Check for updates

Transcription factor genetics and biology in predisposition to bone marrow failure and hematological malignancy

Jiarna R. Zerella[1,2], Claire C. Homan[2,3], Peer Arts[2,3],
Anna L. Brown[1,2,3], Hamish S. Scott[1,2,3] and
Christopher N. Hahn[1,2,3]*

[1]Adelaide Medical School, Faculty of Health and Medical Sciences, University of Adelaide, Adelaide,
SA, Australia, [2]Centre for Cancer Biology, SA Pathology and University of South Australia, Adelaide,
SA, Australia, [3]Department of Genetics and Molecular Pathology, SA Pathology, Adelaide, SA, Australia

Transcription factors (TFs) play a critical role as key mediators of a multitude of developmental pathways, with highly regulated and tightly organized networks crucial for determining both the timing and pattern of tissue development. TFs can act as master regulators of both primitive and definitive hematopoiesis, tightly controlling the behavior of hematopoietic stem and progenitor cells (HSPCs). These networks control the functional regulation of HSPCs including self-renewal, proliferation, and differentiation dynamics, which are essential to normal hematopoiesis. Defining the key players and dynamics of these hematopoietic transcriptional networks is essential to understanding both normal hematopoiesis and how genetic aberrations in TFs and their networks can predispose to hematopoietic disease including bone marrow failure (BMF) and hematological malignancy (HM). Despite their multifaceted and complex involvement in hematological development, advances in genetic screening along with elegant multi-omics and model system studies are shedding light on how hematopoietic TFs interact and network to achieve normal cell fates and their role in disease etiology. This review focuses on TFs which predispose to BMF and HM, identifies potential novel candidate predisposing TF genes, and examines putative biological mechanisms leading to these phenotypes. A better understanding of the genetics and molecular biology of hematopoietic TFs, as well as identifying novel genes and genetic variants predisposing to BMF and HM, will accelerate the development of preventative strategies, improve clinical management and counseling, and help define targeted treatments for these diseases.

KEYWORDS

transcription factor, hematological malignancies (HM), bone marrow failure (BMF), germline, pathogenic variant

Introduction

Over the last decade, the genomic revolution in combination with access to well curated clinical patient information and samples has enabled identification of rare germline syndromes of overlapping and diverse clinical manifestations with different propensities for development of gene-specific bone marrow failure (BMF) and/or hematological malignancy (HM). Around 200 genes are currently included on high evidence germline targeted sequencing panels for HM predisposition, bleeding, and platelet disorders and BMF syndromes (1). The accruing identification of these predisposition genes helps to further inform leukemic biology and disease causation, allows for earlier diagnosis and genetic counseling for individuals who are at higher risk for the disease, permits the development of new and more effect treatments and, may improve our ability to develop preventative measures, targeting populations at higher risk of developing a HM.

Amongst the genes associated with HM and BMF, transcription factors (TFs) are commonly identified as mediators of hereditary predisposition and play a role in leukemogenesis by frequently exhibiting recurrent, somatically acquired chromosomal abnormalities and smaller point mutations and indels. The human genome contains over 1,800 genes that encode for TFs that display complex combinatorial interactions resulting in homeostatic transcriptional networks with positive and negative feedback loops to precisely regulate gene expression for cell fate trajectories and transitions, and cellular responses to environmental triggers (2, 3). Unironically, it is TFs that orchestrate gene expression regulation at each stage of hematopoietic development, including stem cell formation and maintenance, and lineage commitment and homeostasis (4, 5). Despite the identification of numerous master hematopoietic TFs as predisposition genes (inherited or *de novo*) (*RUNX1, CEBPA, GATA2, ETV6, PAX5, IKZF1*), it is likely that unrecognized predisposition genes and variants including in TFs will be discovered considering the complexity of the transcriptional network in normal hematopoietic stem cells (HSC), and the supporting bone marrow microenvironment. Stratifying BMF and/or HM TF predisposition genes currently relies on observed recurrence and functional evaluation, and excludes ontology, environmental factors, and somatic data (6). The prospective inclusion of these could provide a more comprehensive and accurate assessment of the genetic factors that contribute to BMF and/or HM development as well as for instance identification and inclusion of polygenic risk factors derived from multiple genetic variants of differing penetrance. Different types of variants in known BMF and HM predisposition genes display different pathogenic phenotypes and levels of penetrance (7).

Given the intricate interconnections among TFs and their multifaceted involvement in hematopoiesis, it is challenging to delineate the precise mechanism by which each TF aberration contributes to disease. It is suspected that TFs can contribute to disease using a wide range of mechanisms, including initiating the activation or repression of gene expression directly or through the recruitment of cofactors, epigenome changes, initiating new

chromatin looping interactions between enhancers and target promoters, and altering the chromatin landscape through the repositioning of nucleosomes (8). This review will cover putative mechanisms leading to HM oncogenicity, including aberrations to hematopoietic TF networks, lymphoid and myeloid HM and BMF predisposition genes, and stress selection, the prospective inclusion of alternative data to TF stratification, as well as the potential of undiscovered predisposition TFs. Given other recent comprehensive reviews of BMF and/or HM predisposition (9–11) and the subjectivity of novel candidate inclusions to TF BMF and/or HM predisposition genes, this review will primarily focus on DNA binding predisposition TFs currently found in high evidence Genomics England PanelApp panels (Hematological malignancies cancer susceptibility [Version 3.3], Bleeding and Platelet Disorders [Version 1.16] Cytopenia's and congenital anemias [Version 1.111]) (Table 1) and candidate TFs with an important role in normal hematopoiesis, reported leukemic association and/or oncogenic potential (Tables 2, 3). As our knowledge of molecular TF biology expands, so will our ability to definitively establish HM disease diagnosis or reoccurrence, predict prognosis and response to therapy, tailor treatments, and ultimately implement prevention, management, surveillance, and treatment strategies (33–35).

