, **Figure 6-3G**), suggesting an aberrant activation of Wnt signaling associated with *ASXL1* mutations. This is consistent with findings from previous studies, including those by Lin et al. (2023) and Metzeler et al. (2011)^{18,43}. Wnt signaling is crucial for stem cell maintenance and differentiation, processes that are fundamentally disrupted in both cancerous conditions like AML and developmental disorders like BOS. The significant upregulation of these genes in the context of *ASXL1* mutations suggests a converging pathway that might be pivotal in the pathophysiology of both diseases. This linkage suggests that the Wnt pathway, often implicated in cancer and developmental disorders, could be aberrantly activated in conditions associated with *ASXL1* mutations, thereby driving disease processes through a common molecular mechanism. These findings highlight the potential for targeting Wnt signaling as a targeted therapeutic strategy.

Interestingly, the specific upregulation of *ASXL1* in AML-*ASXL1* samples points to possible disease-specific mechanisms of *ASXL1*-associated malignancies. This differential expression pattern offers insights into the distinct molecular pathways activated in AML compared to BOS, which may have implications for targeted therapies.

Overall, the consistent dysregulation of key genes across different diseases suggests that they are not merely artifacts of disease-specific processes but are potentially pivotal drivers of pathophysiology associated with *ASXL1* mutations. The discovery of these cross-disease biomarkers offers a promising avenue for further research and development of diagnostic tools and therapeutic strategies. Therapeutics targeting *VANGL2* and *LRP6* could potentially be used to advance pan-disease therapies resulting from *ASXL1* mutations, irrespective of the traditional disease classification.

Truncating *ASXL1* mutations drive dysregulation of *HOX* genes

The epigenetic alterations delineated in our study, particularly through DNAm profiling using a BOS-specific DNAm episinature, demonstrated that AML-*ASXL1* and BOS share a closer epigenetic landscape compared to AML samples harboring mutations in other genes (**Figure 6-5A, Figure 6-5B**). This supports the hypothesis that *ASXL1* mutations exert a unique epigenetic influence, potentially driven by the impact of *ASXL1* mutations on chromatin modification processes.

The dysregulation of *ASXL1*, whether by loss or mutation, is established to disrupt recruitment of *ASXL1* to polycomb repressive complex 2 (PRC2), leading to derepression of *HOXA* genes⁵⁴. *ASXL1* interacts with EZH2, one of the key components of PRC2, which is responsible for depositing the H3K27me3 histone modification, and thus regulating chromatin accessibility⁵⁴. This modification is crucial for the repression of gene expression during development. EZH2 also interacts with DNA methyltransferases (DNMTs)^{29,54}, thus aberrant interactions caused by mutations in *ASXL1* could drive epigenetic dysregulation in BOS and AML-*ASXL1*.

Central to our findings are the epigenetic and corresponding transcriptomic dysregulation of *HOX* genes, particularly in the posterior *HOXA* cluster and the *HOXB* cluster. These genes play critical roles in development and hematopoiesis and are known to be tightly regulated by epigenetic mechanisms. The hypermethylation of *HOXB4* 3'UTR in both BOS and AML-*ASXL1* samples (**Figure 6-5C**) and its correlation with increased gene expression (**Figure 6-5D**) align with the role of *HOXB4* in enhancing hematopoietic stem cell regeneration. Such epigenetic modifications potentially mediate the oncogenic and developmental pathways activated in AML and BOS, respectively.

Our analyses also identified hypomethylation at specific CpG sites upstream of the *HOXA11* TSS in BOS and AML-*ASXL1* samples (**Figure 6-5E, Figure 6-7A**), which supports the current understanding that truncating *ASXL1* mutations lead to the derepression of posterior

HOXA genes, which are instrumental in developmental and cellular differentiation processes. However, this derepression is particularly pronounced in AML-*ASXL1*, where significant upregulation of *HOXA11* transcript expression is in stark contrast to the non-dysregulated expression in BOS (**Figure 6-5F**). This pattern of differential expression across disease states provides key insights into disease-specific epigenetic regulation by *ASXL1* mutations. Similar trends were observed with other posterior *HOXA* genes, such as *HOXA5* and *HOXA9*, further supporting the selective impact of these mutations on gene expression profiles in AML compared to BOS.

The significant upregulation of key genes within the *HOXA* cluster in AML-*ASXL1* highlights a mutation-driven mechanism that potentially contributes to leukemogenesis through the deregulation of genes critical for hematopoietic differentiation and proliferation.

Differential isoform usage of *RUNX3* in BOS and AML-*ASXL1*

Through DEU analysis, we showed that while BOS is caused by truncating *ASXL1* mutations, there is no difference in *ASXL1* exon usage or isoform expression. Many of the most significant genes affected by DEU in BOS samples had functions involving embryonic and neural crest development, while the genes affected in AML-*ASXL1* had immune- or cancer-related function. This is reflective of the phenotypes of the disorders because while BOS patients have severe neurodevelopmental defects, AML-*ASXL1* patients do not, and instead experience immunodeficiencies associated with cancer.

Key to our results was the identification of dysregulation of *RUNX3* in both BOS and AML-*ASXL1*. The *RUNX3* gene, originally named acute myeloid leukemia gene 2 (*AML2*), is located on human chromosome 1p36.1. It is highly expressed in the hematopoietic system, including the spleen, thymus and peripheral blood mononuclear cells, and has been found to play a role in myeloid and B-cell differentiation⁵³. Furthermore, it acts as a tumor suppressor and apoptosis inducer in many forms of cancer, including myeloid leukemias⁵⁵. There are three

genes in the *RUNX* family— *RUNX1*, *RUNX2*, and *RUNX3*— which all contain the highly conserved runt domain (RD) that covers the N-terminal region. *RUNX1* and *RUNX2* are well-studied as drivers of organ development, and found to be associated with various human disorders including acute myeloid leukemia and insufficient osteogenesis, respectively ⁵³.

While much is known on the functional and disease-causing roles of *RUNX1* and *RUNX2*, less is known on *RUNX3*. *RUNX3* facilitates the binding of RUNX proteins to DNA and its partner subunit Core Binding Factor β (CBF β), as well as helps regulate transcription ⁵³. It is highly expressed in the hematopoietic system, including the spleen, thymus and peripheral blood mononuclear cells, and is speculated to play a role in myeloid and B-cell differentiation ⁵³. The gene encodes the RUNX3 transcription factor that plays a role in regulating cell differentiation, growth and tissue development. Furthermore, *RUNX3* acts as a tumor suppressor and apoptosis inducer in many forms of cancer, including myeloid leukemias ⁵⁵.

RUNX3 contains six exons and spans 67 kb, half of this being the first intron. Its transcription is regulated by two promoter regions, P1 and P2. The P1 promoter regulates transcription of the RUNX3/p44 transcript (exons 2-6, ENST00000308873.11), while the P2 promoter regulates RUNX3/p46 (exons 1-6, ENST00000399916.5 and ENST00000338888.4); the longer *RUNX3*/p46 isoform contains 19 additional unique amino acids encoded by exon 1 (NCBI; Bangsow et al. 2001). In BOS blood, we identified higher usage of exon 1 which is only included in the p46 transcript, indicating increased expression of the *RUNX3*/p46 transcript in this disease. On the contrary, in AML-ASXL1 blood, we found higher usage of the exons that are expressed in the *RUNX3*/p44 transcript and lower expression of exon 1, indicating increased expression of the *RUNX3*/p44 transcript in this disease.

Recent studies have revealed the extensive role of *RUNX3* in neuronal development in the mammalian nervous system, which may suggest the discovery of *RUNX3* dysregulation in BOS, a disease associated with disruption to normal neural development ⁵⁵. *RUNX3* was further identified as a super-enhancer that enhances the transcription of key oncogenes in various

cancers, and has shown to play a role in AML progression, which may explain the similar discovery of *RUNX3* dysregulation in AML-*ASXL1*. *RUNX3* mRNA and protein levels were highly expressed in AML patient cells compared to controls and positively associated with poor prognosis in AML, similar to *ASXL1*⁵⁶.

Overall, the dual function of *RUNX3* in neural development and as a tumor suppressor in the hematopoietic system, combined with our finding of differential isoform expression of *RUNX3/p46* and *RUNX3/p44* transcripts in BOS and AML-*ASXL1* blood, suggest that *RUNX3* isoforms may play a key role in the distinct pathogenesis of these diseases. Quantification of *RUNX3* isoform usage through single-cell RNA sequencing would further validate our findings, which hold potential as therapeutic targets for both diseases.

Conclusion

This comparative analysis provides a deeper understanding of the complex molecular underpinnings of *ASXL1* mutations in BOS and AML, highlighting shared and unique molecular features. By delineating the genetic, epigenetic, and transcriptomic impacts of these mutations, our study not only advances the understanding of the molecular pathology of these conditions but also sets the stage for the development of targeted therapeutic strategies that address the specific molecular alterations associated with *ASXL1* mutations.

Overall, our study advocates for a gene-centric approach in understanding the molecular basis of diseases associated with *ASXL1* mutations.

Our study contributes to a broader understanding of how genetic mutations can influence disease across traditional phenotypic boundaries. This not only challenges existing paradigms of disease classification but also opens new avenues for innovative therapeutic strategies that are based on molecular commonalities rather than clinical manifestations alone. This gene-centric perspective could redefine therapeutic strategies and promote the development of more precise and effective treatments for genetically driven disorders.

Figures and Tables

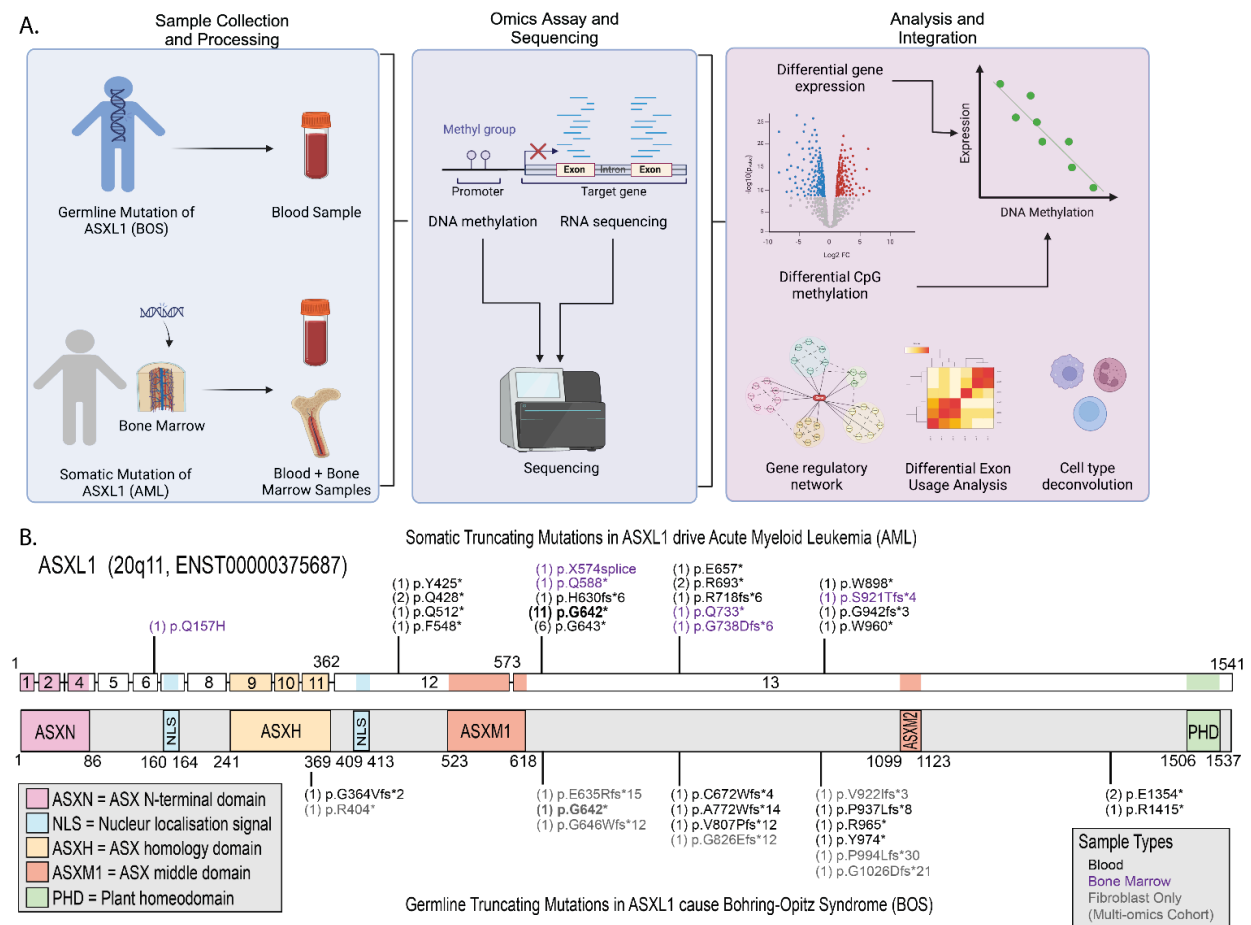


Figure 6-1: Study design and genetic landscape of ASXL1 mutations

(A) Detailed workflow from the collection of patient samples through to the generation and analysis of RNA-sequencing (RNA-seq) and DNA methylation (DNAm) data. Blood samples were collected from Bohring-Opitz syndrome (BOS) patients, and control individuals, and RNA-seq and DNAm were conducted. In combination with blood and bone marrow samples from acute myeloid leukemia with ASXL1 mutations (AML-ASXL1) or without (AML), RNA-seq and DNAm analysis was conducted to examine differential gene expression, differential CpG methylation, dysregulated gene regulatory networks, differential exon usage, and cell type deconvolution.

(B) The *ASXL1* gene, on chromosome 20q11, is illustrated, highlighting gene domains and the loci of germline mutations that cause BOS (below the gene) and somatic mutations that drive AML-*ASXL1* (above the gene) in this study. Common mutation sites (bold font), illustrate the genetic intersections between BOS and AML-*ASXL1*. Mutations from blood +/- fibroblast samples are denoted in black, bone marrow sample mutations in purple, and fibroblast-only sample mutations from our previous study are shown in grey.

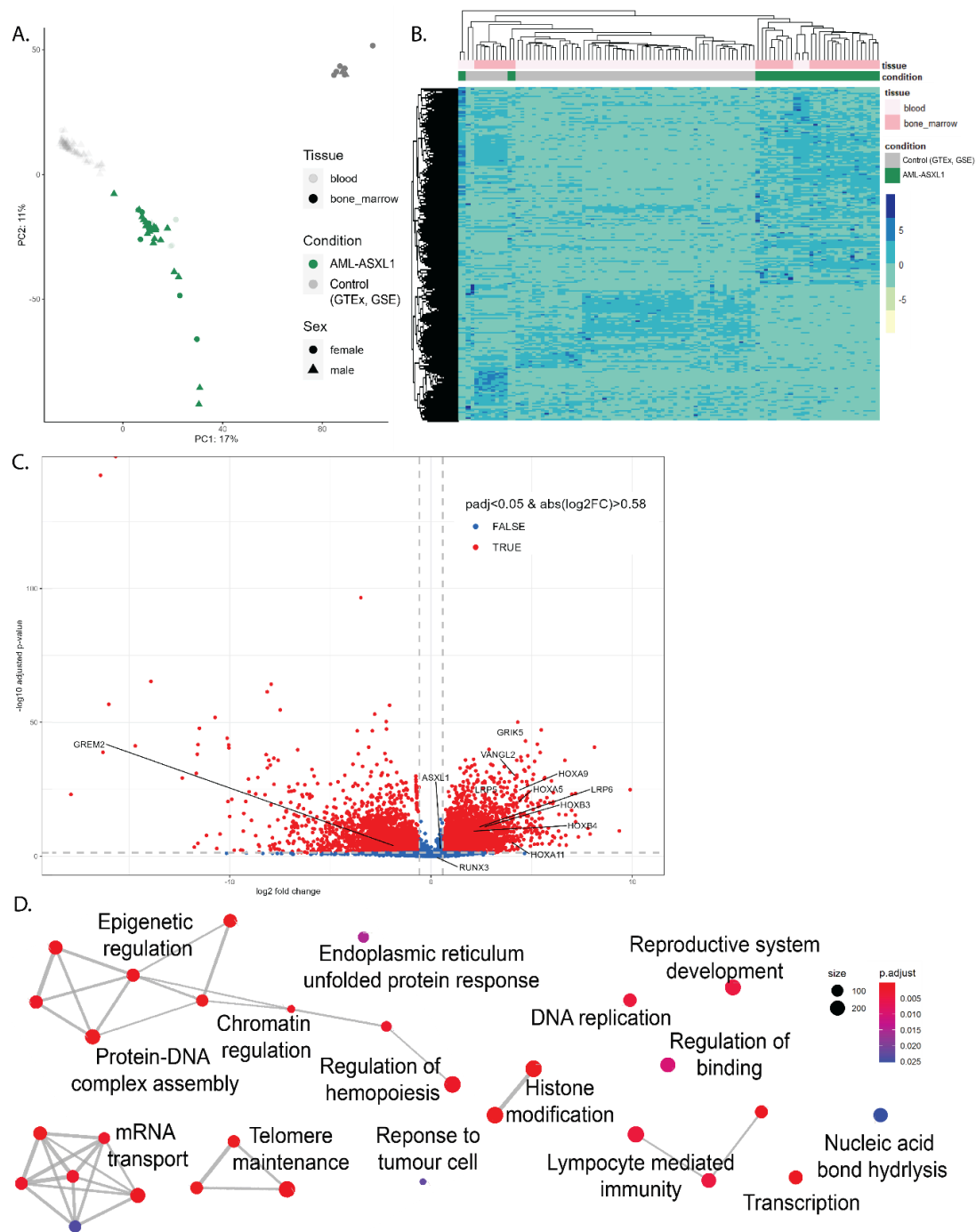


Figure 6-2: Integrative analysis of AML-ASXL1 samples compared to controls reveals distinct transcriptomic signatures.

(A) Principal component analysis (PCA) plot of all AML-ASXL1 samples sourced from BEAT-AML and TCGA, and all control samples sourced from GTEx and GSE120444. Samples are colored by condition (AML-ASXL1 = green, control = grey), shaded by tissue type (bone marrow = dark, blood = light), and the shapes represent sex (female = circle, male = triangle). The PCA plot demonstrates the distinct clustering patterns between AML-ASXL1 and control samples. **(B)** Heatmap showing differentially expressed genes (DEGs) between AML-ASXL1 and control samples. The heatmap highlights significant changes in gene expression patterns, providing insights into the molecular alterations associated with AML-ASXL1. **(C)** Volcano plot comparing AML-ASXL1 samples to controls, with log₂ fold change on the x-axis and -log₁₀ adjusted p-value on the y-axis. Relevant genes are labeled to highlight those with significant expression changes. **(D)** Gene ontology (GO) analysis of the DEGs, illustrating the enrichment of biological processes and pathways affected in AML-ASXL1. This analysis identifies key functional categories that are dysregulated in AML-ASXL1 compared to controls.

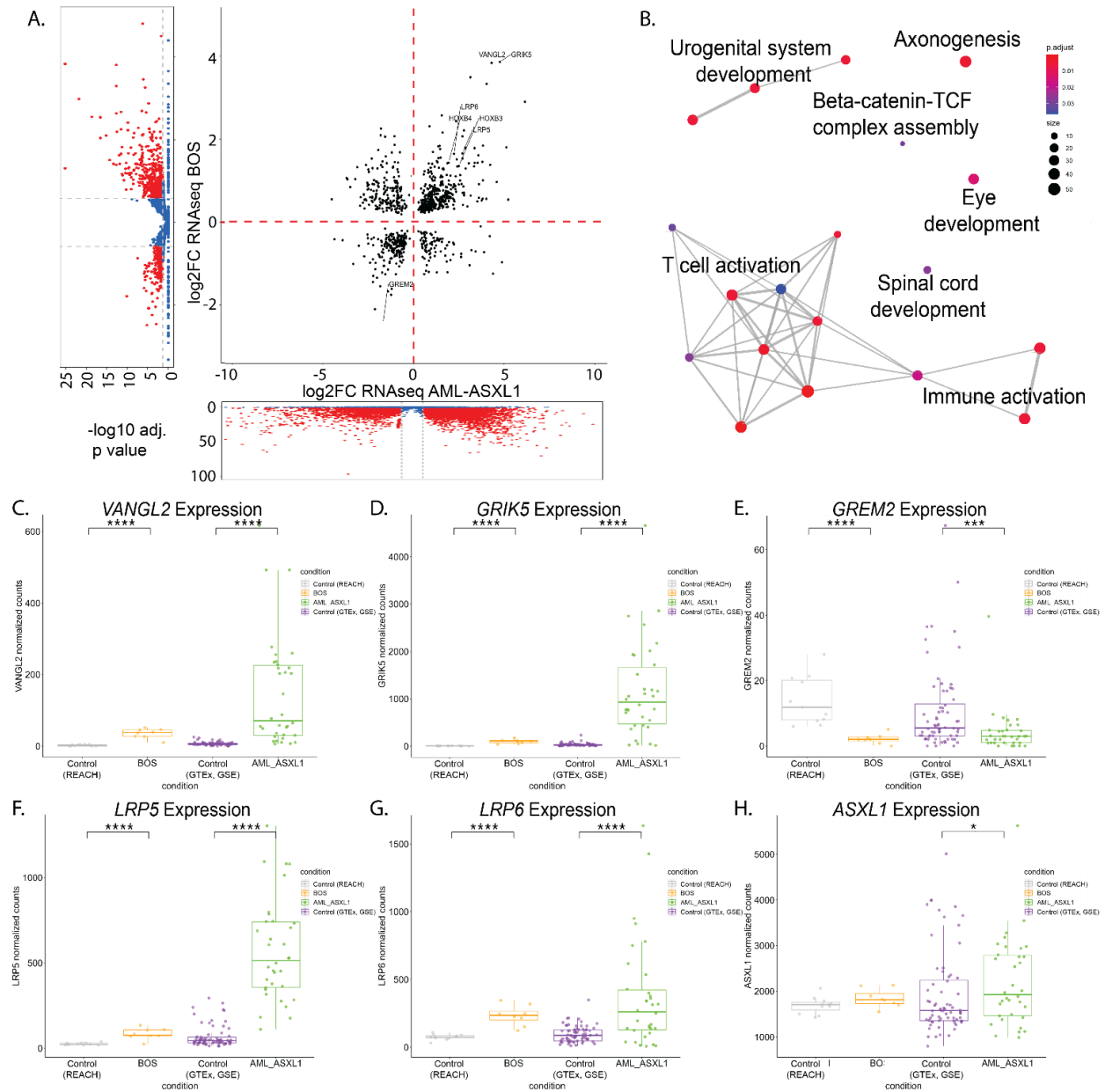


Figure 6-3: Transcriptomic alterations driven by *ASXL1* mutations in BOS and AML-*ASXL1* highlight key biomarkers and Wnt signaling dysregulation.

(A) Fold change integration plot of significant differential expression genes (DEGs, $p_{adj} < 0.05$) with large effect size ($|\log_2FC| \geq 0.58$) across AML-*ASXL1* and BOS compared to their respective controls.

(B) Gene ontology of common DEGs in AML-ASXL1 and BOS revealed enrichment in epigenetic regulation, histone modification, chromatin regulation, and protein-DNA complex assembly among other pathways.