Oncogenic TF networks in hematopoiesis

Hematopoiesis is a complex process in which multiple regulatory TF pathways converge to orchestrate tissue formation, cell movement and cell fate decisions including lineage specification and differentiation. Analysis of genome wide binding patterns identified ten key TF regulators of hematopoietic stem/progenitor cells (HSPC) (i.e., TAL1, LYL1, LMO2, GATA2, RUNX1, MEIS1, PU.1, ERG, FLI1 and GFI1B) (36). The transcriptional role of each TF in HSPCs has been somewhat demonstrated, with combinatorial interaction studies suggesting highly cooperative control of transcription by a core set of seven TFs (i.e., FLI1, ERG, GATA2, RUNX1, SCL/TAL1, LYL1, LMO2) (36). During hematopoiesis, TFs form densely interconnected spatially and temporally regulated networks which either engage regions through cognate motifs or bind indirectly *via* the formation of protein-protein interactions (37). If a single player within the network is mutated, it can lead to TF dysregulation of interacting partners or target genes, impaired differentiation, fuel immature cell population growth and/or trigger inappropriate transcriptional programs; all of which may initiate malignancy. This highlights that small genetic changes in TFs can have a wide-ranging impact, affecting many of the components of the transcriptional network of which they are a part (38).

The TF network regulates multiple aspects of normal HSPC function and thus components of these transcriptional networks are a common target of aberration in leukemias, resulting in selection of cells with a leukemic 'stem cell' signature (39). For example, dysregulation of GATA2, a master regulator of hematopoiesis, causes hematologic pathologies. Around 80% of germline GATA2 carriers develop a myeloid malignancy before the age of 40 years

TABLE 1 Known bone marrow failure and/or hematological malignancy predisposition transcription factors.

TF Groups	Gene symbol	Germline HM gene	Germline blood disorder gene	Canonical Transcript Coding size (aa)	TF Family	DNA binding motif	Human phenotype			Role in hema-topoiesis (downstream differentiation)	Target genes
							OMIM		Mondo		ENCODE and ChEA Consensus TFs from ChIP-X
							Reported phenotype	Zygosity	Reported Phenotype #		
Best characterized BMF/HM genes	**RUNX1**	**Yes**	**Yes**	480	RUNX	TGTGG	AML Platelet disorder	AD	AML hereditary thrombocytopenia B-lymphoblastic leukemia/lymphoma with t(12;21)(p13.2;q22.1)	HSC, CMP, Megakaryocyte, Platelet, B-cell, T-cell	GATA2, GATA1, MAX, BRCA1, TCF3, CREB1, UBTF, TAF1, ZBTB7A
	CEBPA	**Yes**	**No**	358	bZIP	CCAAT T(T/G) NNGNAA (T/G)	AML	AD	AML	Granulocyte, Eosinophil, Monocyte, Neutrophil	ATF2, CREB1, CIN3A, CEBPD, NFYA, CREB1, NFYB, SIX5, GABPA
	GATA2	**Yes**	**Yes**	480	GATA	(A/T)GATA (A/G)	Emberger syndrome Immunodeficiency 21 AML MDS	AD	MDS/AML deafness-lymphedema-leukemia syndrome monocytopenia with susceptibility to infections	HSC, Mast cell, Erythrocyte, Basophil, Megakaryocyte	GATA1, RUNX1, SMAD4, TP63, NFE2L2, PPARG, UBTF, SPI1, RCOR1
PAX5, MECOM, GATA1 and STAT3 BMF/ HM genes	**PAX5**	Yes	**No**	391	PAX	GGCTGAG	ALL	AD	ALL	Myeloid/lymphoid progenitor), B-cell	Not included in dataset
	MECOM (MDS1/ EVI1)	Yes	Yes	1239	ZF	AAGA(C/T) AAGATAA	Radioulnar synostosis with amegakaryoctiyc thrombocytopenia	AD	AML MECOM-associated syndrome radioulnar synostosis with amegakaryyocytic thrombocytopenia	HSC	SPI1, RUNX1, GATA1, GATA2, FLI1, ZNF384, FOS, RCOR1, TCF3
	STAT3	Yes	Yes	770	STAT	TTCC(C/G) GGAA	Autoimmune disease Hyper-IgE recurrent infection syndrome	AD AR	STAT3-related early-onset multisystem autoimmune disease hyper-IgE recurrent infection syndrome 1	HSC HPC	AR, SMAD4, SUZ12, REST, TCF3, SOX2, POU5F1, NFE2L2, NANOG
	GATA1	Yes	Yes	413	GATA	(A/T)GATA (A/G)	Anemia Hemolytic anemia Leukemia Thrombocytopenia with beta-thalassemia	XLR	GATA1-Related X-Linked Cytopenia AML thrombocytopenia transient myeloproliferative syndrome	Eosinophil, Platelet, RBC Basophil, Mast-cell, Erythrocytes, Dendritic cell,	GATA2, RUNX1, SMAD4, AR, ZBTB7A, SPI1, UBTF, ESR1, STAT3

(Continued)

TABLE 1 Continued

TF Groups	Gene symbol	Germline HM gene	Germline blood disorder gene	Canonical Transcript Coding size (aa)	TF Family	DNA binding motif	Human phenotype				Role in hema-topoiesis (downstream differentiation)	Target genes
							OMIM		Mondo			ENCODE and ChEA Consensus TFs from ChIP-X
							Reported phenotype	Zygosity	Reported Phenotype #			
							Thrombocytopenia with/without dyserythropoeitic anemia				Megakaryocyte, Macrophage, MPC	
ETS BMF/ HM genes	ETV6 (TEL)	Yes	Yes	452	ETS	(G/C) CGGAAGT (G/A)	AML Thrombocytopenia	AD	Otosclerosis		HSC Megakaryocyte	Not included in dataset
	FLI1	No	Yes	452	ETS	AGGAA(G/A)	Bleeding disorder	AR AD	Nephronophthisis, Thrombocytopenia, Otosclerosis, paragangliomas, anencephaly, odontochondrodysplasia, schizophrenia, trigonocephaly, thrombocythemia, chondrocalcinosis		Platelet	FLI1, ELF1, SPI1, TAF1, ATF2, RUNX1, BRCA1, CREB1, NRF1
IKAROS BMF/HM genes	IKZF1	Yes	Yes	519	IKAROS	GGAAA	Immunodeficiency	AD	No Phenotypes reported		T-cell, HSC, CLP	ZKSCAN1, VDR, RCOR1, REST, SUZ12, NFE2L2, SPI1, UBTF
	IKZF3	Yes	Yes	509	IKAROS	(T/G)GGAA	Immunodeficiency	AD	No Phenotypes reported		B-cell, T-cell	Not included in data set
	IKZF5	No	Yes	419	IKAROS	GNNTGTNG	Thrombocytopenia	AD	No Phenotypes reported		HSC	Not included in data set

Germline HM and/or blood disorder predisposition genes, found in high evidence Genomics England PanelApp panels: Hematological malignancies cancer susceptibility [Version 3.3] and/or Bleeding and Platelet Disorders [Version 1.16] as well Cytopenias and congenital anemias [Version 1.111]. Known interacting partners were collated through STRING (12), target genes through ChEA from Chip-X (13), and roles in hematopoiesis literature (14–24). Autosomal dominant, AD; Autosomal recessive, AR.

TABLE 2 Oncogenic potential of hematopoietic-related TFs.

	TF gene symbol	Functional domain	MGI Mouse BMF/HM phenotype	Embryonic lethality			Constraint				COSMIC no. of LOFs	Malignant fusions		Biallelic Variants Observed	Known CHIP genes
				Het mice	Hom mice	Human	LOEUF	pLI	Domain	Average constraint		Somatic	HM specific		
Known BMF/HM TFs	**RUNX1**	RUNT TAD	Yes	No	Yes	NA/ Miscarriage observed	0.44	0.65	RUNT	0.48, intolerant	819 (23.9%)	AML1/MDS1, AML1/ETO, AML1/MDS1/ EAI1, AML1/ FOG2, RUNX1/ YTHDF2, RUNX1/SH3D19, RUNX1/ZNF687	AML1-ETO AML1-FOG2 AML-TEL	Yes	Yes (rare)
	CEBPA	TAD1 TAD2 DBD Zip	Yes	No	Die ~4-10 h after birth	NA	1.18	0.55	Zip	0.5032, intolerant	711 (43.1%)	No somatic lesions found	No HM specific malignant fusions found	Yes	No
	GATA2	TAL1 TAL2 ZF 1 ZF 2	Yes	No	Yes	NA/ Miscarriage observed	0.29	0.98	ZF 1 ZF 2	0.13, highly intolerant 0.149, highly intolerant	63 (6.67%)	inv(3)(q21q26)	inv(3)(q21q26)	No (rare)	No
	PAX5	OP, HD TA ID	Yes	No	Yes	NA	0.17	1	Paired domain	0.23, intolerant	80 (4.3%)	No somatic lesions found	PAX5/NOL4L	Yes	No
	MECOM/ EVI1	ZF (x10) AD	Yes	No	Yes	Yes	0.14	1	ZF, ZF, ZF	0.24, intolerant 0.28, intolerant 0.08, highly intolerant	115 (8.69%)	EVI1/RBPH1	EVI1/GR6	No	No
	STAT3	NTD, CC, DBD, LD, SH2, TAD	Yes	No	Yes	NA/ Miscarriage observed	0.10	1	NTD, CC, DBD, LD, SH2, TAD	0.19, intolerant	56 (3.67%)	No somatic lesions found	No HM specific malignant fusions found	No	Yes
	GATA1	AD ZF 1 ZF 2	Yes	No	Yes	Yes (Males)/ Miscarriage observed	0.32	0.95	ZF 1, ZF 2	0.18, intolerant 0.17, highly intolerant	269 (34.1%)	No somatic lesions found	MYB-GATA1	No	No
	ETV6 (TEL)	HLH ETS	Yes	No	Yes	NA/ Miscarriage observed	0.32	0.97	ETS	0.173, highly tolerant	126 (7.3%)	ETV6/PDGFRB, ETV6/MDS2, ETV6/NTRK3,	ETV6/AML1, ETV6/JAK2, ETV6/RUNX1, ETV6/MN1,	Yes	No

(Continued)

TABLE 2 Continued

	TF gene symbol	Functional domain	MGI Mouse BMF/HM phenotype	Embryonic lethality			Constraint				COSMIC no. of LOFs	Malignant fusions		Biallelic Variants Observed	Known CHIP genes
				Het mice	Hom mice	Human	LOEUF	pLI	Domain	Average constraint		Somatic	HM specific		
													ETV6/ACS2, ETV6/ABL2, ETV6/ARNT, ETV6/BTL, ETV6/PER1		
	FLI1	ETS	Yes	No	No	NA	0.28	0.99	ETS	0.0953, highly intolerant	22 (1.97%)	FLI1/EWS	No HM specific malignant fusions found	Yes	No
	IKZF1	ZF (x6)	Yes	No	Yes	NA	0.16	1	ZF	0.1086, highly intolerant	80 (3.41%)	No somatic lesions found	IKAROS/BCL6	No	No
	IKZF3	ZF (x6)	Yes	No	No	NA	0.29	0.98	ZF	0.31, intolerant	29 (2.80%)	No somatic lesions found	No HM specific malignant fusions found	No	No
	IKZF5	ZF (x5)	Yes	No	No	NA	0.46	0.78	ZF	No data available	32 (13.56%)	No somatic lesions found	No HM specific malignant fusions found	No	No
Strong candidate BMF/HM TFs	GATA3	TAL1 TAL2 ZF 1 ZF 2	Yes	No	Yes	Yes	0.39	0.9	ZF 1, ZF 2	0.12, highly intolerant, 0.31, intolerant	784 (43.1%)	No somatic lesions found	No HM specific malignant fusions found	No	No
	SPI1 (PU.1)	AD PEST ETS	Yes	No	Yes	NA	0.24	0.98	ETS	0.1261, highly intolerant	6 (1.86%)	No somatic lesions found	No HM specific malignant fusions found	No	No
	ERG	PNT ETS	Yes	No	Yes	NA	0.33	0.96	ETS	0.2526, intolerant	66 (3.42%)	ERG1, ERG2, ERG/TMPRSS2, ERG/EWS, ERG/FUS	ERG/FUS	No	No