Transcript expression analysis of BOS and AML-ASXL1 samples and their respective controls identified key biomarkers and showed significant **(C)** upregulation in *VANGL2* **(D)** upregulation in *GRIK5*, and **(E)** downregulation in *GREM2*. Wnt signaling co-receptors **(F)** LRP5 and **(G)** LRP6 were also significantly upregulated. (H) *ASXL1* was significantly upregulated in AML-ASXL1 but not in BOS.

ns or no stars denote $p\text{-value} > .05$, * denote $p\text{-value} \leq .05$, ** denote $p\text{-value} < .01$, *** denote $p\text{-value} < .001$, **** denote $p\text{-value} < .0001$

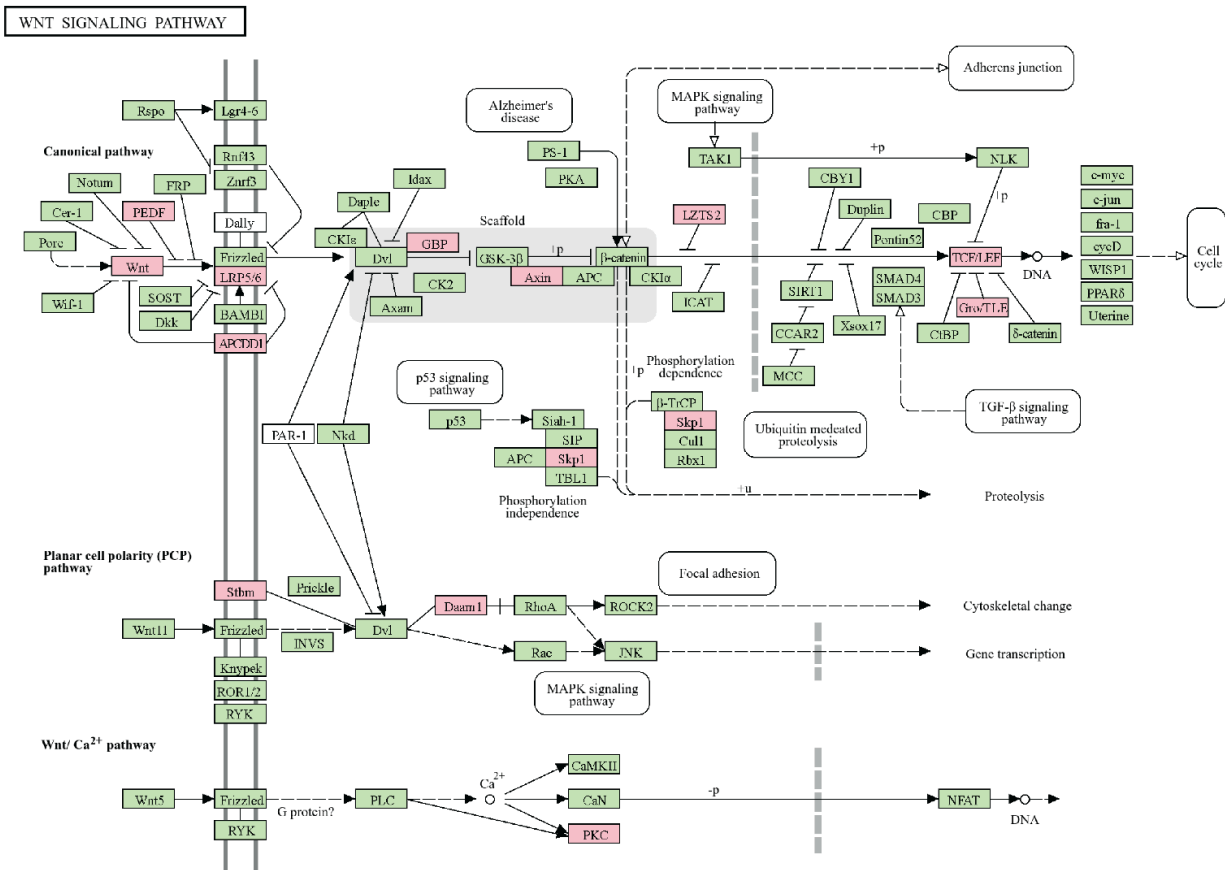


Figure 6-4: KEGG pathway analysis highlights Wnt signaling dysregulation in AML-ASXL1 and BOS.

KEGG Mapper visualization of differentially expressed genes (DEGs) within the Wnt signaling pathway, identifying transcriptomic dysregulation in both AML-ASXL1 and Bohring-Opitz syndrome (BOS). The analysis demonstrates the shared molecular alterations in Wnt pathway components, highlighting key dysregulated genes and their roles in disease pathogenesis.

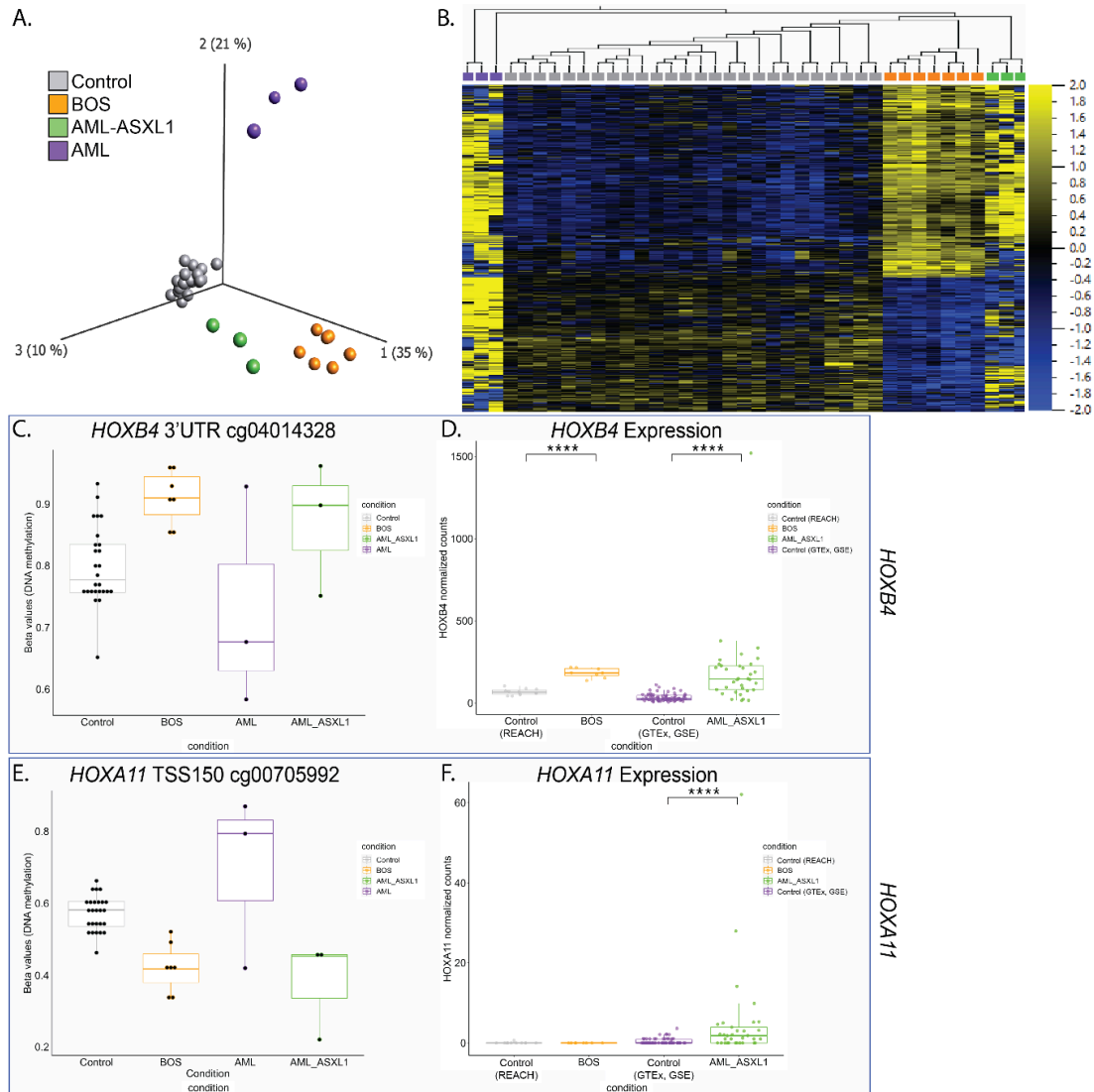


Figure 6-5: Epigenetic landscape of *ASXL1* mutations in BOS and AML-*ASXL1* highlights de-repression of *HOX* genes.

(A) Principal component analysis (PCA) plot demonstrates the closer clustering of acute myeloid leukemia with *ASXL1* mutations (AML-*ASXL1*) samples (green, n=3) with Bohring-Opitz syndrome (BOS) samples (orange, n=8) compared to AML samples with somatic mutations in other genes (purple, n=3), and control samples (grey, n=26), illustrating shared epigenetic landscapes driven by *ASXL1* mutation status.

(B) Heatmap representing DNA methylation patterns using the 413 episites identified in the BOS DNAm episignature depicts a consistent clustering pattern of the unique clustering of AML-ASXL1 samples alongside BOS samples.

(C) DNAm β values for HOXB4 3'UTR at CpG site cg04014328 highlight the hypermethylation in both BOS and AML-ASXL1 patients compared to controls and other AML samples. **(D)** RNA-seq data demonstrates significant upregulation of *HOXB4* and in BOS and AML-ASXL1 compared to their respective controls.

(E) DNAm β values for HOXA11 TSS1500 at CpG site cg00705992 highlight the hypomethylation in both BOS and AML-ASXL1 patients compared to controls and other AML samples. **(F)** RNA-seq data demonstrates significant upregulation of *HOXA11* in AML-ASXL1 but not in BOS.

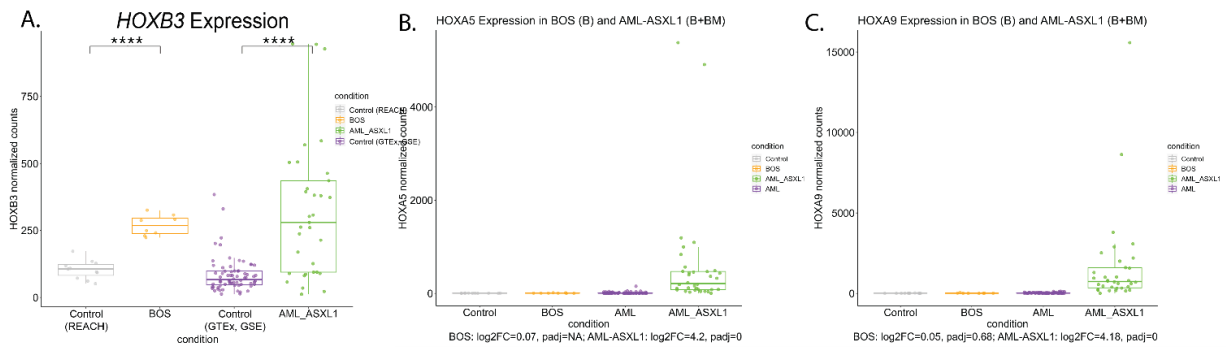


Figure 6-6: RNA-seq of HOX gene expression boxplots

RNA-seq data in BOS and AML-ASXL1 compared to their respective controls demonstrates **(A)** significant upregulation of *HOXB3* in both diseases, **(B)** significant upregulation of *HOXA5* in AML-ASXL1 only, and **(C)** significant upregulation of *HOXA9* in AML-ASXL1 only.

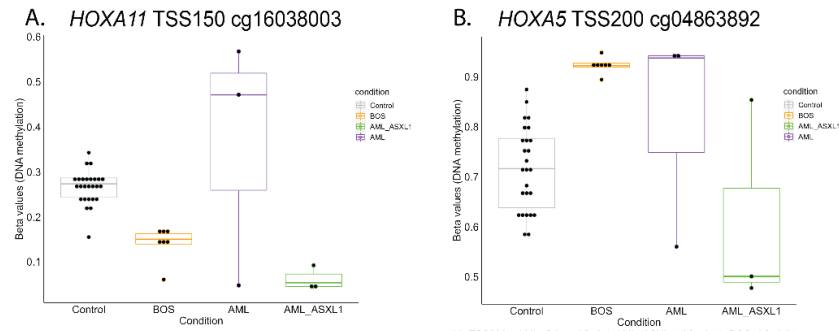


Figure 6-7: DNAm of posterior *HOXA* TSS CpG site

(A) DNAm β values for *HOXA11* TSS1500 at CpG site cg16038003 highlight the hypomethylation in both BOS and AML-ASXL1 patients compared to respective controls. **(B)** DNAm β values for *HOXA5* TSS200 at CpG site cg0486392 highlights the hypermethylation in BOS and hypomethylation in AML-ASXL1 patients compared to respective controls.

(A) The DEXSeq framework was applied to RNA-seq data and used to obtain exon counts and conduct differential exon usage (DEU) analysis in patients vs controls, as well as isoform usage and gene ontology analysis. An integrated analysis was conducted between BOS and AML-ASXL1 samples.

(B) Fold change plot of significant DEUs ($\text{padj} < 0.05$) in BOS patients compared to controls (orange background) highlights RUNX3 as a key affected gene with multiple DEUs. **(C)** Fitted expression exon usage plot for RUNX3 in BOS blood, with significant DEUs are indicated with a pink line. Cases are shown in red and controls in blue. The boxed transcript highlights the primary transcript observed in BOS blood based on DEU.

(D) Fold change plot of significant DEUs ($\text{padj} < 0.05$) in AML-ASXL1 patients compared to controls (green background) also highlights RUNX3 as an affected gene with multiple DEUs.

(E) Fitted expression exon usage plot for RUNX3 in AML-ASXL1 blood. Cases are shown in red and controls in blue. The boxed transcript highlights the primary transcript observed in AML-ASXL1 blood based on DEU.

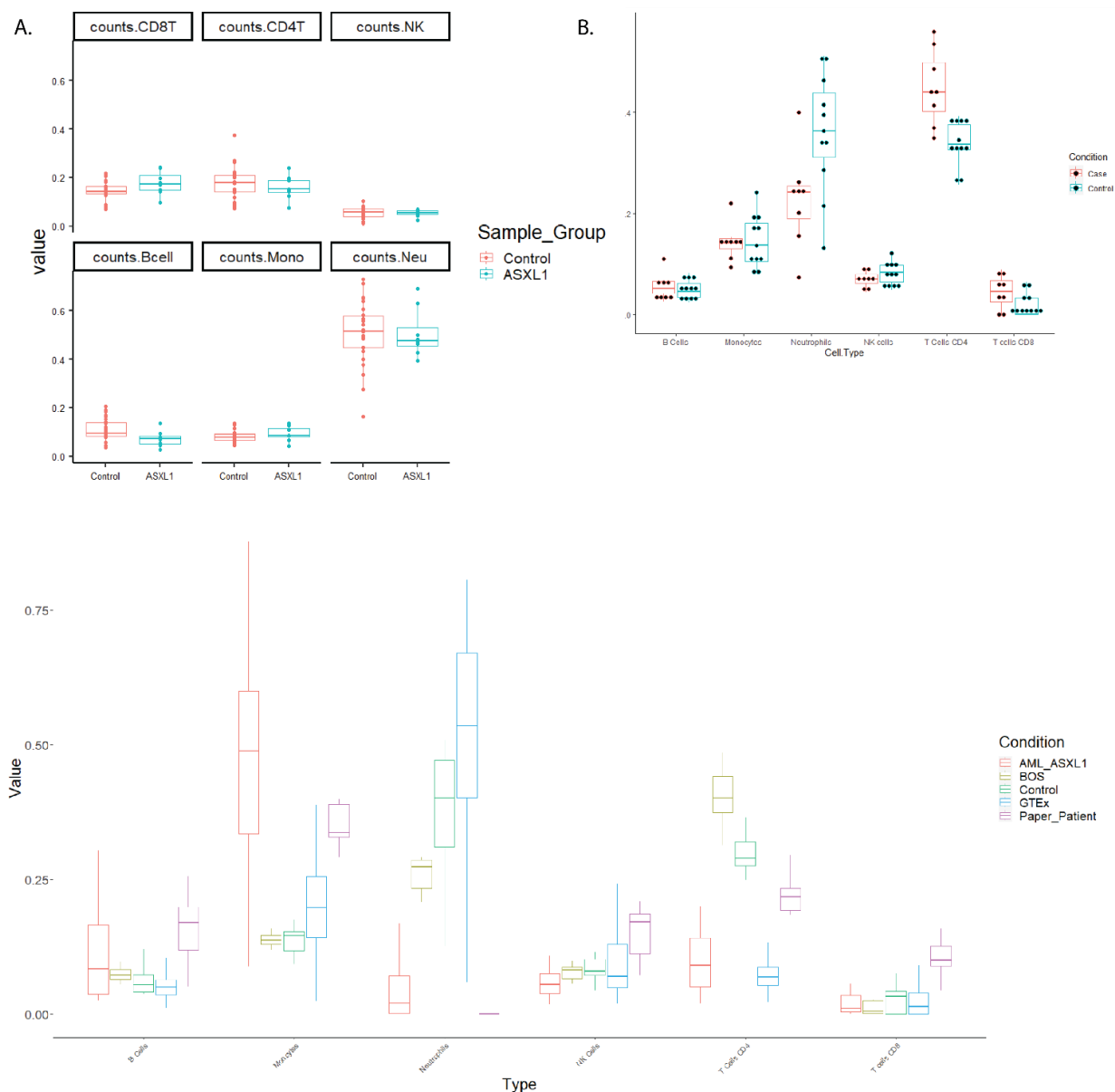


Figure 6-9: Cell type distribution in blood and bone marrow samples reveals no significant differences in immune cell proportions in BOS.

(A) Deconvolution of DNA methylation (DNAm) blood data, comparing the proportion of immune cell types between controls (red) and BOS samples (blue). The analysis includes CD8+ T cells, CD4+ T cells, NK cells, B cells, monocytes, and neutrophils. **(B)** Deconvolution of RNA-seq blood data, comparing the proportion of immune cell types between BOS and control samples. The analysis includes CD8+ T cells, CD4+ T cells, NK cells, B cells, monocytes, and

neutrophils. **(C)** Comparison of cell type distributions for RNA-seq data from blood and bone marrow samples across BOS, controls (blood), AML-ASXL1, and controls (bone marrow).

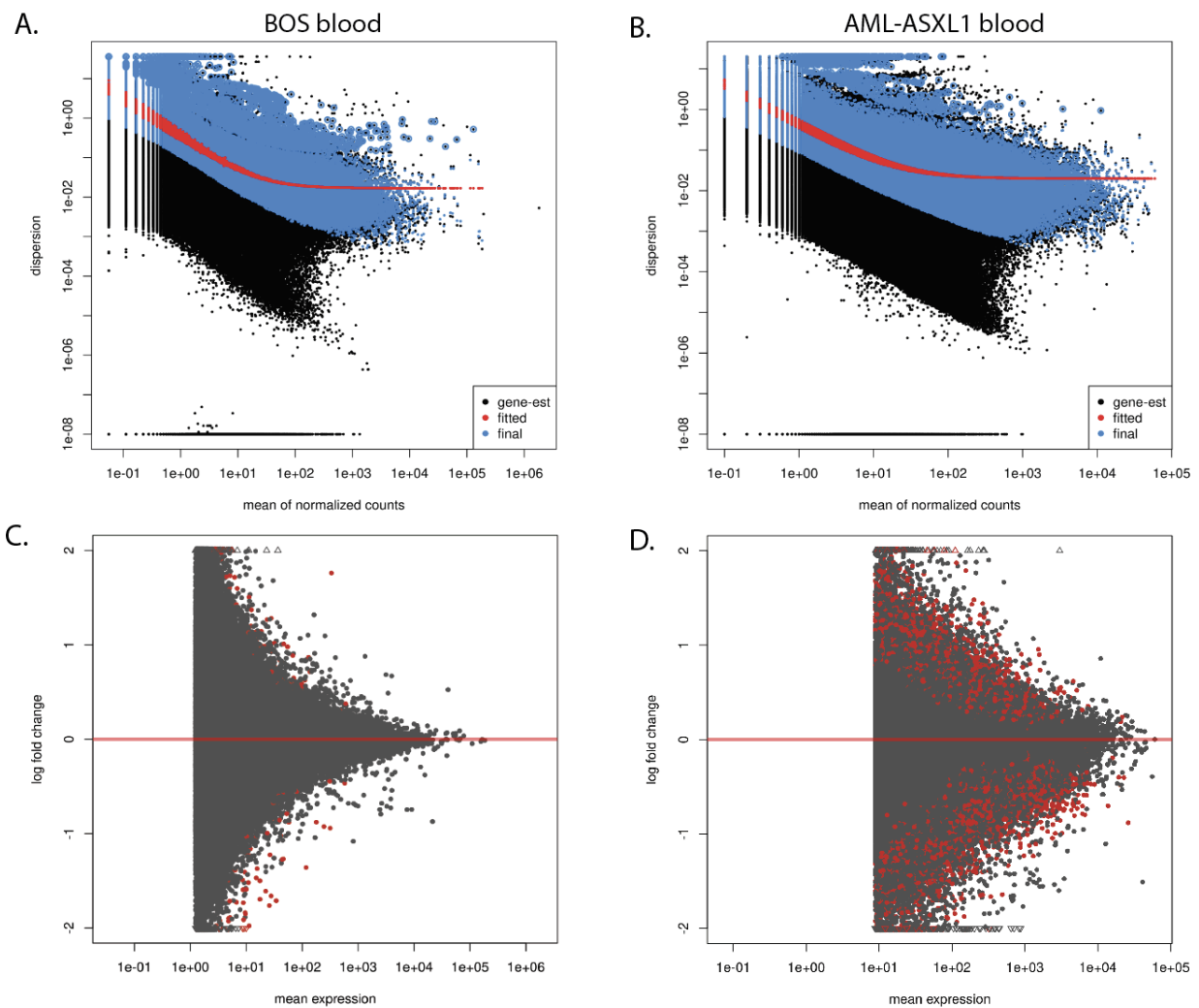


Figure 6-10: DEXSeq quality control analysis.

Dispersion (per-exon dispersion estimates versus mean normalized count) plot for **(A)** BOS blood and **(B)** AML-ASXL1 blood. Initial per-exon dispersion estimates are shown in black, fitted mean-dispersion function in red, and shrunk values in blue.

MA (logarithm of fold change versus mean normalized count per exon) plot with an FDR=0.05 for **(C)** BOS blood and **(D)** AML-ASXL1 blood.