Oncogenic potential of known BMF and/or HM predisposition TFs, as well as strong candidate BMF and/or HM predisposition TFs mapped using mouse phenotypes, constraint scores, variants observed and malignant fusions. Mouse phenotypes collated through Mouse MGI (25–27), and mousephenotype.org (28), constraint scoring (LOEUF, pLI and No. of LOF) extracted from gnomAD (29), domain constraint calculated from MetaDome (30), and somatic mutations obtained from COSMIC (31).
NA, not applicable.

TABLE 3 Candidate inclusions to BMF/HM predisposition gene lists.

Gene symbol	Canonical Transcript Coding size (aa)	TF Family	Human phenotype		Constraint				gnomAD No. of LOF (different/total)	COSMIC		Murine embryonic lethality		Malignant Fusions	
			Phenotype	Zygosity	LOEUF	pLI score	Domain	Average domain constraint		point mutations in Hematopoietic tissue/total samples (%)	No. of LOF	Het	Hom	Somatic	HM specific
ETV2	342	ETS	None reported	N/A	1.42	0	ETS	0.5563, SI	28/121	7/5924 (0.12%)	10/155 (6.45%)	No	Yes	None observed	None observed
ELF1	501	ETS	None reported	N/A	0.58	0.01	ETS	0.4664, I	17/19	53/8011 (0.66%)	39/659 (4.7%)	No	No	None observed	None observed
FEV	238	ETS	None reported	N/A	1.32	0.01	ETS	0.3402, I	5/7	4/5924 (0.07%)	4/99 (4.0%)	No	No	FEV/EWS	None observed
ETS1	485	ETS	None reported	N/A	0.39	0.78	ETS	0.221, I	7/7	136/6148 (2.21%)	32/1086 (2.9%)	No	No	None observed	None observed
ETS2	469	ETS	None reported	N/A	0.2	1	ETS	0.23, I	5/8	22/5924 (0.37%)	17/401 (4.2%)	No	Yes	None observed	None observed
IKZF2	526	IKAROS	None reported	N/A	0.29	0.99	ZF	0.22, I	6/157	71/6587 (1.08%)	35/1124 (3.11%)	No	No	None observed	None observed
IKZF4	585	IKAROS	None reported	N/A	0.38	0.9	ZF	0.11, HI	9/11	20/5924 (0.34%)	13/423 (3.07%)	No	No	None observed	None observed
GFI1B	330	Proto-oncogene	Bleeding disorder, platelet-type, 17*	AD, AR	0.86	0	ZF ZF ZF ZF	0.53, SI 0.51, I 0.34, I 0.7241, N	16/52	28/5924 (0.47%)	11/500 (2.2%)	No	Yes	None observed	None observed
TAL1 (SCL)	331	bHLH	Leukemia, T-cell acute lymphocytic	Somatic	0.7	0.57	BHLH	0.35, I	10/47	19/6120 (0.31%)	34/398 (8.54%)	No	Yes	None observed	None observed
LYL1	280	bHLH	Leukemia, T-cell acute lymphoblastoid	Not recorded	0.61	0.77	BHLH	0.50, I	3/4	7/5924 (0.12%)	8/127 (6.3%)	No	No	None observed	None observed
LMO2	227	bHLH	Leukemia, acute T-cell	Not recorded	1.028	0.16	LIM 1 LIM 2	0.54, SI 0.391, I	9/13	31/6037 (0.51%)	6/358 (1.68%)	No	Yes	None observed	None observed
MEIS1	390	Homeo-box	None reported	N/A	0.14	1	MEIS Homeo-box	0.19, I 0.045, HI	2/2	102/5924 (1.72%)	18/1094 (1.65%)	No	Yes	None observed	None observed

(Continued)

TABLE 3 Continued

Gene symbol	Canonical Transcript Coding size (aa)	TF Family	Human phenotype		Constraint					gnomAD No. of LOF (different/total)	COSMIC		Murine embryonic lethality		Malignant Fusions	
			Phenotype	Zygosity	LOEUF	pLI score	Domain	Average domain constraint			point mutations in Hematopoietic tissue/total samples (%)	No. of LOF	Het	Hom	Somatic	HM specific
TCF12	706	bHLH	Craniosynostosis 3, Hypogonadotropic hypogonadism 26 with/without anosmia	AD, AR	0.37	0.68	Helix-loop-Helix DNA BD	0.2837, I		24/29	113/6172 (1.83%)	67/1550 (4.3%)	No	Postnatal lethality within two weeks	TCF12/NR4A3	None observed
GFI1	422	Proto-oncogene	Neutropenia, severe congenital 2*	AD	0.56	0.25	ZF 1 ZF 2 ZF 3	0.2678, I 0.6286, SI 0.6491, SI		11/13	15/6073 (0.25%)	20/308 (6.5%)	No	No	None observed	None observed
FOS	380	FOS	None reported	N/A	0.65	0.26	bZIP TF	0.3376, I		6/6	8/5924 (0.14%)	16/211 (7.6%)	No	No	None observed	None observed
HOXA9	272	Homeo-box	None reported	N/A	1.62	0	Homeo-box Domain	0.4315, I		13/41	14/6233 (0.22%)	5 (2.33%)	No	No	None observed	HOXA9/NUP98, HOXA9/MSI2
FOSB	338	FOS	None reported	N/A	0.3	0.98	bZIP TF	0.16, HI		6/6	9/5924 (0.15%)	11 (3.45%)	No	No	None observed	None observed
FOXC2	501	Forkhead box	Lymphedema-distichiasis syndrome with renal disease and diabetes mellitus	AD	0.79	0.13	Forkhead Domain	0.21, I		6/7	3/5924 (0.05%)	5 (1.66%)	No	Yes	None observed	None observed
EKLF (KLF1)	362	KLF	[Hereditary persistence of fetal hemoglobin], Dyserythropoietic anemia, congenital, type IV, Blood group–Lutheran inhibitor, - Ineffective erythropoiesis	AD	1.13	0	ZF	0.83, N		16/34	5/5924 (0.1%)	2 (1%)	No	Yes	None observed	None observed

(Continued)

TABLE 3 Continued

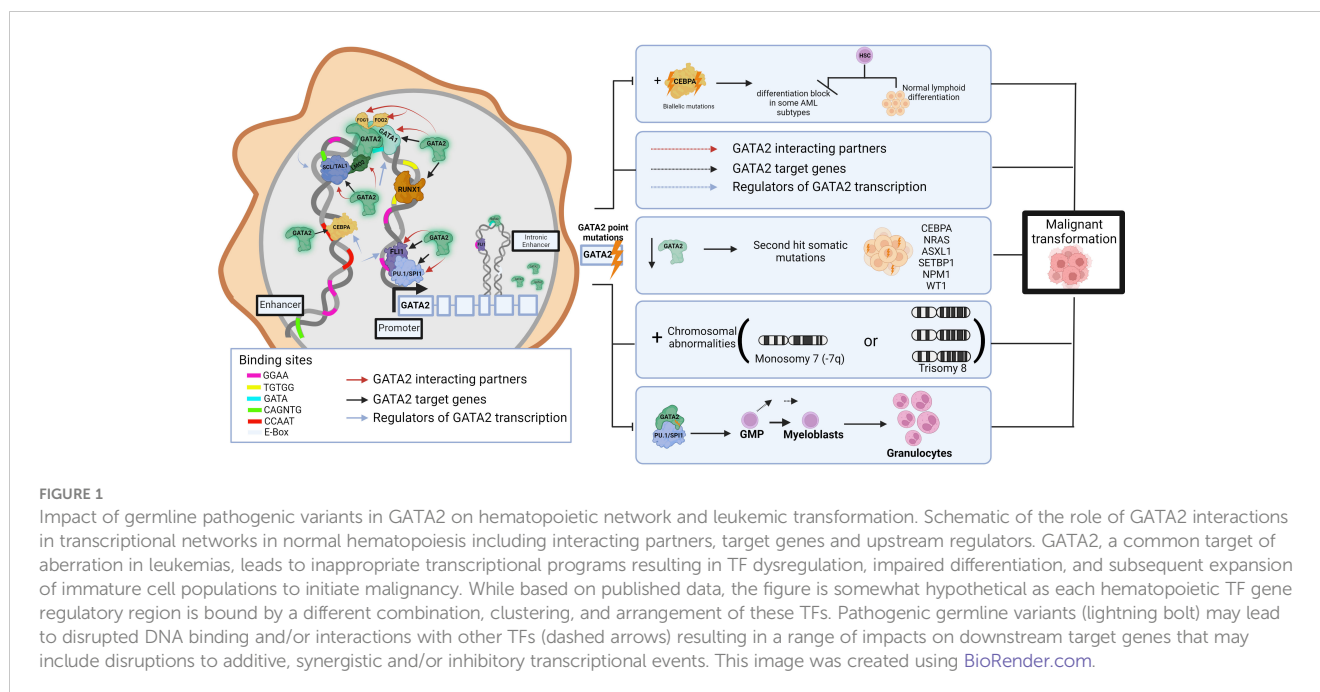| Gene symbol	Canonical Transcript Coding size (aa)	TF Family	Human phenotype		Constraint				gnomAD No. of LOF (different/total)	COSMIC		Murine embryonic lethality		Malignant Fusions	
			Phenotype	Zygosity	LOEUF	pLI score	Domain	Average domain constraint		point muta-tions in Hematopoietic tissue/total samples (%)	No. of LOF	Het	Hom	Somatic	HM specific
NFKB1	969	NF-kappaB	Immunodeficiency, common variable, 12,	AD	0.16	1	Ankyrin repeat 1, Ankyrin Repeat 2, Death Domain	0.52, I 0.57, SI 0.63, SI	10/82	67/6967 (0.96%)	18 (1.98%)	No	No	None observed	

*GFI1 (S36N) polymorphism is associated with incidence and/or prognosis of AML, MDS and multiple myeloma (32).
Candidate inclusions to BMF/HM predisposition gene lists, not found in high evidence Genomics England PanelApp panels: Hematological malignancies cancer susceptibility [Version 3.3] and/or Bleeding and Platelet Disorders [Version 1.16] + Cytopenias and congenital anemias [Version 1.111]. TFs are isolated by involvement in hematopoiesis, oncogenic potential, and previous disease association. Human phenotype is collated through OMIM. Slightly intolerant, SI; intolerant, I; highly intolerant, HI; neutral, N; loss-of-function observed over expected upper bound fraction, LOUEF.
NA, not applicable.