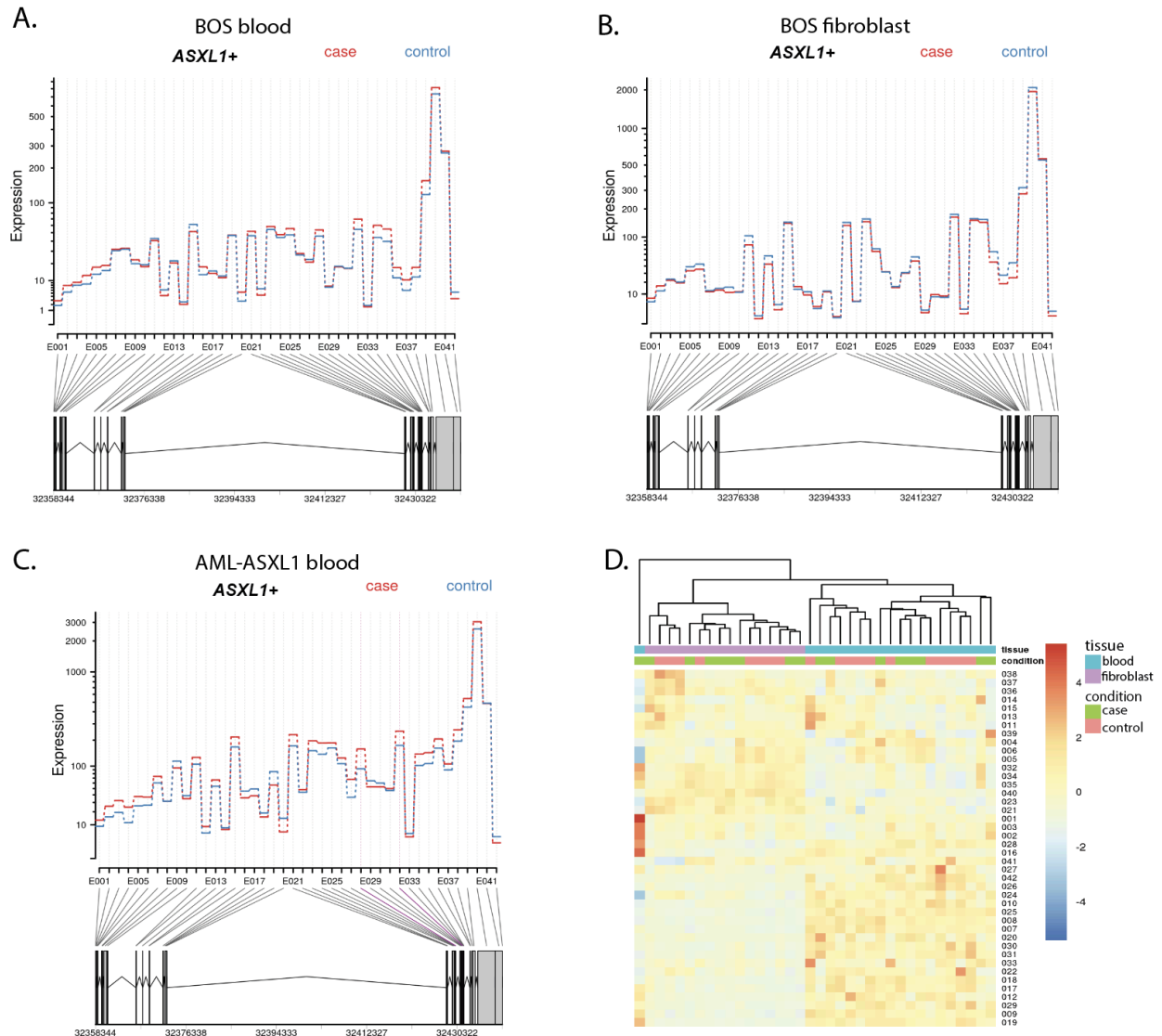


Figure 6-11: ASXL1 exon usage

ASXL1 fitted expression exon usage plots for **(A)** BOS blood, **(B)** BOS fibroblast, and **(C)** AML-ASXL1 blood.

(D) Unsupervised hierarchical cluster heatmap of ASXL1 exon bin usage in BOS blood and fibroblast samples shows clustering by tissue type and not disease status (blue = blood, purple = fibroblast, green = case, pink = control).

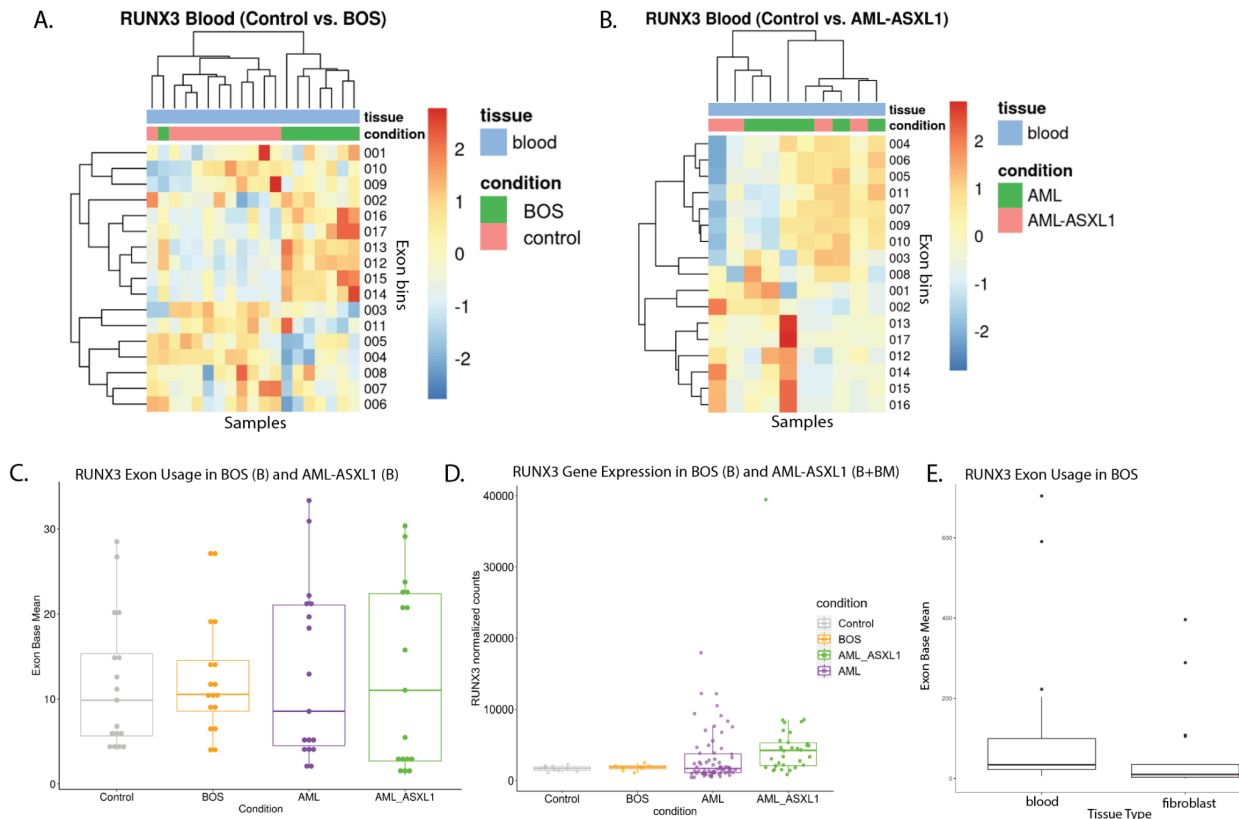


Figure 6-12: *RUNX3* exon usage and gene expression

Unsupervised hierarchical cluster heatmap of *RUNX3* exon bin usage in (A) BOS blood and (B) AML-ASXL1 blood. Cases are shown in green and controls in pink. (C) Boxplot comparing mean exon usage of *RUNX3* in Controls, BOS, AML, and AML-ASXL1 (blood samples). (D) Boxplot comparing normalized *RUNX3* gene expression in control blood, BOS blood, AML blood and bone marrow, and AML-ASXL1 blood and bone marrow. (E) Boxplot comparing mean exon usage of *RUNX3* exon bins shows *RUNX3* is more highly expressed in BOS blood than fibroblast.

Patient ID	Sample ID	Tissue	Gene	Mutation	Sex
2020	BA2006	Bone Marrow	ASXL1	G642fs*	Female
2026	BA2122	Bone Marrow	ASXL1	Q512*	Female
2026	BA2122	Bone Marrow	ASXL1	G643fs*	Female
2035	BA2212	Bone Marrow	ASXL1	G642fs*	Male
2066	BA2066	Bone Marrow	ASXL1	R693*	Female
2073	BA2510	Bone Marrow	ASXL1	W960*	Male
2075	BA2023	Bone Marrow	ASXL1	G642fs*	Male
2080	BA2089	Bone Marrow	ASXL1	G643fs*	Male
2139	BA2885	Bone Marrow	ASXL1	G643fs*	Female
2150	BA2620	Bone Marrow	ASXL1	F548*	Female
2161	BA2903	Bone Marrow	ASXL1	G642fs*	Male
2234	BA2454	Bone Marrow	ASXL1	R693*	Male
2266	BA2685	Bone Marrow	ASXL1	G642fs*	Male
2283	BA2538	Bone Marrow	ASXL1	G642fs*	Male
2285	BA2806	Bone Marrow	ASXL1	G643fs*	Male
2288	BA2681	Bone Marrow	ASXL1	G643fs*	Female
2288	BA2681	Bone Marrow	ASXL1	C789fs*	Female
2290	BA2597	Bone Marrow	ASXL1	G643fs*	Female
2327	BA2970	Bone Marrow	ASXL1	W898*	Male
2338	BA2198	Bone Marrow	ASXL1	H630fs*	Male
2385	BA2377	Bone Marrow	ASXL1	G942fs*	Male
2395	BA2798	Bone Marrow	ASXL1	Q428*	Male
2395	BA2883	Bone Marrow	ASXL1	Q428*	Male
2400	BA2520	Bone Marrow	ASXL1	R718fs*	Male
2453	BA2622	Bone Marrow	ASXL1	G642fs*	Male
2453	BA2758	Bone Marrow	ASXL1	G642fs*	Male
2478	BA3010	Bone Marrow	ASXL1	G642fs*	Male
2480	BA2489	Bone Marrow	ASXL1	G642fs*	Male
2498	BA2618	Bone Marrow	ASXL1	Y425*	Male
2547	BA2135	Bone Marrow	ASXL1	G642fs*	Female
2552	BA2694	Bone Marrow	ASXL1	E657*	Male
2576	BA2636	Bone Marrow	ASXL1	G643fs*	Male
2821	TCGA-AB-2821	Blood	ASXL1	Q733*	Male
2807	TCGA-AB-2807	Blood	ASXL1	G738Dfs*	Female
2864	TCGA-AB-2864	Blood	ASXL1	S921Tfs*	Female
2927	TCGA-AB-2927	Blood	ASXL1	S921Tfs*	Female
2917	TCGA-AB-2917	Blood	ASXL1	*574_splice	Female
2854	TCGA-AB-2854	Blood	ASXL1	Q157H	Female
2949	TCGA-AB-2949	Blood	ASXL1	Y591*	Male

Table 6-1: Demographic and clinical characteristics of AML-ASXL1 patients

This table presents the demographic and clinical data of acute myeloid leukemia patients with *ASXL1* mutations (AML-ASXL1). Sample ID, tissue type, gender, *ASXL1* mutation, and relevant clinical parameters are provided.

Patient ID	Sample ID	Tissue	Sex	Status
SRR1328407	SAMN02790970	Whole Blood	Male	Control
SRR1316015	SAMN02789629	Whole Blood	Male	Control
SRR1318241	SAMN02792305	Whole Blood	Male	Control
SRR1318815	SAMN02788608	Whole Blood	Male	Control
SRR1319772	SAMN02793372	Whole Blood	Male	Control
SRR1320137	SAMN02788845	Whole Blood	Male	Control
SRR1416560	SAMN02788246	Whole Blood	Male	Control
SRR1420558	SAMN02789249	Whole Blood	Male	Control
SRR1370742	SAMN02789404	Whole Blood	Male	Control
SRR1363742	SAMN02788427	Whole Blood	Male	Control
SRR1354665	SAMN02792742	Whole Blood	Male	Control
SRR1357103	SAMN02789616	Whole Blood	Male	Control
SRR1386967	SAMN02787737	Whole Blood	Male	Control
SRR1396191	SAMN02791308	Whole Blood	Male	Control
SRR1459913	SAMN02790182	Whole Blood	Male	Control
SRR1498313	SAMN02793351	Whole Blood	Male	Control
SRR1440421	SAMN02793436	Whole Blood	Male	Control
SRR1440852	SAMN02790059	Whole Blood	Male	Control
SRR1443459	SAMN02790476	Whole Blood	Male	Control
SRR1466556	SAMN02792569	Whole Blood	Male	Control
SRR1084577	SAMN02464958	Whole Blood	Male	Control
SRR1088168	SAMN02465325	Whole Blood	Male	Control
SRR1434610	SAMN02793229	Whole Blood	Male	Control
SRR1435560	SAMN02790844	Whole Blood	Male	Control
SRR1435833	SAMN02790634	Whole Blood	Male	Control
SRR1409510	SAMN02790391	Whole Blood	Male	Control
SRR1413892	SAMN02787702	Whole Blood	Male	Control
SRR1444852	SAMN02791605	Whole Blood	Male	Control
SRR1343501	SAMN02793494	Whole Blood	Male	Control
SRR1348114	SAMN02792329	Whole Blood	Male	Control
SRR1382615	SAMN02792084	Whole Blood	Male	Control
SRR1331684	SAMN02791813	Whole Blood	Male	Control
SRR1332007	SAMN02788335	Whole Blood	Male	Control
SRR1336706	SAMN02787240	Whole Blood	Male	Control
SRR1069024	SAMN02464834	Whole Blood	Male	Control
SRR1070813	SAMN02465620	Whole Blood	Male	Control
SRR1072345	SAMN02465058	Whole Blood	Male	Control
SRR1073679	SAMN02465841	Whole Blood	Male	Control

SRR1076847	SAMN02465593	Whole Blood	Male	Control	
SRR1080998	SAMN02465406	Whole Blood	Male	Control	
SRR1472976	SAMN02792209	Whole Blood	Male	Control	
SRR1474999	SAMN02793323	Whole Blood	Male	Control	
SRR1317898	SAMN02792605	Whole Blood	Female	Control	
SRR1320363	SAMN02793676	Whole Blood	Female	Control	
SRR1099260	SAMN02465352	Whole Blood	Female	Control	
SRR1481351	SAMN02792786	Whole Blood	Female	Control	
SRR1367499	SAMN02790782	Whole Blood	Female	Control	
SRR1402368	SAMN02789440	Whole Blood	Female	Control	
SRR1357757	SAMN02790231	Whole Blood	Female	Control	
SRR1389133	SAMN02790370	Whole Blood	Female	Control	
SRR1457574	SAMN02791656	Whole Blood	Female	Control	
SRR1341164	SAMN02790180	Whole Blood	Female	Control	
SRR1377714	SAMN02791400	Whole Blood	Female	Control	
SRR1471331	SAMN02788139	Whole Blood	Female	Control	
SRR1470798	SAMN02793810	Whole Blood	Female	Control	
SRR1414559	SAMN02787904	Whole Blood	Female	Control	
SRR1440491	SAMN02790383	Whole Blood	Female	Control	
SRR1475930	SAMN02789759	Whole Blood	Female	Control	
SRR1312187	SAMN02791764	Whole Blood	Male	Control	
SRR1457375	SAMN02791492	Whole Blood	Male	Control	
A	SAMN10127754	Bone Marrow	Female	Control	
B	SAMN10127753	Bone Marrow	Male	Control	
C	SAMN10127752	Bone Marrow	Female	Control	
H	SAMN10127751	Bone Marrow	Female	Control	
J	SAMN10127750	Bone Marrow	Female	Control	
O	SAMN10127747	Bone Marrow	Male	Control	
T	SAMN10127749	Bone Marrow	Female	Control	
U	SAMN10127748	Bone Marrow	Female	Control	
2822	TCGA-AB-2822	Blood	Male	AML	*
2826	TCGA-AB-2826	Blood	Female	AML	*
2919	TCGA-AB-2919	Blood	Female	AML	*

Table 6-2: Control sample characteristics for AML-ASXL1 analysis

Demographic information for control samples used in the AML-ASXL1 analysis. This table includes sample ID, tissue type, and sex.

Patient ID	Sample ID	Total Reads	Alt Allele Read	Freq. Across Cancer	Allele Freq	ExAC Freq	In dbSNP	In COSMIC
2020	BA2006	43	14	0.023	0.33	0.0014	TRUE	TRUE
2026	BA2122	87	49	0.0016	0.56	0.000041	TRUE	TRUE
2026	BA2122	22	2	0.0164	0.09	0.0013	TRUE	TRUE
2035	BA2212	28	11	0.023	0.39	0.0014	TRUE	TRUE
2066	BA2066	97	18	0.0066	0.19	0.000058	TRUE	TRUE
2073	BA2510	80	39	0.0016	0.49	0	FALSE	FALSE
2075	BA2023	33	9	0.023	0.27	0.0014	TRUE	TRUE
2080	BA2089	10	6	0.0016	0.60	0.000008	TRUE	FALSE
2139	BA2885	12	2	0.0164	0.17	0.0013	TRUE	TRUE
2150	BA2620	324	136	0.0016	0.42	0	FALSE	FALSE
2161	BA2903	10	5	0.023	0.50	0.0014	TRUE	TRUE
2234	BA2454	142	71	0.0066	0.50	0.000058	TRUE	TRUE
2266	BA2685	25	7	0.023	0.28	0.0014	TRUE	TRUE
2283	BA2538	32	8	0.023	0.25	0.0014	TRUE	TRUE
2285	BA2806	25	2	0.0164	0.08	0.0013	TRUE	TRUE
2288	BA2681	24	2	0.0164	0.08	0.0013	TRUE	TRUE
2288	BA2681	223	68	0.0016	0.30	0	FALSE	FALSE
2290	BA2597	24	3	0.0164	0.13	0.0013	TRUE	TRUE
2327	BA2970	454	191	0.0016	0.42	0.000017	TRUE	FALSE
2338	BA2198	57	6	0.0066	0.11	0.00014	TRUE	TRUE
2385	BA2377	231	89	0.0016	0.39	0	FALSE	FALSE
2395	BA2798	314	136	0.0033	0.43	0	FALSE	FALSE
2395	BA2883	253	107	0.0033	0.42	0	FALSE	FALSE
2400	BA2520	71	10	0.0016	0.14	0.000008	TRUE	FALSE
2453	BA2622	25	5	0.023	0.20	0.0014	TRUE	TRUE
2453	BA2758	19	4	0.023	0.21	0.0014	TRUE	TRUE
2478	BA3010	33	13	0.023	0.39	0.0014	TRUE	TRUE
2480	BA2489	48	15	0.023	0.31	0.0014	TRUE	TRUE
2498	BA2618	208	96	0.0016	0.46	0	FALSE	FALSE
2547	BA2135	43	14	0.023	0.33	0.0014	TRUE	TRUE
2552	BA2694	12	6	0.0016	0.50	0	FALSE	TRUE
2576	BA2636	16	7	0.0164	0.44	0.0013	TRUE	TRUE

Table 6-3: Allele frequency and mutation details for AML-ASXL1 samples

Detailed genetic information for AML-ASXL1 samples at the *ASXL1* mutation sites for bone marrow samples. This table includes sample ID, alternate allele read, total read, cancer allele frequency, ExAC frequency, and mutation registration in dbSNP or COSMIC databases.

Gene ID	log2FC	pvalue	padj	Gene Name	Approved Name
ENSG00000152954.12	-24.07	<1.00E-150	<1.00E-150	NRSN1	neurensin 1
ENSG00000163046.15	-21.03	<1.00E-150	<1.00E-150	ANKRD30BL	ankyrin repeat domain 30B like
ENSG00000254772.10	-17.89	8.76E-26	7.13E-24	EEF1G	eukaryotic translation elongation factor 1 gamma
ENSG00000270188.1	-16.42	1.28E-146	5.63E-143	MTRNR2L11	MT-RNR2 like 11
ENSG00000124333.16	-16.30	2.54E-42	1.39E-39	VAMP7	vesicle associated membrane protein 7
ENSG00000153266.12	-16.01	8.30E-61	1.62E-57	FEZF2	FEZ family zinc finger 2
ENSG00000281887.3	-15.68	0.00E+00	0.00E+00	GIMAP1-GIMAP5	GIMAP1-GIMAP5 readthrough
ENSG00000126550.8	-14.70	7.81E-45	5.29E-42	HTN1	histatin 1
ENSG00000244115.1	-13.92	1.67E-69	4.91E-66	DNAJC25-GNG10	DNAJC25-GNG10 readthrough
ENSG00000258724.1	-12.36	3.03E-32	5.38E-30	PINX1	PIN2 (TERF1) interacting telomerase inhibitor 1
ENSG00000145826.9	-11.76	1.17E-04	4.02E-04	LECT2	leukocyte cell derived chemotaxin 2
ENSG00000145075.13	-11.66	4.61E-34	1.05E-31	CCDC39	coiled-coil domain containing 39
ENSG00000124208.16	-11.61	1.45E-41	7.09E-39	TMEM189-UBE2V1	TMEM189-UBE2V1 readthrough
ENSG00000168530.16	-11.58	3.21E-06	1.54E-05	MYL1	myosin light chain 1
ENSG00000183206.17	-11.57	2.55E-45	1.87E-42	POTEC	POTE ankyrin domain family member C
ENSG00000261796.1	-11.50	1.38E-51	1.52E-48	ISY1-RAB43	ISY1-RAB43 readthrough
ENSG00000270181.3	-11.16	1.85E-09	1.64E-08	BIVM-ERCC5	BIVM-ERCC5 readthrough
ENSG00000154438.7	-10.73	9.26E-56	1.25E-52	ASZ1	ankyrin repeat, SAM and basic leucine zipper domain containing 1
ENSG00000102290.22	-10.64	3.99E-04	1.21E-03	PCDH11X	protocadherin 11 X-linked
ENSG00000161652.12	-10.48	5.41E-10	5.27E-09	IZUMO2	IZUMO family member 2
ENSG00000204694.11	-10.13	8.84E-48	7.41E-45	OR11A1	olfactory receptor family 11 subfamily A member 1
ENSG00000145194.18	-10.09	5.19E-23	3.06E-21	ECE2	endothelin converting enzyme 2
ENSG00000180083.10	-10.06	3.58E-45	2.52E-42	WFDC11	WAP four-disulfide core domain 11
ENSG00000274996.1	-10.05	4.70E-44	2.85E-41	AC006449.1	NA
ENSG00000185823.3	-10.00	4.06E-17	1.18E-15	NPAP1	nuclear pore associated protein 1
ENSG00000205678.7	-9.90	7.82E-24	4.90E-22	TECRL	trans-2,3-enoyl-CoA reductase like
ENSG00000241911.1	9.89	1.12E-27	1.14E-25	TRBVB	T cell receptor beta variable B (pseudogene)
ENSG00000171560.16	-9.78	2.23E-03	5.56E-03	FGA	fibrinogen alpha chain
ENSG00000131864.10	-9.45	1.94E-03	4.93E-03	USP29	ubiquitin specific peptidase 29
ENSG00000121351.7	-9.37	2.27E-03	5.65E-03	IAPP	islet amyloid polypeptide

Table 6-4: Differential gene expression in AML-ASXL1 samples compared to controls (blood + bone marrow)

This table shows the 30 differentially expressed genes (DEGs) with largest effect size from RNA sequencing comparing AML-ASXL1 samples and control samples from blood and bone marrow, adjusted for tissue and sex. The gene ID, log2 fold change, p-value, adjusted p-value, and gene name are listed.

Gene ID	Gene Name	Approved Name	log2FC_BOS	pvalue_BOS	padj_BOS	log2FC_AML_ASXL1	pvalue_AML_ASXL1	padj_AML_ASXL1
ENSG00000105737.9	GRIK5	glutamate ionotropic receptor kainate type subunit 5	3.83	5.56E-30	7.04E-26	4.69	1.06E-46	8.50E-44
ENSG00000162738.6	VANGL2	VANGL planar cell polarity protein 2	3.80	1.15E-16	2.25E-13	4.24	1.89E-32	3.43E-30
ENSG00000139910.20	NOVA1	NOVA alternative splicing regulator 1	3.46	3.06E-09	8.44E-07	3.09	1.43E-07	9.00E-07
ENSG00000173894.11	CBX2	chromobox 2	3.30	9.77E-17	2.25E-13	3.97	1.28E-23	7.90E-22
ENSG00000107447.8	DNTT	DNA nucleotidyltransferase	2.87	8.35E-08	1.42E-05	6.04	3.21E-09	2.73E-08
ENSG00000148143.13	ZNF462	zinc finger protein 462	2.55	2.87E-23	1.82E-19	1.54	1.71E-04	5.63E-04
ENSG00000147100.11	SLC16A2	solute carrier family 16 member 2	2.41	1.92E-13	1.57E-10	2.31	1.36E-07	8.60E-07
ENSG00000171811.14	CFAP46	cilia and flagella associated protein 46	2.29	1.06E-10	4.54E-08	0.87	5.22E-03	1.17E-02
ENSG00000143502.15	SUSD4	sushi domain containing 4	2.23	1.83E-13	1.55E-10	-1.76	7.05E-09	5.67E-08
ENSG00000166341.8	DCHS1	dachsous cadherin-related 1	2.19	1.53E-16	2.76E-13	2.74	9.86E-16	2.35E-14
ENSG00000171357.6	LURAP1	leucine rich adaptor protein 1	2.14	1.08E-06	1.22E-04	1.59	7.01E-04	2.00E-03
ENSG00000172867.4	KRT2	keratin 2	2.07	1.14E-06	1.27E-04	-1.78	5.63E-06	2.57E-05
ENSG00000137731.14	FXYD2	FXYD domain containing ion transport regulator 2	2.05	4.43E-12	2.61E-09	-1.04	1.30E-04	4.40E-04
ENSG00000179855.8	GIPC3	GIPC PDZ domain containing family member 3	2.05	6.42E-21	2.71E-17	2.65	1.20E-12	1.80E-11
ENSG00000160307.10	S100B	S100 calcium binding protein B	2.04	1.73E-05	1.13E-03	-1.54	1.47E-05	6.14E-05
ENSG00000124196.5	GTSF1L	gametocyte specific factor 1 like	1.85	6.78E-08	1.18E-05	-1.02	2.02E-02	3.88E-02
ENSG00000130508.1	PXDN	peroxidase	1.82	4.19E-04	1.22E-02	3.77	1.65E-14	3.24E-13
ENSG00000124615.20	MOC51	molybdenum cofactor synthesis 1	1.82	4.10E-15	4.94E-12	1.43	2.17E-06	1.08E-05
ENSG00000158966.15	CACHD1	cache domain containing 1	1.81	1.04E-12	7.10E-10	1.36	3.75E-06	1.78E-05
ENSG00000114654.7	EFC1	EF-hand and coiled-coil domain containing 1	1.78	3.33E-05	1.91E-03	5.04	1.05E-31	1.74E-29
ENSG00000167880.7	EVPL	envoplakin	1.77	1.15E-09	3.84E-07	2.81	3.60E-17	1.05E-15
ENSG00000266524.3	GDF10	growth differentiation factor 10	1.76	2.85E-06	2.66E-04	3.31	6.25E-10	5.98E-09
ENSG00000072832.14	CRMP1	collapsin response mediator protein 1	1.74	8.69E-08	1.47E-05	1.45	1.50E-05	6.25E-05
ENSG00000091129.20	NRCAM	neuronal cell adhesion molecule	1.72	2.33E-07	3.29E-05	-1.90	6.03E-08	4.09E-07
ENSG00000183010.16	PYCR1	pyrroline-5-carboxylate reductase 1	1.67	2.86E-04	9.48E-03	1.06	2.16E-03	5.43E-03
ENSG00000164867.11	NOS3	nitric oxide synthase 3	1.67	2.21E-13	1.75E-10	-0.73	7.77E-03	1.67E-02
ENSG00000072195.15	SPEG	striated muscle enriched protein kinase	1.66	1.75E-15	2.61E-12	-1.17	2.90E-07	1.72E-06
ENSG00000162337.12	LRP5	LDL receptor related protein 5	1.64	6.47E-12	3.58E-09	2.65	3.09E-24	2.01E-22
ENSG00000070018.9	LRP6	LDL receptor related protein 6	1.63	6.90E-16	1.17E-12	2.21	2.01E-12	2.89E-11
ENSG0000003096.14	KLHL13	kelch like family member 13	1.59	1.19E-09	3.93E-07	3.52	8.84E-20	3.61E-18
ENSG00000211640.4	IGLV6-57	immunoglobulin lambda variable 6-57	1.57	1.10E-04	4.89E-03	-1.50	1.54E-03	4.00E-03
ENSG00000173040.13	EVC2	EvC ciliary complex subunit 2	1.54	6.72E-10	2.46E-07	1.48	5.42E-05	2.00E-04
ENSG00000157510.14	AFA1L1	actin filament associated protein 1 like 1	1.53	1.34E-07	2.10E-05	1.37	5.70E-04	1.66E-03
ENSG00000117266.15	CDK18	cyclin dependent kinase 18	1.52	3.80E-10	1.47E-07	2.66	5.78E-31	8.84E-29
ENSG00000136002.18	ARHGEF4	Rho guanine nucleotide exchange factor 4	1.51	1.43E-18	4.52E-15	-2.25	1.17E-18	4.17E-17
ENSG00000072840.13	EVC	EvC ciliary complex subunit 1	1.50	8.47E-10	2.98E-07	2.35	3.12E-11	3.71E-10
ENSG00000169242.12	EFNA1	ephrin A1	1.45	6.97E-10	2.49E-07	1.27	1.04E-04	3.60E-04
ENSG00000240864.3	IGKV1-16	immunoglobulin kappa variable 1-16	1.44	1.63E-03	2.87E-02	-1.67	4.34E-05	1.63E-04
ENSG00000169083.17	AR	androgen receptor	1.44	1.93E-06	1.95E-04	3.74	3.06E-20	1.32E-18
ENSG00000115468.12	EFHD1	EF-hand domain family member D1	1.43	8.25E-07	9.83E-05	-1.01	5.14E-04	1.51E-03
ENSG00000102038.15	SMARCA1	SWI/SNF related, matrix assoc., actin dep. reg. of chromatin	1.43	1.62E-05	1.07E-03	1.32	3.16E-05	1.23E-04
ENSG00000135960.10	EDAR	ectodysplasin A receptor	1.42	9.74E-16	1.54E-12	-0.84	1.63E-02	3.22E-02
ENSG00000142798.20	HSPG2	heparan sulfate proteoglycan 2	1.41	6.35E-11	2.92E-08	3.48	5.76E-27	5.54E-25
ENSG00000182742.6	HOXB4	homeobox B4	1.41	1.50E-17	4.23E-14	1.91	4.22E-11	4.91E-10
ENSG00000170624.13	SGCD	sarcoglycan delta	-1.53	5.07E-04	1.38E-02	-1.81	4.27E-06	2.01E-05
ENSG00000055732.13	MUCOLN3	mucopolip 3	-1.60	4.28E-09	1.14E-06	-1.19	2.98E-03	7.19E-03
ENSG00000183762.12	KREMEN1	kringle containing transmembrane protein 1	-1.66	1.05E-04	4.77E-03	-1.39	1.83E-05	7.44E-05
ENSG00000277586.2	NEFL	neurofilament light	-1.74	7.15E-07	8.71E-05	-1.22	2.10E-03	5.27E-03
ENSG00000183439.8	TRIM61	tripartite motif containing 61	-2.08	9.26E-08	1.52E-05	-2.09	3.47E-08	2.46E-07
ENSG00000180875.5	GREM2	gremlin 2, DAN family BMP antagonist	-2.48	3.26E-08	6.02E-06	-1.67	9.96E-05	3.47E-04

Table 6-5: Comparative differentially expressed genes between AML-ASXL1 (blood + bone marrow) and BOS RNA-seq analyses

Integration of differential gene expression analyses from AML-ASXL1 blood and bone marrow samples and Bohring-Opitz syndrome (BOS) blood. This comparison displays the 50 most differentially expressed genes (DEGs), adjusted for tissue and sex, and includes gene ID, log2 fold changes for both datasets, adjusted p-values, and gene names.

GO ID	Description	GeneRatio	BgRatio	pvalue	p.adjust
GO:0042110	T cell activation	51/771	484/18654	8.80E-10	4.35E-06
GO:0030098	lymphocyte differentiation	42/771	368/18654	2.66E-09	6.58E-06
GO:0050863	regulation of T cell activation	33/771	332/18654	3.11E-06	3.85E-03
GO:0007409	axonogenesis	42/771	482/18654	4.56E-06	4.52E-03
GO:0022407	regulation of cell-cell adhesion	39/771	439/18654	6.34E-06	5.23E-03
GO:0050870	positive regulation of T cell activation	24/771	213/18654	8.52E-06	5.61E-03
GO:0002429	immune response-activating pathway	40/771	466/18654	1.10E-05	5.61E-03
GO:0002757	immune response-activating signal transduction	40/771	466/18654	1.10E-05	5.61E-03
GO:0031295	T cell costimulation	11/771	55/18654	1.27E-05	5.61E-03
GO:0009952	anterior/posterior pattern specification	24/771	220/18654	1.47E-05	5.61E-03
GO:0001822	kidney development	28/771	283/18654	1.87E-05	5.89E-03
GO:0001655	urogenital system development	31/771	330/18654	1.90E-05	5.89E-03
GO:0001654	eye development	33/771	383/18654	5.90E-05	1.39E-02
GO:0002460	adaptive immune response based on somatic recombination	31/771	357/18654	8.49E-05	1.75E-02
GO:0045580	regulation of T cell differentiation	17/771	150/18654	1.59E-04	2.71E-02
GO:0021510	spinal cord development	14/771	109/18654	1.59E-04	2.71E-02
GO:1904837	beta-catenin-TCF complex assembly	7/771	30/18654	1.77E-04	2.93E-02
GO:2000106	regulation of leukocyte apoptotic process	12/771	86/18654	2.10E-04	3.07E-02
GO:0007159	leukocyte cell-cell adhesion	30/771	364/18654	2.73E-04	3.66E-02

Table 6-6: Gene ontology analysis for common DEGs in AML-ASXL1 (blood + bone marrow) and BOS

Gene ontology (GO) analysis for the common differentially expressed genes (DEGs) identified in both AML-ASXL1 (blood + bone marrow samples) and BOS (blood) analyses, adjusted for tissue and sex. This table lists the top 20 GO terms enriched among the commonly dysregulated genes, GO ID, description, gene ratio, and adjusted p-value.

Site ID	AML_Other_Mean β	AML_ASXL1_Mean β	BOS_Mean β	Control_Mean β	Chr	Position	Gene Name	Feature
cg11673244	0.74	0.22	0.28	0.44	chr22	22,007,100	MIR130B	TSS200
cg00130808	0.92	0.81	0.53	0.66	chr15	37,178,909	LOC145845	TSS200
cg00443981	0.83	0.61	0.75	0.93	chr17	58,499,679	C17orf64	TSS200
cg00642460	0.74	0.81	0.59	0.24	chr5	177,000,000	PFN3	TSS200
cg01343041	0.74	0.86	0.71	0.51	chr2	24,397,787	C2orf84	TSS200
cg04413853	0.74	0.49	0.57	0.72	chr17	58,499,706	C17orf64	TSS200
cg04863892	0.81	0.61	0.92	0.72	chr7	27,183,375	HOXA5	TSS200
cg05650680	0.20	0.17	0.25	0.10	chr8	1,765,217	MIR596	TSS200
cg06399164	0.51	0.49	0.55	0.65	chr3	42,695,035	ZBTB47	TSS200
cg07145284	0.60	0.19	0.20	0.33	chr11	66,084,631	CD248	TSS200
cg09173348	0.73	0.59	0.36	0.49	chr17	37,322,431	ARL5C	TSS200
cg09973105	0.73	0.25	0.31	0.44	chr4	155,000,000	RNF175	TSS200
cg10347828	0.80	0.86	0.71	0.54	chr19	49,244,040	RASIP1	TSS200
cg11126134	0.07	0.44	0.51	0.27	chr13	31,480,304	C13orf33	TSS200
cg12131208	0.77	0.46	0.58	0.77	chr17	58,499,700	C17orf64	TSS200
cg12671744	0.73	0.18	0.41	0.57	chr1	46,859,774	FAAH	TSS200
cg13614409	0.71	0.83	0.85	0.66	chr13	31,506,752	C13orf26	TSS200
cg14014506	0.41	0.74	0.57	0.42	chr9	117,000,000	KIF12	TSS200
cg15649236	0.45	0.63	0.61	0.41	chr17	46,657,504	MIR10A	TSS200
cg15821939	0.13	0.07	0.30	0.11	chr8	1,765,387	MIR596	TSS200
cg16501323	0.62	0.77	0.75	0.62	chr21	45,705,618	AIRE	TSS200
cg16717549	0.79	0.93	0.94	0.70	chr21	45,705,699	AIRE	TSS200
cg17277199	0.82	0.88	0.80	0.58	chr2	24,397,845	C2orf84	TSS200
cg17356252	0.72	0.83	0.77	0.52	chr21	45,705,686	AIRE	TSS200
cg17385448	0.68	0.32	0.33	0.43	chr1	15,911,702	AGMAT	TSS200
cg17501395	0.66	0.21	0.13	0.25	chr6	150,000,000	ZC3H12D	TSS200
cg17628249	0.73	0.52	0.55	0.66	chr17	58,499,854	C17orf64	TSS200
cg18082788	0.69	0.16	0.06	0.17	chr6	150,000,000	ZC3H12D	TSS200
cg18261491	0.47	0.14	0.25	0.40	chr1	46,859,780	FAAH	TSS200
cg18587476	0.77	0.89	0.78	0.54	chr2	24,397,810	C2orf84	TSS200
cg18689454	0.79	0.92	0.88	0.51	chr21	45,705,694	AIRE	TSS200
cg19759481	0.83	0.61	0.92	0.76	chr7	27,183,401	HOXA5	TSS200
cg20359349	0.85	0.37	0.52	0.64	chr17	41,994,531	C17orf88	TSS200
cg21122199	0.64	0.40	0.45	0.57	chr17	58,499,720	C17orf64	TSS200
cg23255283	0.89	0.76	0.77	0.58	chr18	74,240,582	LOC284276	TSS200
cg23995914	0.18	0.41	0.28	0.15	chr4	10,459,228	ZNF518B	TSS200
cg10424681	0.70	0.82	0.60	0.38	chr6	4,079,350	C6orf201	TSS200
cg00123104	0.63	0.38	0.49	0.33	chr1	9,884,732	CLSTN1	TSS200
cg23451221	0.47	0.46	0.42	0.53	chr10	135,000,000	PRAP1	TSS200
cg00262344	0.88	0.75	0.81	0.67	chr18	23,713,741	PSMA8	TSS200
cg03162994	0.64	0.81	0.82	0.67	chr18	23,713,729	PSMA8	TSS200
cg06377543	0.94	0.75	0.87	0.75	chr18	23,713,657	PSMA8	TSS200
cg00705992	0.69	0.38	0.42	0.57	chr7	27,226,329	HOXA11	TSS1500
cg16038003	0.36	0.06	0.14	0.27	chr7	27,226,148	HOXA11	TSS1500
cg01572694	0.42	0.65	0.76	0.61	chr17	46,657,555	MIR10A	TSS1500
cg02850715	0.94	0.87	0.81	0.70	chr11	130,000,000	ST14	TSS1500
cg02886509	0.69	0.60	0.79	0.59	chr11	2,020,556	H19	TSS1500
cg02916332	0.81	0.62	0.88	0.75	chr7	27,183,591	HOXA5	TSS1500

cg03368099	0.78	0.55	0.87	0.63	chr7	27,184,521	HOXA5	TSS1500
cg03996735	0.79	0.93	0.93	0.79	chr11	2,020,104	H19	TSS1500
cg06911238	0.84	0.21	0.69	0.85	chr1	43,020,682	FAAH	TSS1500
cg06913600	0.70	0.40	0.50	0.63	chr2	43,020,682	HAAO	TSS1500
cg07267782	0.51	0.57	0.33	0.45	chr6	54,711,227	FAM83B	TSS1500
cg07525299	0.98	0.93	0.98	0.82	chr11	2,293,171	ASCL2	TSS1500
cg07600533	0.95	0.94	0.97	0.82	chr22	50,986,031	KLHDC7B	TSS1500
cg08745665	0.95	0.96	0.97	0.86	chr4	10,460,009	ZNF518B	TSS1500
cg08949974	0.48	0.25	0.28	0.39	chr4	40,632,860	RBM47	TSS1500
cg08950537	0.51	0.85	0.73	0.54	chr10	95,361,933	RBP4	TSS1500
cg11197905	0.60	0.63	0.70	0.54	chr4	10,459,435	ZNF518B	TSS1500
cg12499235	0.98	0.97	0.98	0.86	chr11	2,293,173	ASCL2	TSS1500
cg14014955	0.85	0.73	0.92	0.78	chr7	27,183,701	HOXA5	TSS1500
cg15474615	0.27	0.30	0.39	0.27	chr3	134,000,000	KY	TSS1500
cg17393016	0.55	0.66	0.69	0.58	chr17	79,283,390	C17orf55	TSS1500
cg17569124	0.82	0.52	0.91	0.67	chr7	27,183,643	HOXA5	TSS1500
cg19871235	0.74	0.93	0.55	0.68	chr12	71,552,569	TSPAN8	TSS1500
cg20275928	0.81	0.70	0.89	0.76	chr17	79,283,959	C17orf55	TSS1500
cg23476401	0.72	0.72	0.80	0.63	chr11	2,020,296	H19	TSS1500
cg23651356	0.61	0.28	0.33	0.45	chr6	153,000,000	RGS17	TSS1500
cg23989336	0.76	0.63	0.83	0.71	chr17	79,283,915	C17orf55	TSS1500
cg25307665	0.83	0.57	0.95	0.69	chr7	27,183,694	HOXA5	TSS1500
cg25506432	0.85	0.82	0.83	0.72	chr7	27,184,065	HOXA5	TSS1500
cg26023912	0.78	0.67	0.85	0.73	chr7	27,184,369	HOXA5	TSS1500
cg21096966	0.49	0.47	0.65	0.54	chr14	45,432,458	KLHL28	TSS1500
cg21626848	0.86	0.54	0.36	0.47	chr17	39,969,267	SC65	TSS1500
cg05017199	0.95	0.77	0.54	0.38	chr7	150,000,000	TMEM176B	TSS1500
cg17815049	0.42	0.15	0.31	0.45	chr1	247,000,000	CNST	TSS1500
cg04120329	0.66	0.28	0.42	0.53	chr10	135,000,000	PRAP1	TSS1500
cg06891458	0.63	0.73	0.53	0.41	chr15	41,951,943	MGA	TSS1500
cg07689920	0.46	0.86	0.76	0.65	chr11	73,356,316	PLEKHB1	TSS1500
cg07880854	0.51	0.47	0.73	0.53	chr2	113,000,000	FBLN7	TSS1500
cg08217285	0.81	0.27	0.29	0.41	chr17	27,917,879	GIT1	TSS1500
cg09862219	0.18	0.26	0.37	0.19	chr2	113,000,000	FBLN7	TSS1500
cg10234998	0.60	0.14	0.13	0.35	chr6	29,601,491	GABBR1	TSS1500
cg16462183	0.57	0.40	0.69	0.55	chr5	54,281,733	ESM1	TSS1500
cg21481950	0.65	0.21	0.31	0.45	chr6	29,601,498	GABBR1	TSS1500
cg25414605	0.86	0.41	0.35	0.48	chr11	67,182,672	ATPGD1	TSS1500
cg10099638	0.32	0.33	0.29	0.43	chr21	47,517,484	COL6A2	TSS1500
cg03520644	0.53	0.53	0.26	0.38	chr1	104,000,000	COL11A1	TSS1500
cg19362196	0.80	0.77	0.77	0.65	chr1	26,233,615	STMN1	TSS1500
cg18081258	0.90	0.65	0.40	0.50	chr14	21,494,161	NDRG2	TSS1500

Table 6-7: DNA methylation at transcriptional start sites (TSS)

DNA methylation (DNAm) at transcriptional start site (TSS) CpG sites, 200bp (TSS200) and 1500bp upstream (TSS1500), comparing average mean DNAm beta (β) values. Samples analyzed are from patients with AML with non-ASXL1 mutations (AML-other, n=3), AML with ASXL1 mutations (AML-ASXL1, n=3), Bohring-Opitz syndrome (BOS, n=8), and controls (n=26). The gene, CpG ID, chromosome, and position for each CpG site are provided.

Site ID	AML_Other_Mean β	AML_ASXL1_Mean β	BOS_Mean β	Control_Mean β	Chr	Position	Gene Name	Feature
cg23936031	0.87	0.66	0.94	0.73	chr7	27183133	HOXA5	1stExon
cg04014328	0.73	0.87	0.91	0.80	chr17	46653615	HOXB4	3'UTR
cg09549073	0.77	0.64	0.86	0.70	chr7	27183274	HOXA5	5'UTR
cg00705992	0.69	0.38	0.42	0.57	chr7	27226329	HOXA11	TSS1500
cg16038003	0.36	0.06	0.14	0.27	chr7	27226148	HOXA11	TSS1500
cg02916332	0.81	0.62	0.88	0.75	chr7	27183591	HOXA5	TSS1500
cg03368099	0.78	0.55	0.87	0.63	chr7	27184521	HOXA5	TSS1500
cg14014955	0.85	0.73	0.92	0.78	chr7	27183701	HOXA5	TSS1500
cg17569124	0.82	0.52	0.91	0.67	chr7	27183643	HOXA5	TSS1500
cg25307665	0.83	0.57	0.95	0.69	chr7	27183694	HOXA5	TSS1500
cg25506432	0.85	0.82	0.83	0.72	chr7	27184065	HOXA5	TSS1500
cg26023912	0.78	0.67	0.85	0.73	chr7	27184369	HOXA5	TSS1500
cg04863892	0.81	0.61	0.92	0.72	chr7	27183375	HOXA5	TSS200
cg19759481	0.83	0.61	0.92	0.76	chr7	27183401	HOXA5	TSS200

Table 6-8: DNA methylation analysis at *HOX* gene sites across different sample groups

Average mean DNA methylation (DNAm) beta (β) values at *HOX* gene sites for four sample groups: AML with non-ASXL1 mutations (AML-other, n=3), AML with ASXL1 mutations (AML-ASXL1, n=3), Bohring-Opitz syndrome (BOS, n=8), and controls (n=26).