(40), and it is shown that the progression to malignancy can occur through multi-step secondary events targeting various hematopoietic pathways. These include (but are not limited to) acquired biallelic mutations in CEBPA (41, 42) which lead to differentiation blocks in some AML subtypes (38), NRAS, ASXL1, SETBP1, NPM1, and WT1 secondary mutations (43), chromosomal abnormalities (monosomy 7 or trisomy 8) (44)), disrupting the regeneration abilities of enhancers (45), HSC exhaustion following repeated infections, and changes in DNA binding affinities that result in disruptions to interaction partners/target genes. (Figure 1) For example, the GATA2 p.T354M missense variant may contribute to the leukemogenic process in two ways, by partial loss of GATA2 transactivation activity and simultaneous increased affinity to PU.1 (46), thereby potentially interfering with differentiation and driving cells towards granulocytic disease (7). Further, GATA2 p.R396Q leads to complete loss of DNA binding and transactivation, and inability to maintain the undifferentiated characteristics of HSCs (47). Single-cell transcriptomics identified 196 TF regulatory networks in AML patient cells and used this data to identify specific TF sub-networks that play a key role in determining differentiation trajectories in hematopoiesis (37, 48). For example, FLI11/ERG constitute a sub-network that promotes endothelial cell fate, whilst another RUNX1/GATA2 sub-network has been implicated in the promotion of cells to a hematopoietic cell fate (49). Cell-type transitions such as these are under tight transcriptional control, and loss of this control by any TF aberration can also lead to proliferation and leukemic transformation (38). Interestingly, analysis of heptad expression identified RUNX1, FLI1, LMO2, GATA2, ERG and LYL1 levels were significantly higher in groups of poorer overall survival, suggesting that high expression levels of the heptad contribute to the relative immaturity or stemness of the AML transcriptome (4, 39).

Evidently, networks can be disrupted *via* multiple pathways and result in dysregulation of the transcriptional network, a potential reason for phenotypic diversity observed in germline carriers. Given this knowledge it is intriguing to hypothesize that germline variants in genes of the core hematopoietic TF network may be implicated in leukemogenesis. Some are currently included in high evidence Genomics England PanelApp panels for HM and/or BMF susceptibility (i.e., *GATA2, RUNX1, FLI1, GFI1B*), while the significance of others (i.e., *ERG, SPI1* (encoding PU.1 protein)) remain to be determined.

Gene-specific mechanisms of oncogenicity

While TFs that predispose to BMF and/or HM may reside and act in common networks, no two TFs predispose to exactly the same phenotypes with the same hematopoietic disorder(s) incidence rate, average age of onset or propensity to solid cancers. Modern efforts to characterize genes for germline predisposition to HM has yielded a much larger genetic susceptibility than previously thought (50). Observed germline occurrence has so far allowed for the identification of predisposition germline mutations encoding

FIGURE 1
Impact of germline pathogenic variants in GATA2 on hematopoietic network and leukemic transformation. Schematic of the role of GATA2 interactions in transcriptional networks in normal hematopoiesis including interacting partners, target genes and upstream regulators. GATA2, a common target of aberration in leukemias, leads to inappropriate transcriptional programs resulting in TF dysregulation, impaired differentiation, and subsequent expansion of immature cell populations to initiate malignancy. While based on published data, the figure is somewhat hypothetical as each hematopoietic TF gene regulatory region is bound by a different combination, clustering, and arrangement of these TFs. Pathogenic germline variants (lightning bolt) may lead to disrupted DNA binding and/or interactions with other TFs (dashed arrows) resulting in a range of impacts on downstream target genes that may include disruptions to additive, synergistic and/or inhibitory transcriptional events. This image was created using BioRender.com.

master hematopoietic TFs such as RUNX1, CEBPA, GATA2, ETV6, PAX5 and IKZF1 (IKAROS). The majority of these segregate as autosomal dominant in nature. In the following section we review the known predisposition BMF and/or HM TF genes.
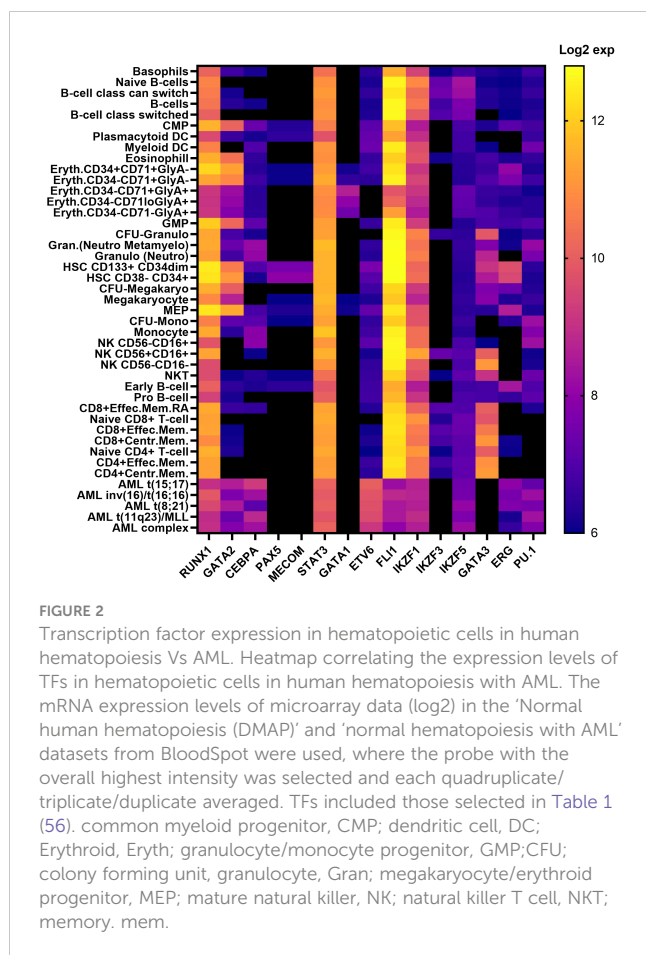
Best characterized TF BMF and/or HM predisposition genes: RUNX1, GATA2 and CEBPA