Gene ID	Exon Bin	Gene Name	Gene Type	Chromosome	Exon Base Mean	pvalue	padj	BOS Mean	Control Mean	log2FC
ENSG00000020633	E012	RUNX3	protein_coding	chr1	25.93	2.40E-18	8.32E-13	10.31	5.88	-1.76
ENSG00000020633	E013	RUNX3	protein_coding	chr1	27.92	5.35E-16	9.27E-11	10.54	6.23	-1.66
ENSG00000102934	E005	PLLP	protein_coding	chr16	28.29	3.12E-13	2.70E-08	9.47	7.23	-0.85
ENSG00000020633	E015	RUNX3	protein_coding	chr1	22.24	4.50E-11	3.11E-06	9.49	5.66	-1.61
ENSG00000020633	E014	RUNX3	protein_coding	chr1	34.31	2.68E-10	1.55E-05	11.57	6.79	-1.71
ENSG00000073849	E019	ST6GAL1	protein_coding	chr3	27.99	1.02E-08	3.93E-04	6.90	10.33	1.28
ENSG00000132386	E013	SERPINF1	protein_coding	chr17	10.01	1.71E-08	5.40E-04	6.54	3.20	-2.13
ENSG00000142279	E010	WTIP	protein_coding	chr19	12.00	1.57E-08	5.40E-04	6.76	5.16	-0.82
ENSG00000171444	E014	MCC	protein_coding	chr5	6.69	2.03E-08	5.87E-04	5.63	3.14	-1.73
ENSG00000008323	E019	PLEKHG6	protein_coding	chr12	4.53	3.75E-08	8.06E-04	1.77	4.38	2.64
ENSG00000106003	E001	LFNG	protein_coding	chr7	4.76	3.53E-08	8.06E-04	4.92	2.23	-2.32
ENSG00000155966	E007	AFF2	protein_coding	chrX	332.57	3.96E-08	8.06E-04	18.08	26.44	1.76
ENSG00000132386	E012	SERPINF1	protein_coding	chr17	11.26	6.99E-08	1.10E-03	6.88	3.55	-1.98
ENSG00000149091	E001	DGKZ	protein_coding	chr11	10.88	7.93E-08	1.12E-03	6.81	4.11	-1.52
ENSG00000125107	E018	CNOT1	protein_coding	chr16	21.04	9.71E-08	1.25E-03	6.22	8.55	0.99
ENSG00000079691	E008	CARMIL1	protein_coding	chr6	54.76	1.70E-07	2.10E-03	10.06	13.40	0.98
ENSG00000073849	E018	ST6GAL1	protein_coding	chr3	18.51	6.74E-07	6.14E-03	5.82	8.43	1.14
ENSG00000253729	E006	PRKDC	protein_coding	chr8	79.68	8.13E-07	7.22E-03	12.51	14.27	0.47
ENSG00000150093	E043	ITGB1	protein_coding	chr10	31.66	1.27E-06	1.05E-02	7.34	10.10	1.02
ENSG00000000460	E001	C1orf112	protein_coding	chr1	117.11	1.63E-06	1.23E-02	18.35	12.83	-1.36
ENSG00000020633	E016	RUNX3	protein_coding	chr1	9.40	1.72E-06	1.24E-02	6.30	3.93	-1.41
ENSG00000077235	E006	GTF3C1	protein_coding	chr16	12.04	2.34E-06	1.59E-02	4.58	6.81	1.20
ENSG00000109452	E044	INPP4B	protein_coding	chr4	5.54	2.90E-06	1.93E-02	1.98	4.85	2.63
ENSG00000149428	E026	HYOU1	protein_coding	chr11	5.13	3.36E-06	2.20E-02	2.72	4.90	1.73
ENSG00000064933	E043	PMS1	protein_coding	chr2	5.56	5.78E-06	3.03E-02	5.13	3.06	-1.52
ENSG00000072163	E008	LIMS2	protein_coding	chr2	57.41	5.87E-06	3.03E-02	13.27	10.26	-0.88
ENSG00000099917	E001	MED15	protein_coding	chr22	12.28	5.50E-06	3.03E-02	6.98	4.98	-1.03
ENSG00000100479	E028	POLE2	protein_coding	chr14	15.26	5.79E-06	3.03E-02	5.19	7.30	1.04
ENSG00000102934	E004	PLLP	protein_coding	chr16	70.23	5.35E-06	3.03E-02	11.79	14.17	0.65
ENSG00000115459	E017	ELMOD3	protein_coding	chr2	22.13	5.56E-06	3.03E-02	6.61	8.52	0.79

Table 6-9: Significant differential exon usage (DEU) exon bins in Bohring-Opitz syndrome blood samples.

This table lists the most significant DEU exon bins identified in the comparison between Bohring-Opitz syndrome (BOS) samples and control samples. Each row represents an exon bin with detailed information on its statistical significance, including adjusted p-values and fold changes, highlighting the key exonic regions differentially used in BOS.

Gene ID	Exon Bin	Gene Name	Gene Type	Chromosome	Exon Base Mean	pvalue	padj	AML Mean	AML-ASXL1 Mean	log2FC
ENSG00000136051	E023	WASHC4	protein_coding	chr12	299.02	1.07E-03	2.02E-14	19.61	21.96	0.53
ENSG00000177565	E021	TBL1XR1	protein_coding	chr3	165.12	3.06E-03	1.66E-13	16.04	18.57	0.61
ENSG00000198231	E011	DDX42	protein_coding	chr17	337.55	1.68E-03	5.63E-12	20.56	22.97	0.54
ENSG00000067900	E035	ROCK1	protein_coding	chr18	208.50	1.24E-03	1.80E-11	17.19	19.90	0.64
ENSG00000079246	E039	XRCC5	protein_coding	chr2	674.15	1.43E-03	1.80E-11	26.99	25.02	-0.42
ENSG00000160796	E013	NBEAL2	protein_coding	chr3	707.87	9.47E-04	1.80E-11	25.64	27.54	0.41
ENSG00000169764	E073	UGP2	protein_coding	chr2	117.95	3.49E-03	1.80E-11	16.04	12.34	-0.99
ENSG00000177565	E025	TBL1XR1	protein_coding	chr3	126.18	2.48E-03	1.80E-11	14.64	16.90	0.57
ENSG00000067900	E036	ROCK1	protein_coding	chr18	261.08	2.65E-03	1.86E-11	18.47	21.46	0.69
ENSG00000152332	E009	UHMK1	protein_coding	chr1	204.43	5.92E-03	1.86E-11	17.69	19.23	0.36
ENSG00000177565	E026	TBL1XR1	protein_coding	chr3	113.74	2.38E-03	1.95E-11	14.05	16.33	0.59
ENSG00000145725	E026	PPIP5K2	protein_coding	chr5	155.13	1.39E-03	2.04E-11	15.64	18.14	0.61
ENSG00000174579	E001	MSL2	protein_coding	chr3	737.08	7.47E-03	3.35E-11	27.79	23.85	-0.84
ENSG00000143437	E002	ARNT	protein_coding	chr1	354.19	6.22E-03	6.58E-11	23.24	19.12	-0.93
ENSG00000123358	E043	NR4A1	protein_coding	chr12	59.85	4.56E-03	6.89E-11	12.85	8.72	-1.33
ENSG00000177565	E024	TBL1XR1	protein_coding	chr3	194.08	2.92E-03	7.25E-11	17.27	19.25	0.47
ENSG00000155096	E020	AZIN1	protein_coding	chr8	355.84	1.10E-03	1.89E-10	21.27	23.03	0.39
ENSG00000100201	E017	DDX17	protein_coding	chr22	2235.72	9.75E-04	2.17E-10	34.32	34.68	0.07
ENSG0000018510	E025	AGPS	protein_coding	chr2	1538.46	4.31E-03	5.15E-10	32.81	29.66	-0.66
ENSG00000130939	E046	UBE4B	protein_coding	chr1	284.78	7.80E-03	6.27E-10	22.17	17.77	-1.01
ENSG00000125107	E072	CNOT1	protein_coding	chr16	679.81	2.85E-03	7.57E-10	25.11	27.43	0.50
ENSG00000138768	E013	USO1	protein_coding	chr4	228.97	1.74E-03	9.18E-10	18.11	20.35	0.52
ENSG00000140262	E073	TCF12	protein_coding	chr15	439.99	6.92E-03	1.52E-09	24.64	20.49	-0.92
ENSG00000158290	E028	CUL4B	protein_coding	chrX	95.53	6.92E-03	2.07E-09	12.15	15.99	1.03
ENSG00000102893	E031	PHKB	protein_coding	chr16	324.73	1.21E-03	5.33E-09	20.58	22.32	0.39
ENSG00000136051	E024	WASHC4	protein_coding	chr12	281.32	4.05E-03	5.33E-09	19.14	21.76	0.60
ENSG00000163527	E008	STT3B	protein_coding	chr3	478.97	2.53E-03	6.36E-09	22.78	24.94	0.47
ENSG00000115942	E021	ORC2	protein_coding	chr2	94.95	7.66E-03	6.75E-09	12.18	16.02	1.03
ENSG00000134900	E047	TPP2	protein_coding	chr13	441.63	6.68E-03	7.09E-09	24.47	20.55	-0.87

Table 6-10: Significant differential exon usage (DEU) exon bins in AML-ASXL1 blood samples.

This table presents the most significant DEU exon bins identified in the comparison between AML-ASXL1 samples and control samples. The table provides comprehensive details on the differential usage of exon bins, including statistical significance metrics, which underscore the exonic regions with altered usage in AML-ASXL1.

Gene ID	Exon Bin	Gene Name	BOS_exonBaseMean	BOS_padj	log2FC(control/case)	AML_exonBaseMean	AML_padj	log2FC(control/case)
ENSG00000020633	E001	RUNX3	5.74	1.00E+00	0.32	3.89	1.00E+00	-0.93
ENSG00000020633	E002	RUNX3	590.27	1.00E+00	-0.01	1433.61	9.75E-03	-0.89
ENSG00000020633	E003	RUNX3	703.99	6.50E-01	0.20	1283.41	5.32E-01	-0.11
ENSG00000020633	E004	RUNX3	222.34	5.72E-01	0.27	266.86	1.64E-01	0.22
ENSG00000020633	E005	RUNX3	203.29	9.81E-01	0.20	359.73	3.38E-02	0.31
ENSG00000020633	E006	RUNX3	99.44	1.00E+00	0.20	341.57	6.44E-02	0.29
ENSG00000020633	E007	RUNX3	84.03	1.00E+00	0.22	412.75	7.13E-02	0.35
ENSG00000020633	E008	RUNX3	6.69	9.98E-01	0.45	11.55	9.44E-01	-0.06
ENSG00000020633	E009	RUNX3	60.59	1.00E+00	0.20	247.26	7.44E-02	0.58
ENSG00000020633	E010	RUNX3	45.51	1.00E+00	0.21	102.39	1.03E-01	0.75
ENSG00000020633	E011	RUNX3	32.12	8.70E-01	0.46	40.11	5.82E-02	0.85
ENSG00000020633	E012	RUNX3	25.93	8.32E-13	-1.76	5.30	1.00E+00	-1.22
ENSG00000020633	E013	RUNX3	27.92	9.27E-11	-1.66	5.97	1.00E+00	-1.15
ENSG00000020633	E014	RUNX3	34.31	1.55E-05	-1.71	6.37	1.00E+00	-2.12
ENSG00000020633	E015	RUNX3	22.24	3.11E-06	-1.61	3.08	1.00E+00	-1.66
ENSG00000020633	E016	RUNX3	9.40	1.24E-02	-1.41	1.17	1.00E+00	-2.35
ENSG00000020633	E017	RUNX3	10.53	1.18E-01	-1.43	0.87	1.00E+00	-1.13

Table 6-11: Differential exon usage in *RUNX3* across BOS and AML-ASXL1 analyses.

This table displays the base means and log2 fold changes (log2FC) for each exon bin in the *RUNX3* gene, as identified in the differential exon usage analysis comparing BOS to controls and AML-ASXL1 to controls. It includes detailed statistics for each exon bin, facilitating a comparative understanding of exon usage patterns in *RUNX3* across these conditions.

References

1. Bohring, A. *et al.* Severe end of Opitz trigonocephaly (C) syndrome or new syndrome? *Am. J. Med. Genet.* **85**, 438–446 (1999).
2. Russell, B., Tan, W.-H. & Graham, J. M., Jr. Bohring-Opitz Syndrome. in *GeneReviews®* (eds. Adam, M. P. *et al.*) (University of Washington, Seattle, Seattle (WA), 2018).
3. Russell, B. E. *et al.* Clinical findings in 39 individuals with Bohring-Opitz syndrome from a global patient-driven registry with implications for tumor surveillance and recurrence risk. *Am. J. Med. Genet. A* **191**, 1050–1058 (2023).
4. Pratcorona, M. *et al.* Acquired mutations in ASXL1 in acute myeloid leukemia: prevalence and prognostic value. *Haematologica* **97**, 388–392 (2012).
5. Genovese, G. *et al.* Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N. Engl. J. Med.* **371**, 2477–2487 (2014).
6. Gelsi-Boyer, V. *et al.* Mutations in ASXL1 are associated with poor prognosis across the spectrum of malignant myeloid diseases. *J. Hematol. Oncol.* **5**, 12 (2012).
7. Micol, J.-B. & Abdel-Wahab, O. The Role of Additional Sex Combs-Like Proteins in Cancer. *Cold Spring Harb. Perspect. Med.* **6**, (2016).
8. Jaiswal, S. *et al.* Age-related clonal hematopoiesis associated with adverse outcomes. *N. Engl. J. Med.* **371**, 2488–2498 (2014).
9. Dawoud, A. A. Z., Tapper, W. J. & Cross, N. C. P. Clonal myelopoiesis in the UK Biobank cohort: ASXL1 mutations are strongly associated with smoking. *Leukemia* **34**, 2660–2672 (2020).
10. Qi, H., Dong, C., Chung, W. K., Wang, K. & Shen, Y. Deep genetic connection between cancer and developmental disorders. *Hum. Mutat.* **37**, 1042–1050 (2016).
11. Mintz, B. Gene expression in neoplasia and differentiation. *Harvey Lect.* **71**, 193–246 (1978).

12. Bellacosa, A. Developmental disease and cancer: biological and clinical overlaps. *Am. J. Med. Genet. A* **161A**, 2788–2796 (2013).
13. Shipitsin, M. & Polyak, K. The cancer stem cell hypothesis: in search of definitions, markers, and relevance. *Lab. Invest.* **88**, 459–463 (2008).
14. Gibney, E. R. & Nolan, C. M. Epigenetics and gene expression. *Heredity* **105**, 4–13 (2010).
15. Fahrner, J. A. & Bjornsson, H. T. Mendelian disorders of the epigenetic machinery: tipping the balance of chromatin states. *Annu. Rev. Genomics Hum. Genet.* **15**, 269–293 (2014).
16. Nussinov, R., Tsai, C.-J. & Jang, H. How can same-gene mutations promote both cancer and developmental disorders? *Sci Adv* **8**, eabm2059 (2022).
17. Hoischen, A. *et al.* De novo nonsense mutations in ASXL1 cause Bohring-Opitz syndrome. *Nat. Genet.* **43**, 729–731 (2011).
18. Lin, I. *et al.* Multi-omics of Bohring-Opitz syndrome truncating ASXL1 mutations identify canonical and non-canonical Wnt signaling dysregulation. *JCI Insight* (2023) doi:10.1172/jci.insight.167744.
19. Matheus, F. *et al.* Pathological ASXL1 Mutations and Protein Variants Impair Neural Crest Development. *Stem Cell Reports* **12**, 861–868 (2019).
20. Nagase, R. *et al.* Expression of mutant Asxl1 perturbs hematopoiesis and promotes susceptibility to leukemic transformation. *J. Exp. Med.* **215**, 1729–1747 (2018).
21. Fisher, C. L., Berger, J., Randazzo, F. & Brock, H. W. A human homolog of Additional sex combs, ADDITIONAL SEX COMBS-LIKE 1, maps to chromosome 20q11. *Gene* **306**, 115–126 (2003).
22. Gross, N., Peñagaricano, F. & Khatib, H. Integration of whole-genome DNA methylation data with RNA sequencing data to identify markers for bull fertility. *Anim. Genet.* **51**, 502–510 (2020).
23. Zhu, W. *et al.* Integrated single-cell RNA-seq and DNA methylation reveal the effects of air pollution in patients with recurrent spontaneous abortion. *Clin. Epigenetics* **14**, 105 (2022).

24. Cappelli, E., Felici, G. & Weitschek, E. Combining DNA methylation and RNA sequencing data of cancer for supervised knowledge extraction. *BioData Min.* **11**, 22 (2018).
25. The Cancer Genome Atlas Research Network, John N Weinstein, Eric A Collisson, Gordon B Mills, Kenna R Mills Shaw, Brad A Ozenberger, Kyle Ellrott, Ilya Shmulevich, Chris Sander & Joshua M Stuart. The cancer genome atlas pan-cancer analysis project. *Nat. Genet.* **45**, 1113–1120 (2013).
26. Jensen, M. A., Ferretti, V., Grossman, R. L. & Staudt, L. M. The NCI Genomic Data Commons as an engine for precision medicine. *Blood* **130**, 453–459 (2017).
27. Tyner, J. W. *et al.* Functional genomic landscape of acute myeloid leukaemia. *Nature* **562**, 526–531 (2018).
28. Oetjen, K. A. *et al.* Human bone marrow assessment by single-cell RNA sequencing, mass cytometry, and flow cytometry. *JCI Insight* **3**, (2018).
29. Awamleh, Z. *et al.* DNA methylation signature associated with Bohring-Opitz syndrome: a new tool for functional classification of variants in ASXL genes. *Eur. J. Hum. Genet.* (2022) doi:10.1038/s41431-022-01083-0.
30. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
31. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
32. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
33. Du, P. *et al.* Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics* **11**, 587 (2010).
34. Wagner, G. P., Kin, K. & Lynch, V. J. Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. *Theory Biosci.* **131**, 281–285 (2012).
35. Anders, S., Reyes, A. & Huber, W. Detecting differential usage of exons from RNA-seq data. *Genome Res.* **22**, 2008–2017 (2012).

36. Newman, A. M. *et al.* Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat. Biotechnol.* **37**, 773–782 (2019).
37. Steen, C. B., Liu, C. L., Alizadeh, A. A. & Newman, A. M. Profiling Cell Type Abundance and Expression in Bulk Tissues with CIBERSORTx. *Methods Mol. Biol.* **2117**, 135–157 (2020).
38. Ashburner, M. *et al.* Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **25**, 25–29 (2000).
39. Gene Ontology Consortium *et al.* The Gene Ontology knowledgebase in 2023. *Genetics* **224**, (2023).
40. Galea, G. L. *et al.* Cell non-autonomy amplifies disruption of neurulation by mosaic Vangl2 deletion in mice. *Nat. Commun.* **12**, 1159 (2021).
41. Chew, L. J. *et al.* Characterization of the rat GRIK5 kainate receptor subunit gene promoter and its intragenic regions involved in neural cell specificity. *J. Biol. Chem.* **276**, 42162–42171 (2001).
42. Tanwar, V. *et al.* Gremlin 2 promotes differentiation of embryonic stem cells to atrial fate by activation of the JNK signaling pathway. *Stem Cells* **32**, 1774–1788 (2014).
43. Metzeler, K. H. *et al.* ASXL1 mutations identify a high-risk subgroup of older patients with primary cytogenetically normal AML within the ELN Favorable genetic category. *Blood* **118**, 6920–6929 (2011).
44. Cancer Genome Atlas Research Network *et al.* The Cancer Genome Atlas Pan-Cancer analysis project. *Nat. Genet.* **45**, 1113–1120 (2013).
45. Giampaolo, A. *et al.* Key functional role and lineage-specific expression of selected HOXB genes in purified hematopoietic progenitor differentiation. *Blood* **84**, 3637–3647 (1994).
46. Kawagoe, H., Humphries, R. K., Blair, A., Sutherland, H. J. & Hogge, D. E. Expression of HOX genes, HOX cofactors, and MLL in phenotypically and functionally defined subpopulations of leukemic and normal human hematopoietic cells. *Leukemia* **13**, 687–698

(1999).

47. Martinez, P. & Amemiya, C. T. Genomics of the HOX gene cluster. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **133**, 571–580 (2002).
48. McGuire, M. H. *et al.* Pan-cancer genomic analysis links 3'UTR DNA methylation with increased gene expression in T cells. *EBioMedicine* **43**, 127–137 (2019).
49. Björnsson, J. M. *et al.* Reduced proliferative capacity of hematopoietic stem cells deficient in Hoxb3 and Hoxb4. *Mol. Cell. Biol.* **23**, 3872–3883 (2003).
50. Schiedlmeier, B. *et al.* HOXB4's road map to stem cell expansion. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 16952–16957 (2007).
51. Shulgin, A. A., Lebedev, T. D., Prassolov, V. S. & Spirin, P. V. Plasmolipin and Its Role in Cell Processes. *Mol. Biol.* **55**, 773–785 (2021).
52. O'Neill, A. C. *et al.* A Primate-Specific Isoform of PLEKHG6 Regulates Neurogenesis and Neuronal Migration. *Cell Rep.* **25**, 2729–2741.e6 (2018).
53. Bangsow, C. *et al.* The RUNX3 gene--sequence, structure and regulated expression. *Gene* **279**, 221–232 (2001).
54. Abdel-Wahab, O. *et al.* ASXL1 mutations promote myeloid transformation through loss of PRC2-mediated gene repression. *Cancer Cell* **22**, 180–193 (2012).
55. Inoue, K.-I., Shiga, T. & Ito, Y. Runx transcription factors in neuronal development. *Neural Dev.* **3**, 20 (2008).
56. Zhang, W. *et al.* Runt-Related Transcription Factor 3 Promotes Acute Myeloid Leukemia Progression. *Front. Oncol.* **11**, 725336 (2021).

CHAPTER 7

Bloom Syndrome and Carriers: Increased Frequency of Somatic Mutations in Clonal Hematopoiesis of Indeterminate Potential

Abstract

Bloom Syndrome (BSyn) is an autosomal recessive disorder caused by biallelic germline variants in *BLM*, which functions to maintain genomic stability. BSyn patients have poor growth, immune defects, insulin resistance, and a significantly increased risk of malignancies, most commonly hematologic. The malignancy risk in carriers of pathogenic variants in *BLM* (*BLM* variant carriers) remains understudied. Clonal hematopoiesis of indeterminate potential (CHIP) is defined by presence of somatic mutations in leukemia-related genes in blood of individuals without leukemia and is associated with increased risk of leukemia. We hypothesize that somatic mutations driving clonal expansion may be an underlying mechanism leading to increased cancer risk in BSyn patients and *BLM* variant carriers.

To determine whether *de novo* or somatic variation is increased in BSyn patients or carriers, we performed and analyzed exome sequencing on BSyn and control trios.

We discovered that both BSyn patients and carriers had increased numbers of low-frequency, putative somatic variants in CHIP genes compared to controls. Furthermore, *BLM* variant carriers had increased numbers of somatic variants in DNA methylation genes compared to controls. There was no statistical difference in the numbers of *de novo* variants in BSyn probands compared to control probands.

Our findings of increased CHIP in BSyn probands and carriers suggest that one or two germline pathogenic variants in *BLM* could be sufficient to increase the risk of clonal hematopoiesis. These findings warrant further studies in larger cohorts to determine the significance of CHIP as a potential biomarker of aging, cancer, cardiovascular disease, morbidity and mortality.

Introduction

Bloom syndrome (BSyn, OMIM #210900) is a rare autosomal recessive disorder characterized by prenatal and postnatal growth restriction, sun sensitivity, insulin resistance, mild immune deficiency, and increased risk of early-onset malignancy^{1–3}. Predominantly, BSyn cases are caused by homozygous or compound heterozygous variants in the *BLM* gene, with over 70 different pathogenic variants identified, including several founder variants in Ashkenazi, Slavic, Spanish and Portuguese populations^{4,5}.

The BLM protein, a 3'→5' ATP-dependent RecQ DNA helicase, plays crucial roles in genome maintenance, including the stabilization of replication forks³, suppression of DNA crossovers during meiosis⁶, disentanglement of under-replicated DNA strands during metaphase, and management of complex DNA structures such as G complexes⁷. Notably, cells from individuals with BSyn exhibit an excess number of sister-chromatid exchanges (SCE), historically considered the diagnostic test for BSyn.^{8–11} However, increased SCE has not proven predictive of cancer development and there are currently no evidence-based cancer surveillance guidelines and biomarkers for early cancer detection^{8,12}.

The heightened incidence of early-onset cancer in BSyn patients is thought to be driven by genomic instability resulting from biallelic germline variants in *BLM*^{13–16}. For example, a recent BSyn patient with T-cell anaplastic large-cell lymphoma showed an excess of inter- and intra-chromosomal structural rearrangements, 150 and 1,293 respectively, in tumor cells¹⁷.

The Bloom Syndrome Registry reported that 53% of participants had developed cancer, with hematologic malignancies being the most common¹⁶. The most common solid tumors included colon, breast, and oropharyngeal cancers. The well-established cancer risk in BSyn individuals contrasts with the conflicting literature surrounding cancer risk in carriers of pathogenic variants in *BLM* (*BLM* variant carriers)^{5,18–23}. A study conducted in New York and Israeli Ashkenazi Jews identified an increased risk of colorectal cancer in *BLM* variant carriers,

with an odds ratio of 2.34²⁴, and an Australian study in a high-risk breast cancer cohort identified a heterozygous, deleterious, protein truncating variant in *BLM* that co-segregated with cancer in a family¹⁹. A Netherlands cohort of early-onset colorectal cancer reported an increased frequency of heterozygous loss-of-function *BLM* variants¹⁸. Moreover, *in vitro* and *in vivo* experiments found that heterozygous *BLM* mutations increased susceptibility to mesothelioma²⁵.

Clonal hematopoiesis of indeterminate potential (CHIP) is characterized by somatic mutations in leukemia-related “CHIP” genes detected in the blood of individuals without apparent hematologic malignancy^{26,27}. The definition of CHIP has been a subject of ongoing discussion, ranging from a benign precursor state for hematologic neoplasms²⁸, to the presence of a clonally expanded hematopoietic stem cell with a mutation with leukemogenic potential²⁹. CHIP significantly elevates hematologic malignancy risk, with an annual increased risk of leukemia ranging from 0.5% to 1.0%²⁶. Furthermore, CHIP has been implicated in adverse outcomes, such as cardiovascular disease, inflammation, and reduced age-adjusted life expectancy^{30–32}.

The prevalence of CHIP varies with age, with up to 10% of patients aged 60-70 showing evidence of CHIP, while only 1% of those under the age of 50 seem to be affected. This age dependent pattern is attributed to the cumulative effects of mutations and telomere shortening in haematopoietic stem cells and multipotent progenitors over time, resulting in selective fitness and a profound decrease in clonal diversity^{26,32–34}.

Studies suggest increased risk of CHIP in individuals harboring germline mutations in DNA repair and telomere maintenance genes^{33,35,36}. In a study on heterozygous carriers of *POT1*, a gene involved in telomerase-dependent telomere elongation, predisposition to clonal hematopoiesis was observed in an autosomal dominant pattern of inheritance. This increased risk demonstrated age-dependent penetrance, with the hypothesis that longer telomere length allowed clones to sustain clonality³⁷. Patients with short telomere syndrome have also been found to have an increased incidence of CHIP, although there is a wide spectrum of mutation

types and genes in which these mutations occur^{38,39}. This suggests that germline genetics can create a “permissive” environment for clonal evolution³⁸, leading to clonal selection in hematopoietic cells. Furthermore, in a recent large-scale study of UK Biobank, Memorial Sloan Kettering IMPACT, and The Cancer Genome Atlas cohorts, *BLM* was identified as one of 18 clonal hematopoiesis genes also associated with hematopoietic malignancy in the heterozygous state⁴⁰.

Although clonal cytogenetic abnormalities in Bloom Syndrome fibroblast cell lines hint at a neoplastic process, no comprehensive genomic analyses have assessed overall somatic mutational load or CHIP as potential predictive biomarkers for hematologic malignancy^{41,42}. We hypothesize that 1 or 2 germline *BLM* variants may heighten CHIP risk, potentially at earlier ages, correlating with increased malignancy risk identified in BSyn individuals and mildly elevated risk in *BLM* variant carriers. This study used exome sequencing to investigate the correlation between germline *BLM* variants, germline mutational load, and the risk of developing CHIP, potentially at earlier ages. To determine whether *de novo* or somatic variation occurs at rates in BSyn patients and *BLM* variant carriers different from control populations, we conducted exome analysis and variant annotation for these cohorts.

No significant differences in *de novo* mutation were identified in BSyn probands compared to controls; however, both BSyn probands and *BLM* variant carriers exhibited an increased incidence of low frequency, putatively somatic variants in CHIP and DNA methylation (DNAm) genes compared to sex- and age-matched controls. This study sheds new light on the interplay between genetic predispositions and somatic variation, and highlights the need for additional studies to further evaluate the mechanisms and potential clinical implications for patients with one or two *BLM* variants.

Methods

Study participants and ethical approval

All study participants provided informed consent under a protocol for the Bloom Syndrome Registry approved by the Weill Cornell Medical College Institutional Review Board, and a material transfer agreement was obtained. Bloom Syndrome trios in this cohort are described in **Table 7-1**.

Exome sequencing

We performed exome sequencing with the Nextera DNA Flex Pre-Enrichment Library Prep and the Roche NimbleGen exome capture kit following standard protocols. Libraries were indexed, multiplexed and sequenced on a 2x150 Illumina NovaSeq S1 flowcell at the UCLA Technology Center for Genomics and Bioinformatics.

Age- and sex-matched control trios were obtained from the publicly available dbGAP study phs001272.v1.p1 deposited by the Broad Institute Center for Mendelian Genomics. Control trios harbored undiagnosed disease without cancer phenotypes (**Table 7-2**) were processed with the Illumina Nextera Exome Kit and sequenced on an Illumina HiSeq.

Data processing, variant calling, and annotation

All FASTQ files underwent unified quality control, mapping and variant-calling based on GATK best-practices pipeline^{43,44}. Given the reliable identification of high-confidence somatic SNV/indel calls at positions with sufficient sequencing coverage >10X, variant calls were filtered to maintain read depth (DP)>10 over the alternate allele.

Family structure and variants were confirmed with *BLM* gene variant visualization using *Integrative Genomics Viewer v2.9.4* and variant annotation using *VarSeq 2.3.0*.

Variant filtering

Variants were filtered for DP and analyzed for coverage across regions of interest. Subsequent filtering for genotype quality (GQ) and a quality score of “PASS” were included.

Variant allele frequency (VAF) analysis

De novo variants were subsetted based on standard somatic and germline variant allele frequencies (VAF). VAF, representing the percentage of sequencing reads matching a specific DNA variant⁴⁵, was used as a surrogate measure of allele proportion. A VAF<0.3 indicated acquired somatic variants, while VAF≥0.3 indicated likely germline or *de novo* variants^{46–49}.

Given the range of VAF cutoffs in literature, analyses were also conducted using a more stringent VAF cutoff of 0.25, aligning with estimates of somatic singleton mutations in the range $0.1 \leq \text{VAF} \leq 0.25$ from previous studies⁵⁰.

Public data sets.

Control Exome data was obtained from Study ID: phs000178.v11.p8.c1, submitted by the Center for Mendelian Genomics (CMG) - The Broad Institute Joint Center for Mendelian Genomics. The Exomes used are: SRA ID SRS2136666, SRS2813808, SRS2136486, SRS2140039, SRS2140061, SRS2136721, SRS2203482, SRS2202906, SRS2202907, SRS2130875, SRS2136628, SRS2130876, SRS2197363, SRS2197826, SRS2197795, SRS2140305, SRS2137393, SRS2137389, SRS2200570, SRS2200550, SRS2200596, SRS2195820, SRS2195786, SRS2195798, SRS2200588, SRS2200627, SRS2200615, SRS2205811, SRS2195821, SRS2195834, SRS2136679, SRS2136619, SRS2136617, SRS2136629, SRS2136659, SRS2136613, SRS2203316, SRS2202950, SRS2202953, SRS2203490, SRS2203471, SRS2203336, SRS2200551, SRS2200573, SRS2200609, SRS2200624, SRS2200562, SRS2200610, SRS2200576, SRS2200561, SRS2288808, SRS2288810, SRS2288816, SRS2200626, SRS2200613, SRS2288805, SRS2200605.

Results

Trio exome analysis for BSyn and control trios

We performed exome sequencing on 29 peripheral blood DNA samples obtained from the Bloom Syndrome Registry. The cohort consisted of 10 BSyn probands and their biological parents who are obligate carriers of these pathogenic variants in *BLM* (**Table 7-1**). Among the BSyn probands, there were equal numbers of male (n=5) and female (n=5), with ages ranging from 10 months to 36 years of age at the time of sample collection. Half (n=5) of the BSyn probands had a history of developing at least one type of cancer before or after the sample was collected (**Table 7-1**).

The pathogenic *BLM* variants observed in BSyn trios spanned amino acid 25 to 1243 of the *BLM* gene, with most variants clustered in the DEAH Helicase and RecQ Helicase C-terminal domains (**Figure 7-1A**). Family structure and *BLM* variants were confirmed with *Integrative Genomics Viewer v2.9.4* (**Figure 7-1B**) and detailed variant annotation (**Table 7-1**).

No significant difference in mean total reads was identified between BSyn and control trio samples (t-test, p -value=0.268). After mapping to exome targets, BSyn trio samples had a mean of 113.1x coverage compared to controls with 106.5x coverage (t-test, p -value=0.018). This coverage consistency extended across all chromosomes.

Samples were categorized into four key groups: BSyn proband samples (n=10) designated as “affected”, *BLM* variant carrier samples as “carrier” (n=19), and control proband (n=19) and control parents (n=38) as “unaffected.” In our somatic variant analysis, we separately considered control parent and control children as age-matched controls to assess the incidence of CHIP.

BSyn probands and *BLM* variant carriers harbor somatic mutations in CHIP genes

We investigated the occurrence of low frequency, putative somatic variants (VAF < 0.3) in CHIP genes among BSyn probands and *BLM* variant carriers, revealing a median of 2,

contrasting with control cohorts where no somatic CHIP gene variants were detected (median = 0) (Kruskal Wallis, p -value=1.50E-06 to 6.37E-03) (**Figure 7-2A, Table 7-3**). Applying a more stringent VAF cutoff of ≤ 0.25 , based on differing cutoffs for somatic variants in literature, still identified more variants in CHIP genes in BSyn probands and *BLM* variant carriers compared to control cohorts, and notably significant comparisons between *BLM* variant carriers and control cohorts (Kruskal Wallis, p -value=8.82E-06 to 7.37E-03).

We explored multiple read depth cut-offs for loci in CHIP genes . While DP>0 and DP>10 actually identified more loci in control samples than BSyn samples, DP>50 showed no significant difference in pairwise coverage comparison for total loci covered in CHIP genes. A heatmap plotting all loci called for each sample at DP>0 confirmed nearly complete overlap across all samples.

We further categorized variants in CHIP genes into putative somatic or germline based on VAF (**Figure 7-2B**). Consistently, significant differences were observed across all likely somatic variant comparisons between BSyn groups (model mean = 3.70 - 4.80%) and control groups (model mean = 0.30%) (p -value=1.41E-06 to 1.60E-03) . No significant differences were found in the mean proportion of germline and somatic variants between BSyn probands and *BLM* variant carriers (p -value=0.447), nor between control probands and control parents (p -value=0.991) (**Figure 7-2B**).

Our analysis identified no significant correlations between mean somatic and germline variants in CHIP genes and the putative somatic subset. Across the four sample groups, we identified no significant difference between mean somatic and germline number of variants in CHIP genes (**Figure 7-2C, Table 7-3**), type of variant Refseq Genes 110, NCBI), pathogenicity (ClinVar 2023-01-05, NCBI), or CHIP genes to which these variants mapped (Refseq Genes 110, NCBI). One-way ANOVA with random family effect confirmed that total CHIP gene variants followed a normal distribution. No significant differences were observed in the type of variant, variant pathogenicity , or CHIP gene distribution.

We also explored the influence of age on the number of putative somatic variants in CHIP genes in each cohort (**Figure 7-2D**). Linear regression analysis identified very weak linear relationships between age and the frequency of putative somatic CHIP variants at VAF<0.3 in our cohorts ($R^2 = 1.111\text{E-}05$ to 0.009).

The increased presence of low frequency, putative somatic variants in CHIP genes for BSyn probands and *BLM* variant carriers is not attributable to differential coverage at CHIP genes but is likely due to increased somatic variant load in key cancer-related genes such as CHIP genes.

BSyn probands and *BLM* variant carriers harbor somatic variants in DNA methylation genes

Epigenetic alterations have been linked to aging and cancer, and recent findings have found an accelerated aging, DNA methylation episinature in patients and mice with biallelic variants in *BLM*^{51–54}. Furthermore, loss of a different RecQ helicase protein, WRN, results in accelerated rates of DNA damage and increased DNAm aging in patients with Werner Syndrome who exhibit many clinical signs of accelerated aging^{55,56}. To explore the impact of *BLM* mutations on mutational load in DNAm genes, we conducted a comprehensive analysis^{57,58}.

We identified significantly more low frequency, putatively somatic DNAm gene variants (VAF < 0.3) in BSyn probands and *BLM* variant carriers compared to unaffected control cohorts (**Figure 7-3A**). *BLM* variant cohorts exhibited a median of 1 somatic DNAm gene variant, compared to a median of 0 in BSyn proband, control children and control parent cohorts. Statistically, these differences were significant in comparisons between *BLM* variant carrier and control children cohorts (Kruskal Wallis, $p\text{-value}=7.941\text{E-}04$) and *BLM* variant carrier and control parent cohorts (Kruskal Wallis, $p\text{-value}=8.815\text{E-}04$) (**Figure 7-3A, Table 7-4**).

Applying a more stringent VAF cutoff of ≤ 0.25 still identified more variants in DNAm genes in *BLM* variant carriers compared to control cohorts, with significant differences between *BLM* variant carriers and the age-matched control cohort (Kruskal Wallis, $p\text{-value}=0.044$).

We assessed multiple read depth cut-offs for loci in DNAm genes. While $DP > 0$ and $DP > 10$ identified more loci covered in control samples than BSyn samples, $DP > 50$ showed no significant difference in pairwise coverage comparison for total loci in DNAm genes. Loci coverage analysis identified almost complete overlap.

Total variant analysis in DNAm genes was further explored, categorizing variants into putative somatic or germline based on VAF (**Figure 7-3B**). We identified significant differences in the mean proportions of variants, with higher proportions of somatic DNAm gene variants in BSyn proband ($4.3\% \pm 2.3\%$) and *BLM* variant carrier ($5.7\% \pm 1.9\%$) cohorts compared to control parent ($0.0\% \pm 0.0\%$) cohorts (**Figure 7-3B, Table 7-4**). Significant differences were noted between *BLM* variant carriers and control children (mean proportion, p -value= $2.944E-03$), and *BLM* variant carriers and control parents (mean proportion, p -value= $2.944E-03$).

No significant differences were found in the mean number of total variants in DNAm genes using one way ANOVA with random family effect (**Figure 7-3C**). Similarly, in our mean proportion analysis, no significant differences were observed based on the types of variants or DNAm gene in which the variant occurred. Additionally, the effect of age on the frequency of putative somatic variants in DNAm genes was assessed across cohorts (**Figure 7-3D**). Linear regression analysis identified very weak to no linear relationships between age and the frequency of putative somatic DNAm variants at $VAF < 0.3$ in our cohorts ($R^2 = 0.001$ to NA). These findings suggest that the increased presence of low frequency, putatively somatic DNAm gene variants in BSyn probands and *BLM* variant carriers is likely due to presence of the *BLM* variant and not age or family-effects.

BSyn is not associated with increased *de novo* mutation rates

Finally, we assessed whether *BLM* mutations have differential effects on germline or somatic mutation rates, which are driven by different molecular mechanisms. One proband (B380) was excluded from this analysis as sequencing from only one parent was available. High

quality coding variants in probands of each trio (n=9) that were not inherited from either parent were identified, representing *de novo* variants (DNVs, VAF $\geq$ 0.3) and newly arising somatic variants (VAF<0.3) (**Figure 7-4A**). No significant difference was found between BSyn and control cohorts (t-test, p -value=0.124).

Mean proportions of coding variants were assessed based on VAF categories, revealing no significant differences (**Figure 7-4B**). The number and type of coding DNVs (VAF $\geq$ 0.3) showed no significant difference between BSyn probands and control children (**Figure 7-4C**, t-test, p -value=0.628), aligning with reported literature values⁵⁹. Similarly, putative somatic variants (VAF<0.3) showed no significant differences between BSyn probands and control children (**Figure 7-4D**, t-test, p -value=0.083). However, the mean number of somatic variants for BSyn probands was 20.78, compared to a mean number of 16.33 in control children, suggesting a potential trend that a larger sample size would clarify.

We examined whether the *de novo* mutation rate was influenced by parental age at conception. An inverse relationship (slope=-0.961) between paternal age at conception and number of DNVs in the BSyn probands was observed, however this correlation was not robust ($R^2=0.355$). This likely stems from a small sample size and a narrow range of paternal age at proband conception (19-32 years of age). No significant difference in DNVs was identified between BSyn probands with or without a cancer diagnosis. No significant trend was observed between the number of DNVs in BSyn probands with maternal age at conception ($R^2=0.294$, data not shown) or with BSyn proband age at collection ($R^2=0.039$).

Discussion

This study addresses the impact of pathogenic *BLM* variants on the incidence of *de novo* variants and somatic variants. Our findings reveal an increased frequency of low frequency, putatively somatic variants in CHIP (**Figure 7-2A**) and DNAm genes (**Figure 7-3A**) in BSyn probands and *BLM* variant carriers, compared to sex- and age-matched controls. These

variants were predominantly represented by synonymous variants in BSyn probands and splice variants in *BLM* variant carriers. In contrast to prior studies on clonal hematopoiesis, which identified predicted-pathogenic somatic variants predominantly in *DNMT3A*, *ASXL1*, and *TET2*, we identified mainly synonymous and benign splice variants primarily in *NOTCH1* and *CUX1*

26,60,61

The absence of significant differences in mean somatic and germline variants between BSyn probands and *BLM* variant carriers suggests other factors besides *BLM* mutations, such as environmental exposures or other genetic modifiers, may influence mutation patterns. Small sample size limits statistical significance. For example, although the mean number of exome-wide somatic variants for BSyn probands was higher than in control children, it did not attain statistical significance.

The findings of increased low-frequency putative somatic variants in CHIP genes aligns with the established roles of CHIP genes in cancer predisposition. This suggests a potential link between *BLM* mutations and increased somatic mutation load in cancer-related genes. An oral presentation at the American Society of Hematology 2023 Meeting identified *BLM* as one of 18 clonal hematopoiesis genes associated with hematopoietic malignancy in the heterozygous state⁴⁰, consistent with our findings.

This study suggests reduced function of *BLM* may lead to genomic instability, resulting in increased somatic mutations. These identified variants, largely synonymous and benign splice variants, may be a consequence of cell fitness selection, as they are less likely to be functional and more neutral to cell fitness and selection bias. Functional (deleterious) mutations would likely be selected against even in cells with deleterious *BLM* variants, while variants that promote growth and maintenance pathways may experience clonal expansion within the bone marrow. This expansion could contribute to a stable and functional hematopoietic system despite the underlying genetic alterations. Furthermore, cellular mechanisms involved in maintaining genomic stability, such as alternative DNA repair pathways or enhanced

surveillance against accumulation of deleterious genetic alterations, could act as protective factors, potentially compensating for the deficiency caused by *BLM* mutations.

Our findings underscore the potential elevated risk for malignancies in *BLM* variant carriers, as evidenced by 63% (12/19) of *BLM* variant carriers exhibiting 2 or more somatic CHIP variants, a significant association compared to controls. However, conflicting evidence on increased cancer risk in *BLM* variant carriers warrants further investigation^{20,23,62–65}. Deep-amplicon sequencing is recommended to validate variants observed through exome sequencing.

This study further explores the increased risk of CHIP in patients with germline mutations in genes involved in DNA repair and telomere maintenance genes. The *BLM* gene is a DNA helicase involved in replication fork stabilization and the suppression of homologous recombination, which could influence genomic stability, potentially contributing to the observed increase of somatic variants in CHIP-associated genes. This is consistent with existing evidence linking germline mutations in DNA repair and telomere maintenance genes and an increased risk of clonal hematopoiesis^{33,35,36,38,39}.

Epigenome dysregulation is a hallmark of many diseases, including hematologic malignancies^{26,61,66,67}. DNAm genes are known regulators of the dynamic epigenome⁶⁸. We identified significantly increased numbers of somatic variants in DNAm genes in BSyn probands and *BLM* variant carriers, suggesting potential accelerated aging and methylation changes as observed recently by Lee et al.⁵⁴, and similar to findings in patients with mutations in *WRN*⁶⁹. Unfortunately, insufficient DNA precluded DNAm assays. Our findings advocate for future studies and suggest that individuals with *BLM* mutations may benefit from sequencing-based screenings, which include many genes involved in DNAm processes^{26,61,66,67}, for early detection of hematological malignancies.

Contrary to expectations, no significant difference in *de novo* mutation rates between BSyn probands and control cohorts were observed. However, the increased mean number of

somatic variants for BSyn probands implies that a larger sample size may provide more clarity on *de novo* mutation rates.

The study's limitations include the rarity of Bloom Syndrome, resulting in a small sample size that constrains our ability to detect significant differences in *de novo* mutation rates and assess the potential impact of cancer therapy on the prevalence of clonal hematopoiesis. Although cancer treatments conceivably shape the landscape of clonal hematopoiesis ⁷⁰, our study did not find significant correlations between prior cancer diagnosis, including treatments administered before sample collection, and the number of somatic variants in BSyn patients (**Figure 7-4D**). An interesting follow-up would be to examine temporal samples, enabling monitoring of age-related expansions of potential somatic mutations and clarification of whether these represent true clonal somatic mosaicism. Such analyses could enhance statistical power in evaluating variant load in BSyn patients, and distinguishing those with and without cancer diagnoses.

The use of two different exome enrichment methods is another limitation, mitigated through joint computational processing. Additionally, the use of exome sequencing poses challenges in detecting ultra-low-frequency clones associated with CHIP variants, and the nature of exome sequencing with enrichment may lead to biases in coverage and detection of somatic variants. Our findings point to the potential utility of targeted assay for CHIP variants to address these ultra-low-frequency variants.

Finally, the association between true CHIP and cancer risk, as observed in various studies, may be influenced by factors such as smoking ^{26,61,71}. Some studies indicate a strong link between CHIP and a history of smoking, which may be a confounding effect on association studies. Although smoking history is unknown in our BSyn cohort, it is reasonable to assume that the BSyn probands did not have significant pack-year history given their ages.

The complex interplay between BLM mutations and the mutational landscape, revealing associations with somatic mutations in cancer-related genes and DNA methylation alterations,

underscores the need for further investigation. Our findings contribute to the growing literature on the heightened somatic mutation rate and increased cancer risk in carriers for genes important in maintaining genomic integrity, such as *ATM* and a subset of Fanconi Anemia-related genes ^{72–74}. These findings may pave the way for early biomarkers in cancer detection and general health assessment in rare disease patients and carriers. Larger-scale studies with BSyn cohorts are imperative to unravel the mechanisms underpinning *BLM* loss-of-function variants and their contribution to CHIP and cancer risk.