RUNX1, GATA2 and CEBPA are archetypal HM predisposition genes. First reported well over a decade ago, germline RUNX1 (51), GATA2 (46) and CEBPA (52) are the best characterized TF genes linked to HM predisposition. Unlike preleukemic syndromic features often associated with RUNX1 and GATA2, CEBPA variants predispose solely to AML without any recognized preleukemic phenotypes (53, 54). These three predisposition genes harbor point mutations and small indels (both somatic and inherited), resulting in leukemia, in addition to partial or whole gene deletions being observed in RUNX1 and GATA2-driven HM. RUNX1 also presents with over 70 translocation fusion partnerships in patients with HM (55). RUNX1, GATA2 and CEBPA all reside in similar hematopoietic networks, impacting common and different regulatory pathways, however each play unique roles during hematopoiesis and in cell fate decisions, despite the expression levels of RUNX1 in hematopoietic cells presenting significantly higher than both GATA2 and CEBPA (Figure 2) (14–19, 57). RUNX1 binds to the TGTGGT DNA-binding motif as a heterodimer with core binding factor beta (CBFβ) and can activate or repress gene expression depending on its cellular context and in response to extracellular signals from the microenvironment. This includes regulation of chromatin accessibility and its role in recruiting and enacting transcriptional programs with other transcriptional co-factors (58–60). CEBPA and GATA2 are

predominantly myeloid transcription factors which bind to CCAAT and GATA motifs (Table 1), respectively, but have distinct roles in regulating myeloid differentiation. (Table 1) (14, 15, 18–20).

Pathogenic germline variants contributing to RUNX1/GATA2-driven disease are largely premature terminations throughout the coding region or missense variants, clustering in the DNA binding runt homology domain (RHD) of RUNX1 and C-terminus (zinc finger (ZF) 2 domain) of GATA2, predominantly resulting in haploinsufficiency (53, 54). Interestingly, no pathogenic germline GATA2 single nucleotide variants have been described within the N-terminus, except for two rare variants, p. T117= (c.351C>G) (61–63) and p.A286V (c.857C>T) (64, 65), both of which generate strong cryptic splice sites and are better classified as truncating splice variants leading to nonsense mediated decay (53). There are also very rare individual cases of variants in ZF1 (p.H313Y, p.L315P, p.A318T) (53), which intriguingly is the most common site of somatic mutations (66). Strikingly, there is a clear separation of GATA2 germline (ZF2) and somatic (ZF1) missense variants in HM, suggesting different mechanisms of leukemogenesis in germline GATA2 HM. In contrast, distinction of germline RUNX1 and sporadic HM variants is challenging as identical variants are observed in both. Germline deletions of RUNX1, only affecting the RUNX1c isoform support this as being the predominant oncogenic isoform (9).

Germline CEBPA mutations are most commonly frameshift, protein-truncating variants at the N-terminus with rare families harboring C-terminal in frame insertions or deletions (54, 67). The N-terminal premature termination prevents generation of the full-length 42 kDa protein but preserves translation of the smaller 30 kDa isoform (68). These mutations display a high degree of penetrance for AML (90%) compared to C-terminal in-frame indels that display lower penetrance (50%) (52, 54, 68). Individuals with a germline CEBPA N-terminal mutation at AML

FIGURE 2

Transcription factor expression in hematopoietic cells in human hematopoiesis Vs AML. Heatmap correlating the expression levels of TFs in hematopoietic cells in human hematopoiesis with AML. The mRNA expression levels of microarray data (log2) in the 'Normal human hematopoiesis (DMAP)' and 'normal hematopoiesis with AML' datasets from BloodSpot were used, where the probe with the overall highest intensity was selected and each quadruplicate/triplicate/duplicate averaged. TFs included those selected in Table 1 (56). common myeloid progenitor, CMP; dendritic cell, DC; Erythroid, Eryth; granulocyte/monocyte progenitor, GMP;CFU; colony forming unit, granulocyte, Gran; megakaryocyte/erythroid progenitor, MEP; mature natural killer, NK; natural killer T cell, NKT; memory. mem.

diagnosis, often acquire C-terminal CEBPA mutations within the bZIP region (predominantly missense or in-frame indels), highlighting the synergistic effect of these lesions and their selection during clonal expansion (69). In <10% of cases however, homozygous single N- and C- terminal mutations have been reported, arising from copy neutral loss of heterozygosity (69). Like RUNX1, CEBPA germline and somatic variants may look identical, and a germline tissue sample or familial segregation with disease is necessary for confirmation of germline origin.

The dysregulation of RUNX1, GATA2 and CEBPA activity underlies hematologic pathologies, yet the biological mechanisms that establish and maintain its contextually distinct expression patterns are versatile. GATA2 and RUNX1 have been shown to function as both "pioneer" and "master" regulatory TFs in hematopoiesis, opening chromatin for easy access by other TFs and recruiting other TFs to enhance downstream effectors and signaling pathways (70). Their direct associations with other important hematopoietic TFs such as with each other, TAL1 and PU.1 (Table 1) (12, 13), may highlight a mechanism by which missense variants or reduced levels of the wildtype (haploinsufficiency) protein disrupt the stoichiometry required for normal hematopoietic expansion and differentiation processes leading to microenvironments that are conducive to development of cytopenia's or initiation, maintenance and/or progression of malignancy. Complete deletion of either the RUNX1 or GATA2 locus suggests haploinsufficiency as the mechanism of

predisposition; however, not all variants result in complete LOF of one allele. For example, GATA2 p.T354M and RUNX1 p.R204Q missense pathogenic variants display not only partial loss of transactivation activity, but also act in a dominant negative manner (46, 71). It is not clear whether the dominant negative action further decreases the total cellular activity levels in these partial LOF variants to approach that of the complete LOF situation.