Figures and Tables

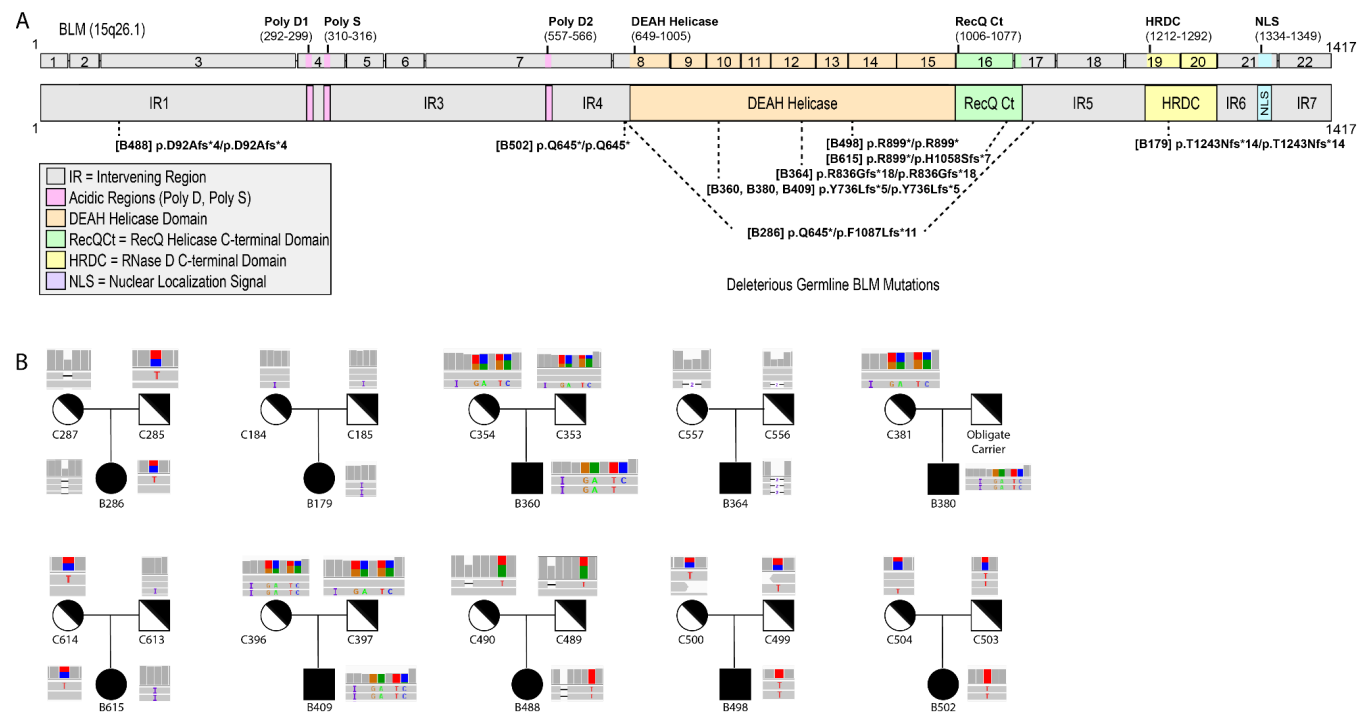


Figure 7-1: Genomic analysis of *BLM* mutations in Bloom syndrome patients and carriers.

(A) Schematic representation of the *BLM* transcript (ENST00000355112.8) and protein (GenBank: BLM; NM000057.4; GRCh38), its functional domains (solid lines above transcript), and mutations (dotted lines below transcript) causing Bloom Syndrome. Variants listed correspond to BSyn probands in this study and are tagged with patient identifiers (B#, see **Table 7-1**). The location of deleterious biallelic variants are shown with one dotted line, while compound heterozygous variants are shown with two dotted lines. **(B)** Variants in each BSyn proband (B#) and BSyn carrier (C#) in our cohort were verified in *Integrative Genomic Viewer v.2.9.4*. Each *IGV* screenshot shows coverage at the *BLM* variant at the top and the first two to three sequencing reads below with reference bases in grey and genetic variants in color. Histograms represent the coverage around the Bsyn variants at that site. Each trio relationship is depicted.

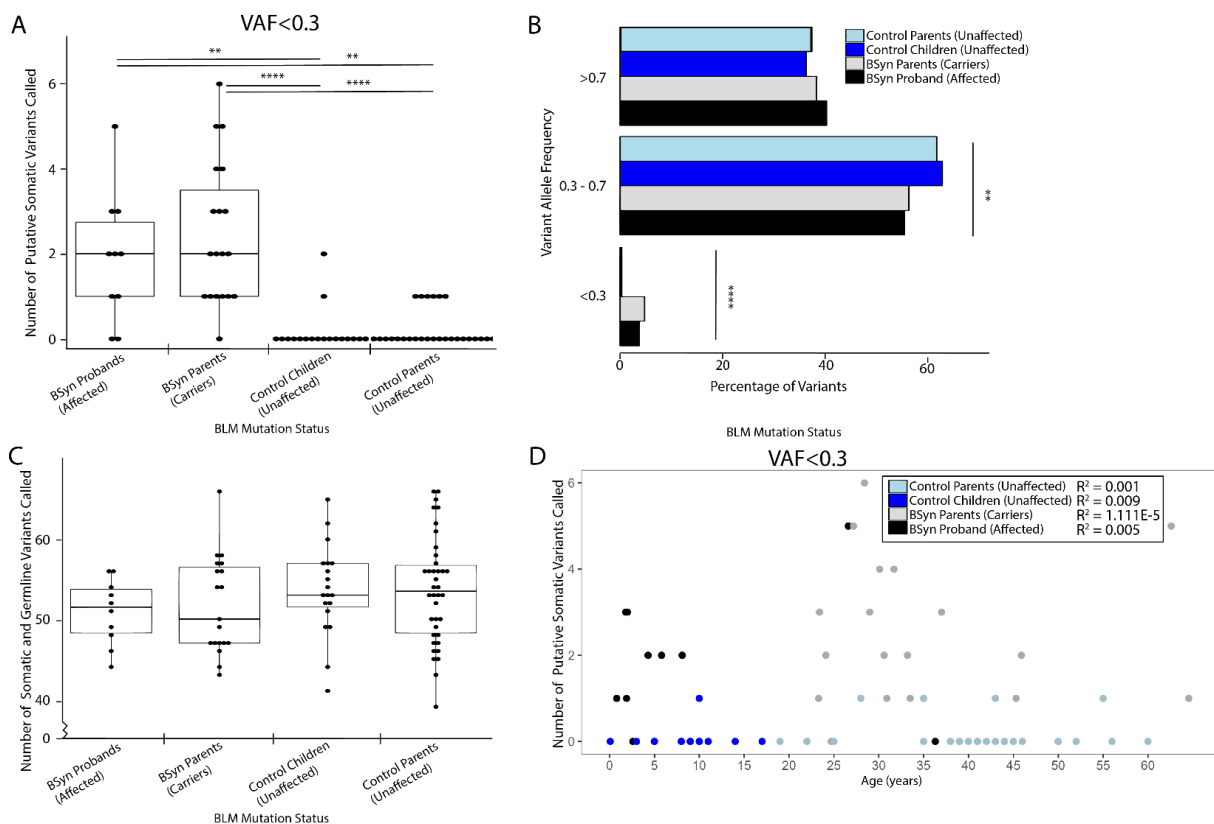


Figure 7-2: Comparative analysis of somatic CHIP gene variants identified significantly more in BSyn probands and carriers.

We examined variants in CHIP genes in exome sequenced samples from four cohorts: BSyn probands (n=10, black), BSyn carrier parents (n=19, grey), control children (n=19, light blue) and control parents (n=38, dark blue). **(A)** Using VarSeq variant calling, we identified the number of putative somatic CHIP variants using a variant allele frequency (VAF) cutoff < 0.3 (**Table 7-2**). **(B)** CHIP gene variants were subsetted based on variant allele frequency (VAF) grouping, and the mean proportions are shown. **(C)** The total number of germline and somatic CHIP gene variants are shown (**Table 7-2**). **(D)** We examined the effect of age on the number of putative somatic variants in CHIP genes that were called. Linear regression analysis was performed separately for each cohort and the R^2 values indicate goodness of fit for each model: $R^2 = 0.005$ (BSyn proband), $R^2 = 1.111E-05$ (carrier), $R^2 = 0.009$ (control child), $R^2 = 0.001$

(control parent). These values suggest a weak linear relationship between age and the frequency of somatic CHIP variants in our cohorts.

ns or no stars denote $p\text{-value} > .05$, * denote $p\text{-value} \leq .05$, ** denote $p\text{-value} < .01$, *** denote $p\text{-value} < .001$, **** denote $p\text{-value} < .0001$

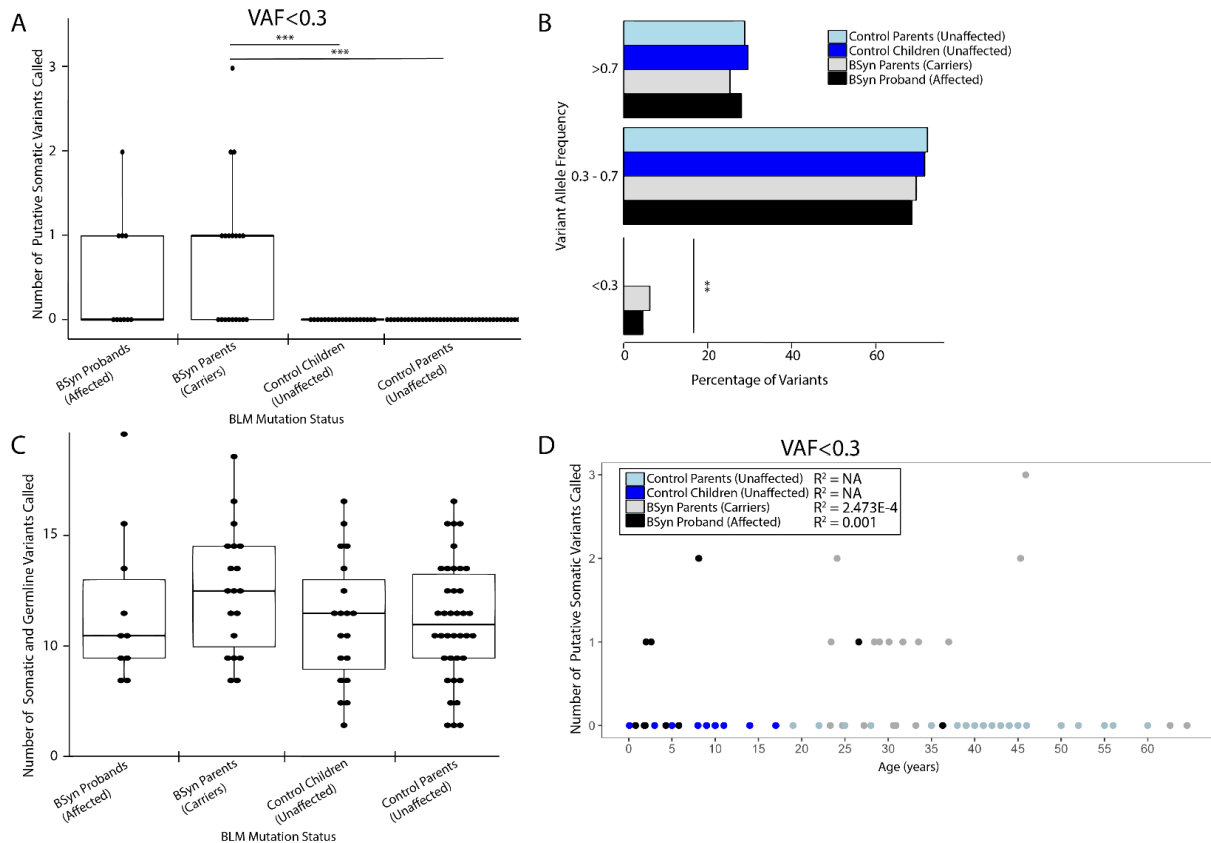


Figure 7-3: Comparative analysis of somatic DNA methylation gene variants identified significantly more in BSyn carriers compared to controls.

We examined variants in DNAm genes in exome sequenced samples from four cohorts: BSyn probands (n=10, black), BSyn carrier parents (n=19, grey), control children (n=19, light blue) and control parents (n=38, dark blue). **(A)** Using VarSeq variant calling, we identified the number of putative somatic DNAm variants using a variant allele frequency (VAF) cutoff < 0.3. **(B)** DNAm gene variants were subsetting based on variant allele frequency (VAF) grouping, and the mean proportions are shown. **(C)** The total number of germline and somatic DNAm gene variants are shown. **(D)** We examined the effect of age on the number of putative somatic variants in DNAm genes that were called. Linear regression analysis was performed separately for each cohort and the R^2 values indicate goodness of fit for each model: $R^2 = 0.001$ (BSyn proband), $R^2 = 2.473\text{E-}04$ (carrier), $R^2 = \text{NA}$ (control child), $R^2 = \text{NA}$ (control parent). These

values suggest a weak to no linear relationship between age and the frequency of somatic CHIP variants in our cohorts.

ns or no stars denote $p\text{-value} > .05$, * denote $p\text{-value} \leq .05$, ** denote $p\text{-value} < .01$, *** denote $p\text{-value} < .001$, **** denote $p\text{-value} < .0001$

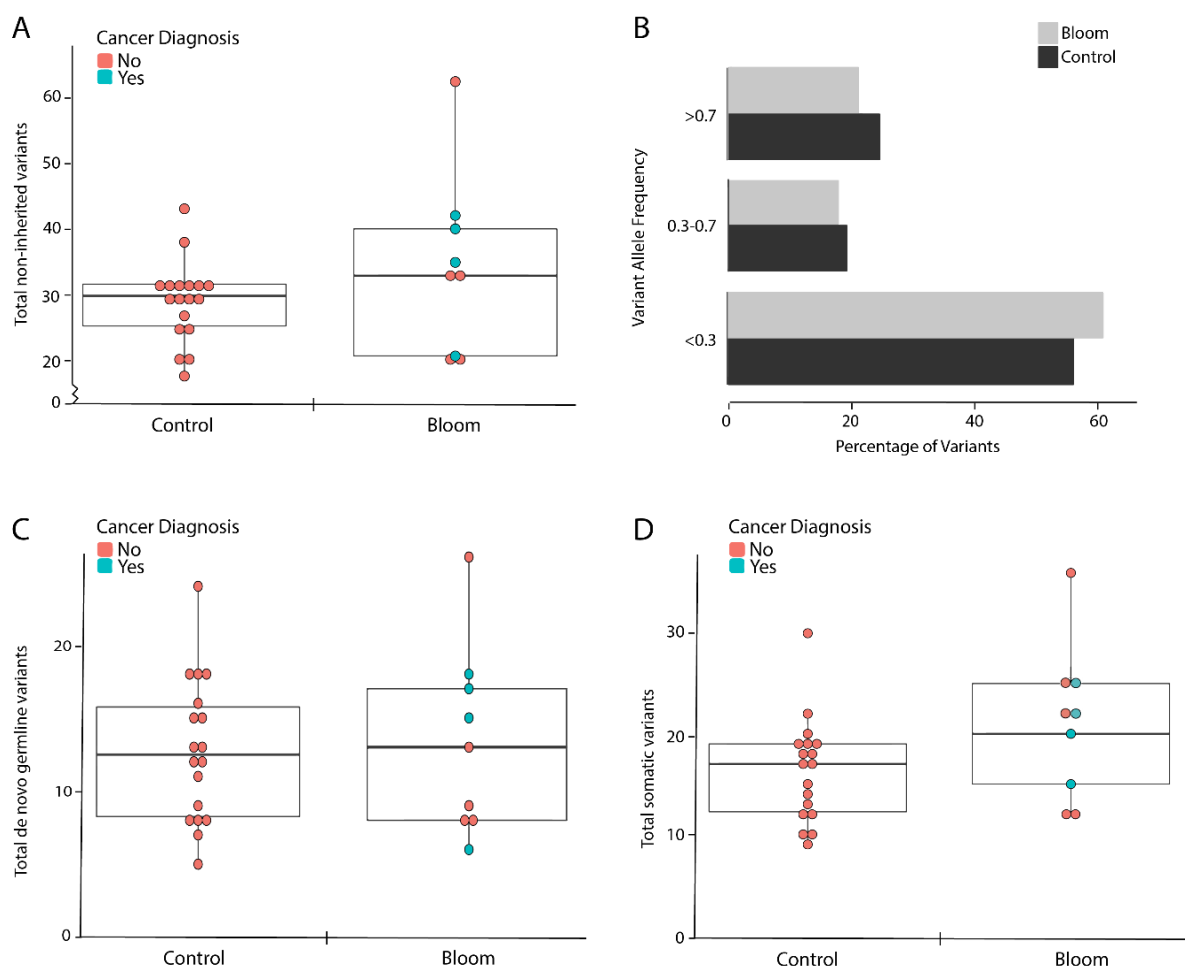


Figure 7-4: De novo and exome-wide somatic variant rates in BSyn and controls.

(A) We examined variants not inherited from either parent in our trio cohorts using VarSeq Trio Exome variant calling. This encompasses both germline *de novo* and newly arising somatic variants in BSyn probands (n=9) and age- and sex-matched control (n=18). T-test was conducted to compare BSyn and control numbers of variants (t-test, p -value=0.187). (B) Non-inherited variants in BSyn probands and age- and sex- matched controls were subsetting based on variant allele frequency (VAF) grouping. (C) The number of germline *de novo* variants ($VAF \geq 0.3$) for each sample are shown. T-test was conducted to compare BSyn and control numbers of variants (t-test, p -value=0.628). (D) The number of somatic variants ($VAF < 0.3$) for each sample are shown. T-test was conducted to compare BSyn and control numbers of

variants (t-test, p -value=0.083).

Pink indicates no cancer diagnosis, green indicates history of cancer. ns denotes p -value>0.05, implying no statistically significant difference between BSyn probands and controls.

Sample ID	Status	Sex	Age Range	Trio Grouping	BSyn Status	HGVS Nomenclature of BLM Mutation (NM_000057.4)	Cancer History
B179	Proband	F	5-9	BLOOM_179	Affected	NM_000057.4:c.[3727_3728insA];[3727_3728insA]	N
C185	Father	M	35-39		Carrier	NM_000057.4:c.3727_3728insA	N
C184	Mother	F	30-34		Carrier	NM_000057.4:c.3727_3728insA	N
B286	Proband	F	1-4	BLOOM_286	Affected	NM_000057.4:c.[1933C>T];[c.3261del]	N
C285	Father	M	30-34		Carrier	NM_000057.4:c.1933C>T	N
C287	Mother	F	25-29		Carrier	NM_000057.4:c.3261del	N
B360	Proband	M	<1	BLOOM_360	Affected	NM_000057.4:c.[2207_2212delinsTAGATTC];[2207_2212delinsTAGATTC]	Y (Acute lymphoblastic leukemia (7 years post-collection))
C353	Father	M	30-34		Carrier	NM_000057.4:c.2207_2212delinsTAGATTC	N
C354	Mother	F	20-24		Carrier	NM_000057.4:c.2207_2212delinsTAGATTC	N
B364	Proband	M	35-39	BLOOM_364	Affected	NM_000057.4:c.[2506_2507del];[2506_2507del]	Y (3x squamous cell carcinoma on skin (1 year pre-collection), acute lymphoblastic leukemia (4 years post-collection))
C556	Father	M	65-69		Carrier	NM_000057.4:c.2506_2507del	N
C557	Mother	F	60-64		Carrier	NM_000057.4:c.2506_2507del	N
B380	Patient	M	1-4	BLOOM_380	Affected	NM_000057.4:c.[2207_2212delinsTAGATTC];[2207_2212delinsTAGATTC]	Y (Gastrointestinal stromal tumour (17 years post-collection), diffuse B cell lymphoma (20 years post-collection))
C381	Mother	F	30-34		Carrier	NM_000057.4:c.2207_2212delinsTAGATTC	N
B409	Proband	M	1-4	BLOOM_409	Affected	NM_000057.4:c.[2207_2212delinsTAGATTC];[2207_2212delinsTAGATTC]	Y (Acute myeloid leukemia (20 years post-collection))
C396	Father	M	25-29		Carrier	NM_000057.4:c.2207_2212delinsTAGATTC	N
C397	Mother	F	20-24		Carrier	NM_000057.4:c.2207_2212delinsTAGATTC	N
B488	Proband	F	1-4	BLOOM_488	Affected	NM_000057.4:c.[275del];[275del]	N
C489	Father	M	25-29		Carrier	NM_000057.4:c.275del	N
C490	Mother	F	30-34		Carrier	NM_000057.4:c.275del	N
B498	Proband	M	5-9	BLOOM_498	Affected	NM_000057.4:c.[2695C>T];[2695C>T]	N
C499	Father	M	30-34		Carrier	NM_000057.4:c.2695C>T	N
C500	Mother	F	35-39		Carrier	NM_000057.4:c.2695C>T	N
B502	Proband	F	25-29	BLOOM_502	Affected	NM_000057.4:c.[1933C>T];[1933C>T]	Y (Bladder carcinoma (6 years post-collection))
C503	Father	M	45-49		Carrier	NM_000057.4:c.1933C>T	N
C504	Mother	F	45-49		Carrier	NM_000057.4:c.1933C>T	N
B615	Proband	F	1-4	BLOOM_615	Affected	NM_000057.4:c.[2695C>T];[3171_3172insT]	N
C613	Father	M	25-29		Carrier	NM_000057.4:c.3171_3172insT	N
C614	Mother	F	20-24		Carrier	NM_000057.4:c.2695C>T	N

Table 7-1: Bloom syndrome patient and *BLM* variant carrier demographics

Bloom Syndrome patient (B#, n=10) and *BLM* variant carrier (C#, n=19) parents included in our studies are identified with their patient IDs. Individual sex, age range, family grouping, *BLM* mutation and prior history of cancer are also shown.

SRA ID	Sample ID	Trio Status	Sex	Age	Ethnicity	BLM Mutation Status
SRS2136666	SRR5463149	Child	M	10-14	Not documented	Unaffected
SRS2813808	SRR6447679	Father	M	60-64	Not documented	Unaffected
SRS2136486	SRR5462651	Mother	F	50-54	Not documented	Unaffected
SRS2140039	SRR5469760	Child	M	10-14	White Caucasian	Unaffected
SRS2140061	SRR5469979	Father	M	50-54	White Caucasian	Unaffected
SRS2136721	SRR5463506	Mother	F	40-44	White Caucasian	Unaffected
SRS2203482	SRR5567876	Child	F	NA	Not documented	Unaffected
SRS2202906	SRR5565258	Mother	F	NA	Not documented	Unaffected
SRS2202907	SRR5565278	Father	M	NA	Not documented	Unaffected
SRS2130875	SRR5456220	Child	F	10-14	White Caucasian	Unaffected
SRS2136628	SRR5462955	Mother	F	50-54	White Caucasian	Unaffected
SRS2130876	SRR5456232	Father	M	40-44	White Caucasian	Unaffected
SRS2197363	SRR5553488	Child	F	5-9	Egyptian	Unaffected
SRS2197826	SRR5556325	Father	M	20-24	Egyptian	Unaffected
SRS2197795	SRR5555687	Mother	F	25-29	Egyptian	Unaffected
SRS2140305	SRR5470425	Child	F	NA	Not documented	Unaffected
SRS2137393	SRR5466100	Mother	F	NA	Not documented	Unaffected
SRS2137389	SRR5466052	Father	M	NA	Not documented	Unaffected
SRS2200570	SRR5561237	Child	M	5-9	White, non-hispanic	Unaffected
SRS2200550	SRR5560909	Mother	F	40-44	White, non-hispanic	Unaffected
SRS2200596	SRR5561706	Father	M	40-44	White, non-hispanic	Unaffected
SRS2195820	SRR5550774	Child	F	1-5	Egyptian	Unaffected
SRS2195786	SRR5550148	Mother	F	15-19	Egyptian	Unaffected
SRS2195798	SRR5550381	Father	M	20-24	Egyptian	Unaffected
SRS2200588	SRR5561584	Child	M	<1	White, non-hispanic	Unaffected
SRS2200627	SRR5562238	Mother	F	40-44	White, non-hispanic	Unaffected
SRS2200615	SRR5562034	Father	M	40-44	White, non-hispanic	Unaffected
SRS2205811	SRR5571393	Child	F	10-14	Egyptian	Unaffected
SRS2195821	SRR5550794	Mother	F	25-29	Egyptian	Unaffected
SRS2195834	SRR5551062	Father	M	35-39	Egyptian	Unaffected
SRS2136679	SRR5463270	Child	F	10-14	White Caucasian	Unaffected
SRS2136619	SRR5462872	Mother	F	50-54	White Caucasian	Unaffected
SRS2136617	SRR5462858	Father	M	40-44	White Caucasian	Unaffected
SRS2136629	SRR5462967	Child	M	5-9	India	Unaffected
SRS2136659	SRR5463095	Mother	F	45-49	India	Unaffected
SRS2136613	SRR5462820	Father	M	45-49	India	Unaffected
SRS2203316	SRR5566591	Child	F	NA	Not documented	Unaffected
SRS2202950	SRR5565931	Mother	F	NA	Not documented	Unaffected
SRS2202953	SRR5565968	Father	M	NA	Not documented	Unaffected

SRS2203490	SRR5567954	Child	F	NA	Not documented	Unaffected
SRS2203471	SRR5567694	Mother	F	NA	Not documented	Unaffected
SRS2203336	SRR5566906	Father	M	NA	Not documented	Unaffected
SRS2200551	SRR5560927	Child	M	10-14	White, non-hispanic	Unaffected
SRS2200573	SRR5561287	Mother	F	35-39	White, non-hispanic	Unaffected
SRS2200609	SRR5561934	Father	M	40-44	White, non-hispanic	Unaffected
SRS2200624	SRR5562184	Child	M	15-19	White, non-hispanic	Unaffected
SRS2200562	SRR5561107	Mother	F	55-59	White, non-hispanic	Unaffected
SRS2200610	SRR5561950	Father	M	55-59	White, non-hispanic	Unaffected
SRS2200576	SRR5561319	Child	M	5-9	White, non-hispanic	Unaffected
SRS2200561	SRR5561089	Mother	F	35-39	White, non-hispanic	Unaffected
SRS2288808	SRR5688669	Father	M	40-44	White, non-hispanic	Unaffected
SRS2288810	SRR5688704	Child	M	5-9	Hispanic/Latino	Unaffected
SRS2288816	SRR5688812	Mother	F	35-39	Hispanic/Latino	Unaffected
SRS2200626	SRR5562220	Father	M	35-39	Hispanic/Latino	Unaffected
SRS2200613	SRR5562000	Child	F	5-9	White, non-hispanic	Unaffected
SRS2288805	SRR5688631	Mother	F	35-39	White, non-hispanic	Unaffected
SRS2200605	SRR5561875	Father	M	40-44	White, non-hispanic	Unaffected

Table 7-2: Statistical calculations for CHIP gene variant analysis

Control children (n=19) and parents (n=38) included in our studies are identified with their sample IDs and SRA accession numbers from dbGaP. Individual sex, age range at sample collection, family grouping, genetic disorder and sequencing information are shown for samples where sample information was provided.

CHIP Gene Variants	Group 1	Group 2	n (Group 1)	Mean (Group 1)	Median (Group 1)	n (Group 2)	Mean (Group 2)
Total Variants	Affected	Carrier	10	50.9	51.5	19	51.74
Total Variants	Affected	Unaffected child	10	50.9	51.5	19	53.68
Total Variants	Affected	Unaffected parent	10	50.9	51.5	38	53.63
Total Variants	Carrier	Unaffected child	19	51.74	50	19	53.68
Total Variants	Carrier	Unaffected parent	19	51.74	50	38	53.63
Total Variants	Unaffected child	Unaffected parent	19	53.68	53	38	53.63
Somatic Variants	Affected	Carrier	10	1.9	2	19	2.47
Somatic Variants	Affected	Unaffected child	10	1.9	2	19	0.158
Somatic Variants	Affected	Unaffected parent	10	1.9	2	38	0.158
Somatic Variants	Carrier	Unaffected child	19	2.47	2	19	0.158
Somatic Variants	Carrier	Unaffected parent	19	2.47	2	38	0.158
Somatic Variants	Unaffected child	Unaffected parent	19	0.158	0	38	0.158

Table 7-3: Statistical calculations for CHIP gene variant analysis

One way (mixed) analysis of variance (ANOVA) model with a random family effect to allow for non-independence between cohorts was conducted for the total number of CHIP gene variants. The mean and median number of variants for each sample group, as well as the standard deviation (SD), pooled standard error of mean (SEM), and degrees of freedom (df) are shown. For each comparison, the adjusted p values are listed.

The non-parametric Kruskal-Wallis method was conducted for the somatic CHIP gene variants across the four sample groupings. The mean and median number of variants for each sample group, as well as the standard deviation (SD) are shown. For each comparison, the adjusted p values are listed.

DNA_m Gene Variants	Group 1	Group 2	n (Group 1)	Mean (Group 1)	Median (Group 1)	n (Group 2)	Mean (Group 2)
Total Variants	Affected	Carrier	10	11.2	10	19	12
Total Variants	Affected	Unaffected child	10	11.2	10	19	10.63
Total Variants	Affected	Unaffected parent	10	11.2	10	38	10.63
Total Variants	Carrier	Unaffected child	19	12	12	19	10.63
Total Variants	Carrier	Unaffected parent	19	12	12	38	10.63
Total Variants	Unaffected child	Unaffected parent	19	10.63	11	38	10.63
Somatic Variants	Affected	Carrier	10	0.5	0	19	0.74
Somatic Variants	Affected	Unaffected child	10	0.5	0	19	0
Somatic Variants	Affected	Unaffected parent	10	0.5	0	38	0
Somatic Variants	Carrier	Unaffected child	19	0.74	1	19	0
Somatic Variants	Carrier	Unaffected parent	19	0.74	1	38	0
Somatic Variants	Unaffected child	Unaffected parent	19	0	0	38	0

Table 7-4: Statistical calculations for DNA_m gene variant analysis

One way (mixed) analysis of variance (ANOVA) model with a random family effect to allow for non-independence between cohorts was conducted for the total number of DNA_m gene variants. The mean and median number of variants for each sample group, as well as the standard deviation (SD), pooled standard error of mean (SEM), and degrees of freedom (df) are shown. For each comparison, the adjusted p values are listed.

The non-parametric Kruskal-Wallis method was conducted for the somatic DNA_m gene variants across the four sample groupings. The mean and median number of variants for each sample group, as well as the standard deviation (SD) are shown. For each comparison, the adjusted p values are listed.

References

1. German, J., Archibald, R. & Bloom, D. CHROMOSOMAL BREAKAGE IN A RARE AND PROBABLY GENETICALLY DETERMINED SYNDROME OF MAN. *Science* **148**, 506–507 (1965).
2. Bloom, D. Congenital telangiectatic erythema resembling lupus erythematosus in dwarfs; probably a syndrome entity. *AMA Am. J. Dis. Child.* **88**, 754–758 (1954).
3. Cunliffe, C., Bassetti, J. A. & Ellis, N. A. Bloom's Syndrome: Clinical Spectrum, Molecular Pathogenesis, and Cancer Predisposition. *Mol. Syndromol.* **8**, 4–23 (2017).
4. Ellis, N. A. *et al.* The Ashkenazic Jewish Bloom syndrome mutation blmAsh is present in non-Jewish Americans of Spanish ancestry. *Am. J. Hum. Genet.* **63**, 1685–1693 (1998).
5. Sokolenko, A. P. *et al.* High prevalence and breast cancer predisposing role of the BLM c.1642 C>T (Q548X) mutation in Russia. *Int. J. Cancer* **130**, 2867–2873 (2012).
6. Wu, L. & Hickson, I. D. The Bloom's syndrome helicase suppresses crossing over during homologous recombination. *Nature* **426**, 870–874 (2003).
7. Manthei, K. A. & Keck, J. L. The BLM dissolvosome in DNA replication and repair. *Cell. Mol. Life Sci.* **70**, 4067–4084 (2013).
8. Chaganti, R. S., Schonberg, S. & German, J. A manyfold increase in sister chromatid exchanges in Bloom's syndrome lymphocytes. *Proc. Natl. Acad. Sci. U. S. A.* **71**, 4508–4512 (1974).
9. Onoda, F., Seki, M., Miyajima, A. & Enomoto, T. Elevation of sister chromatid exchange in *Saccharomyces cerevisiae* sgs1 disruptants and the relevance of the disruptants as a system to evaluate mutations in Bloom's syndrome gene. *Mutat. Res.* **459**, 203–209 (2000).
10. McDaniel, L. D. & Schultz, R. A. Elevated sister chromatid exchange phenotype of Bloom syndrome cells is complemented by human chromosome 15. *Proc. Natl. Acad. Sci. U. S. A.*

- 89**, 7968–7972 (1992).
11. Wang, W. *et al.* Functional relation among RecQ family helicases RecQL1, RecQL5, and BLM in cell growth and sister chromatid exchange formation. *Mol. Cell. Biol.* **23**, 3527–3535 (2003).
 12. Dicken, C. H., Dewald, G. & Gordon, H. Sister chromatid exchanges in Bloom's syndrome. *Arch. Dermatol.* **114**, 755–760 (1978).
 13. Hodson, C. *et al.* Mechanism of Bloom syndrome complex assembly required for double Holliday junction dissolution and genome stability. *Proc. Natl. Acad. Sci. U. S. A.* **119**, (2022).
 14. German, J., Crippa, L. P. & Bloom, D. Bloom's syndrome. III. Analysis of the chromosome aberration characteristic of this disorder. *Chromosoma* **48**, 361–366 (1974).
 15. German, J. Bloom's syndrome. XX. The first 100 cancers. *Cancer Genet. Cytogenet.* **93**, 100–106 (1997).
 16. Sugrañes, T. A. *et al.* Age of first cancer diagnosis and survival in Bloom syndrome. *Genet. Med.* **24**, 1476–1484 (2022).
 17. Ye, X. *et al.* Genomic characterization of lymphomas in patients with inborn errors of immunity. *Blood Adv* (2022) doi:10.1182/bloodadvances.2021006654.
 18. de Voer, R. M. *et al.* Deleterious Germline BLM Mutations and the Risk for Early-onset Colorectal Cancer. *Sci. Rep.* **5**, 14060 (2015).
 19. Thompson, E. R. *et al.* Exome sequencing identifies rare deleterious mutations in DNA repair genes FANCC and BLM as potential breast cancer susceptibility alleles. *PLoS Genet.* **8**, e1002894 (2012).
 20. Antczak, A. *et al.* A common nonsense mutation of the BLM gene and prostate cancer risk and survival. *Gene* **532**, 173–176 (2013).
 21. Baris, H. N. *et al.* Prevalence of breast and colorectal cancer in Ashkenazi Jewish carriers of Fanconi anemia and Bloom syndrome. *Isr. Med. Assoc. J.* **9**, 847–850 (2007).

22. Cleary, S. P. *et al.* Heterozygosity for the BLM(Ash) mutation and cancer risk. *Cancer Res.* **63**, 1769–1771 (2003).
23. Kluźniak, W. *et al.* Inherited Variants in BLM and the Risk and Clinical Characteristics of Breast Cancer. *Cancers* **11**, (2019).
24. Gruber, S. B. *et al.* BLM heterozygosity and the risk of colorectal cancer. *Science* **297**, 2013 (2002).
25. Bononi, A. *et al.* Heterozygous germline *BLM* mutations increase susceptibility to asbestos and mesothelioma. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 33466–33473 (2020).
26. Jaiswal, S. *et al.* Age-related clonal hematopoiesis associated with adverse outcomes. *N. Engl. J. Med.* **371**, 2488–2498 (2014).
27. Schenz, J. *et al.* Increased prevalence of clonal hematopoiesis of indeterminate potential in hospitalized patients with COVID-19. *Front. Immunol.* **13**, 968778 (2022).
28. Steensma, D. P. *et al.* Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes. *Blood* **126**, 9–16 (2015).
29. Marnell, C. S., Bick, A. & Natarajan, P. Clonal hematopoiesis of indeterminate potential (CHIP): Linking somatic mutations, hematopoiesis, chronic inflammation and cardiovascular disease. *J. Mol. Cell. Cardiol.* **161**, 98–105 (2021).
30. Yu, B. *et al.* Supplemental Association of Clonal Hematopoiesis With Incident Heart Failure. *J. Am. Coll. Cardiol.* **78**, 42–52 (2021).
31. Fuster, J. J. & Walsh, K. Somatic Mutations and Clonal Hematopoiesis: Unexpected Potential New Drivers of Age-Related Cardiovascular Disease. *Circ. Res.* **122**, 523–532 (2018).
32. Mitchell, E. *et al.* Clonal dynamics of haematopoiesis across the human lifespan. *Nature* **606**, 343–350 (2022).
33. Bick, A. G. *et al.* Inherited causes of clonal haematopoiesis in 97,691 whole genomes. *Nature* **586**, 763–768 (2020).

34. Hagiwara, K. *et al.* Dynamics of age- versus therapy-related clonal hematopoiesis in long-term survivors of pediatric cancer. *Cancer Discov.* (2023) doi:10.1158/2159-8290.CD-22-0956.
35. Loh, P.-R. *et al.* Insights into clonal haematopoiesis from 8,342 mosaic chromosomal alterations. *Nature* **559**, 350–355 (2018).
36. Loh, P.-R., Genovese, G. & McCarroll, S. A. Monogenic and polygenic inheritance become instruments for clonal selection. *Nature* **584**, 136–141 (2020).
37. DeBoy, E. A. *et al.* Familial Clonal Hematopoiesis in a Long Telomere Syndrome. *N. Engl. J. Med.* **388**, 2422–2433 (2023).
38. Schratz, K. E. *et al.* Cancer spectrum and outcomes in the Mendelian short telomere syndromes. *Blood* **135**, 1946–1956 (2020).
39. Schratz, K. E. *et al.* Somatic reversion impacts myelodysplastic syndromes and acute myeloid leukemia evolution in the short telomere disorders. *J. Clin. Invest.* **131**, (2021).
40. Liu, J. *et al.* Genetic Determinants of Clonal Hematopoiesis and Progression to Hematologic Malignancies in 479,117 Individuals. doi:10.1182/blood-2023-182423.
41. Hoehn, H. & Salk, D. Clonal analysis of stable chromosome rearrangements in Bloom's syndrome fibroblasts. *Cancer Genet. Cytogenet.* **11**, 405–415 (1984).
42. Vencio, E. F. *et al.* Clonal cytogenetic abnormalities in Erdheim-Chester disease. *Am. J. Surg. Pathol.* **31**, 319–321 (2007).
43. McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
44. Nextera Flex for Enrichment Documentation. https://support.illumina.com/sequencing/sequencing_kits/nextera-flex-for-enrichment-kit/documentation.html.
45. Strom, S. P. Current practices and guidelines for clinical next-generation sequencing oncology testing. *Cancer Biol Med* **13**, 3–11 (2016).

46. DeLeonardis, K., Hogan, L., Cannistra, S. A., Rangachari, D. & Tung, N. When Should Tumor Genomic Profiling Prompt Consideration of Germline Testing? *J. Oncol. Pract.* **15**, 465–473 (2019).
47. Friedlaender, A., Tsantoulis, P., Chevallier, M., De Vito, C. & Addeo, A. The Impact of Variant Allele Frequency in EGFR Mutated NSCLC Patients on Targeted Therapy. *Front. Oncol.* **11**, 644472 (2021).
48. Coffee, B. *et al.* A substantial proportion of apparently heterozygous TP53 pathogenic variants detected with a next-generation sequencing hereditary pan-cancer panel are acquired somatically. *Hum. Mutat.* **41**, 203–211 (2020).
49. Wang, Y. *et al.* Comprehensive identification of somatic nucleotide variants in human brain tissue. *Genome Biol.* **22**, 92 (2021).
50. Stacey, S. N. *et al.* Genetics and epidemiology of mutational barcode-defined clonal hematopoiesis. *Nat. Genet.* (2023) doi:10.1038/s41588-023-01555-z.
51. Pan, H. *et al.* Discovery of Candidate DNA Methylation Cancer Driver Genes. *Cancer Discov.* **11**, 2266–2281 (2021).
52. Lee, C.-J. *et al.* Impact of mutations in DNA methylation modification genes on genome-wide methylation landscapes and downstream gene activations in pan-cancer. *BMC Med. Genomics* **13**, 27 (2020).
53. Levine, M. E. *et al.* An epigenetic biomarker of aging for lifespan and healthspan. *Aging* **10**, 573–591 (2018).
54. Lee, J. *et al.* Bloom syndrome patients and mice display accelerated epigenetic aging. *Aging Cell* e13964 (2023).
55. Guastafierro, T. *et al.* Genome-wide DNA methylation analysis in blood cells from patients with Werner syndrome. *Clin. Epigenetics* **9**, 92 (2017).
56. Maierhofer, A. *et al.* Accelerated epigenetic aging in Werner syndrome. *Aging* **9**, 1143–1152 (2017).

57. Rasmussen, K. D. & Helin, K. Role of TET enzymes in DNA methylation, development, and cancer. *Genes Dev.* **30**, 733–750 (2016).
58. Ginno, P. A. *et al.* A genome-scale map of DNA methylation turnover identifies site-specific dependencies of DNMT and TET activity. *Nat. Commun.* **11**, 2680 (2020).
59. Kong, A. *et al.* Rate of de novo mutations and the importance of father's age to disease risk. *Nature* **488**, 471–475 (2012).
60. Acuna-Hidalgo, R. *et al.* Ultra-sensitive Sequencing Identifies High Prevalence of Clonal Hematopoiesis-Associated Mutations throughout Adult Life. *Am. J. Hum. Genet.* **101**, 50–64 (2017).
61. Genovese, G. *et al.* Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N. Engl. J. Med.* **371**, 2477–2487 (2014).
62. Prokofyeva, D. *et al.* Nonsense mutation p.Q548X in BLM, the gene mutated in Bloom's syndrome, is associated with breast cancer in Slavic populations. *Breast Cancer Res. Treat.* **137**, 533–539 (2013).
63. Rashkin, S. R. *et al.* Pan-cancer study detects genetic risk variants and shared genetic basis in two large cohorts. *Nat. Commun.* **11**, 4423 (2020).
64. Laitman, Y. *et al.* The risk for developing cancer in Israeli ATM, BLM, and FANCC heterozygous mutation carriers. *Cancer Genet.* **209**, 70–74 (2016).
65. Climente-González, H. *et al.* Boosting GWAS using biological networks: A study on susceptibility to familial breast cancer. *PLoS Comput. Biol.* **17**, e1008819 (2021).
66. Xie, M. *et al.* Age-related mutations associated with clonal hematopoietic expansion and malignancies. *Nat. Med.* **20**, 1472–1478 (2014).
67. Alhumaid, M. *et al.* Prognostic role of multiparameter flow cytometry-based measurable residual disease assessment in patients with acute myeloid leukemia harboring DNMT3A/TET2/ASXL1 mutation. *Blood* **136**, 8–9 (2020).
68. Hop, P. J. *et al.* Genome-wide identification of genes regulating DNA methylation using

- genetic anchors for causal inference. *Genome Biol.* **21**, 220 (2020).
69. de Renty, C. & Ellis, N. A. Bloom's syndrome: Why not premature aging?: A comparison of the BLM and WRN helicases. *Ageing Res. Rev.* **33**, 36–51 (2017).
70. Bolton, K. L. *et al.* Cancer therapy shapes the fitness landscape of clonal hematopoiesis. *Nat. Genet.* **52**, 1219–1226 (2020).
71. Levin, M. G. *et al.* Genetics of smoking and risk of clonal hematopoiesis. *Sci. Rep.* **12**, 7248 (2022).
72. Goss, K. H. *et al.* Enhanced tumor formation in mice heterozygous for Blm mutation. *Science* **297**, 2051–2053 (2002).
73. Hall, M. J. *et al.* Germline Pathogenic Variants in the Ataxia Telangiectasia Mutated (ATM) Gene are Associated with High and Moderate Risks for Multiple Cancers. *Cancer Prev. Res.* **14**, 433–440 (2021).
74. Del Valle, J. *et al.* Exploring the Role of Mutations in Fanconi Anemia Genes in Hereditary Cancer Patients. *Cancers* **12**, (2020).

CHAPTER 8

Conclusion: Bridging Genes and Diseases: From Insights to Impact

Concluding remarks

This dissertation embarked on a comprehensive exploration of the *ASXL* gene family, with a focus on the profound and multifaceted impact of *ASXL1* mutations on human development and disease. Through a series of detailed studies, we have expanded the understanding of how these mutations disrupt various biological processes and contribute to the pathogenesis of disorders like Bohring-Opitz syndrome and acute myeloid leukemia.

The multi-omics approach employed across different chapters provided a holistic view of the epigenetic, transcriptomic, and metabolic alterations induced by *ASXL1* mutations. Notably, our findings underscore the pervasive influence of these mutations on Wnt signaling pathways (**Chapter 2**), neural crest formation (**Chapter 3**), mitochondrial function (**Chapter 4, Chapter 5**), and metabolic reprogramming (**Chapter 4**). These insights are critical for developing targeted therapies that could mitigate the severe developmental and metabolic anomalies observed in affected individuals.

Furthermore, the comparative analysis between BOS and AML highlighted shared molecular signatures, suggesting common mechanistic pathways that may be exploited for therapeutic benefits across multiple diseases (**Chapter 6**). This suggests a potential for broad-spectrum therapies that target these shared pathways, providing a foundation for cross-disease treatment strategies.

We also established links for genetic predispositions to increased somatic mutations in clonal haematopoiesis of indeterminate potential (CHIP) genes in another rare genetic disease, Bloom Syndrome. In both Bloom Syndrome individuals and carriers, increased numbers of somatic mutations in these CHIP genes provide new insights into the mechanisms underlying malignant potential and help identify potential targets for therapeutic intervention (**Chapter 7**).

Challenges and limitations

Despite these advances, our research encountered challenges, including limited sample availability, an inherent limitation to any rare disease study, phenotypic variability, and finding suitable age-matched healthy controls for pediatric patients. Furthermore, the structural complexity of the *ASXL* genes and lack of specific antibodies for the *ASXL* proteins limited our ability to resolve the *ASXL* genes at the protein level. These limitations underscore the intricate nature of genetic research and highlight areas where further methodological and technological advancements are needed.

Future research directions

The insights gained from this dissertation open numerous avenues for future research. The potential therapeutic targets identified—such as Wnt signaling, MPC1 and MPC2, and the modulation of L-asparagine and pyruvate pathways—offer promising directions for developing targeted interventions for BOS, AML, and the broader *ASXL* gene family and genetic neurodevelopmental disorders. Expanding these studies to include other rare genetic disorders and leveraging emerging techniques like single-cell sequencing, holds promise for breakthroughs in the diagnosis, prevention, and treatment of genetic disorders. The establishment of collaborative groups and increased data and sample sharing, facilitated through initiatives like the *Rare Epigenetic and Chromatin Disorders* (REACH) biobank established through our lab and our collaborators, are crucial steps toward fostering innovation and accelerating research in this field.

In conclusion, this dissertation represents a significant step forward in our understanding of the *ASXL* gene family and its multifaceted roles in human health and disease. By bridging the fields of rare genetic disorders, cancer and metabolism, this research not only contributes to the body of knowledge in genetics and biomedical research but also holds the potential to inform the development of novel therapeutic strategies, ultimately improving patient care and outcomes

in these complex conditions. I hope that my findings have laid a foundation to expand our understanding and one day offer therapeutic options for the patients and families who have made this research possible.