Amidst germline RUNX1, GATA2 and CEBPA variants setting the basis for autosomal dominant predisposition to leukemogenesis, the considerable clinical heterogeneity in disease progression (even within families) suggests they are not transformation-sufficient, and that acquired secondary mutations are required for initiation and maintenance of malignancy (72) and may dictate the type of malignancy and its response to therapies. Ironically, the number of LOF mutations found in Catalog of Somatic Mutations in Cancer [COSMIC (73)] for RUNX1 and CEBPA surpass other hematopoietic TFs by hundreds (Table 2), coinciding to the hypothesis that the acquisition of somatic mutations in these oncogenic genes strongly drives leukemic transformation. The selective advantage of clones is often shaped by the acquisition of new aberrations in distinct genes. For example, STAG2 mutations act as a driver of clonal hematopoiesis with limited leukemic potential in GATA2 deficiency patients, but are rarely seen outside of this context, while SETBP1, RUNX1 and RAS pathway mutations along with monosomy 7 are associated with leukemic transformation (74). Molecular analysis of GATA2 and RUNX1 deficiency by germline predisposition found that the most frequent cytogenetic abnormalities for GATA2 involved monosomy 7 or trisomy 8, and the aggregation of second hit somatic mutations were unique to each gene with GATA2, targeting most frequently ASXL1, NRAS/KRAS, WT1, STAG2, and SETBP1, whereas in RUNX1, most frequently biallelic RUNX1 variants, followed PHF6, BCOR, WT1, and TET2, leading to leukemic transformation (54, 72, 74). Acquisition of somatic RUNX1 variants appears to be a late event associated with disease transformation, with clonal hematopoiesis (driven by BCOR, TET2) observed in pre-leukemic carriers.

Curiously, RUNX1, GATA2 and CEBPA display loss-of-function (LOF) intolerant (pLI), LOF observed/expected upper bound fraction (LOEUF) and domain constraint (metadome) scores no higher than many other TFs associated with hematopoiesis, (Figure 3, Table 2) (29, 30) suggesting that it is likely that other TF genes remain to be identified in BMF and/or HM predisposition.

The ETS family of BMF/HM predisposition TFs

ETS TFs, are a family of TFs which share a winged helix-turn-helix DNA binding domain (ETS domain), which recognize the DNA sequence GGAA/T. ETS TFs (e.g. ETS1, ETS2, ELF1, ERG, ETV2, ETV6, FEV, FLI, PU.1) are involved in the regulation of a variety of genes in hematopoiesis; however despite this involvement, only ETV6 (also known as TEL) and FLI1 have

FIGURE 3

Pathogenicity predictors for TFs implicated in predisposition to BMF and HM. Comparison of pathogenicity predictor score for known predisposition TFs and potential TF genes isolated from Table 1. **(A)** The Probability of being loss-of-function (LOF) intolerant (pLI) score of each gene was collated using the gnomAD database (version 2.1.1). The pLI score reflects the tolerance of a given gene to the LOF based on the number of protein-truncating variants referenced in control databases weighted by the size of the gene and the sequencing coverage. The pLI score ranges from 0-1, where higher the score, the higher the intolerance of the gene. **(B)** Percentage of gene LOF variants in COSMIC database. The total number of somatic LOF variants within a particular gene were calculated by totaling positive mutation data for the selected gene. Variants called 'LOF' included nonsense substitutions, frameshift insertions and frameshift deletions. Percentage calculated using LOF variants over the total number of unique samples of each gene. **(C)** Percentage of point mutations observed in hematopoietic and lymphoid tissues. The distribution of mutations across the primary hematopoietic and lymphoid tissues curated by COSMIC were collated. The percentages were calculated by totaling the number of point mutations of each gene, over the total samples tested.

been included on germline targeted sequencing panels for bleeding and platelet disorders and HM predisposition (1). ETV6 primarily functions as a transcriptional repressor, targeting a wide spectrum of genes, many of which are highly regulated during hematopoiesis, whilst FLI1 plays an essential role in embryogenesis, vascular development and megakaryopoiesis (75–77). Interestingly, germline pathogenic variants in both TFs co-segregate with autosomal dominant bleeding disorders (i.e., mild thrombocytopenia), although ETV6 also presents a HM risk in ~30% of carriers while FLI1 is most commonly autosomal recessive (78–81). As a consequence of the important function of other ETS TFs (e.g. ETS1, ETS2, ELF1, ERG, ETV2, FEV, PU.1) in hematopoiesis (82), including the role of ERG in definitive hematopoiesis, adult HSC function and platelet maintenance (83), and the role of PU.1 in positive regulation of genes in the macrophage, granulocyte, dendritic-cell and B-cell lineages, (Table 4) (84) it is probable that germline genetic changes involving these TFs will also be found to predispose or contribute to aberrant hematopoiesis including cytopenia's and/or HM (85). As with the intron 4 enhancer variants in GATA2 (86), causal genetic variants may well be in gene elements such as promoters, enhancers and/or suppressors that are used or activated in a cell-specific manner at particular stages of hematopoietic lineage development and maturation. In keeping with this, using Genome-wide association studies (GWAS) for some cancers such

as breast cancer, common genetic variants have been identified in intergenic and intronic regions and shown to impact on important regulatory regions such as promoters and enhancers of known and novel genes important for the cancer initiation or development as well as multi-exonic non-coding RNA (mencRNA) genes (87, 88). It is conceivable and likely that rare variants may similarly exist that have even greater impact on cancer including HM predisposition, and such regulatory regions are not screened in WES, not included in panels, and not interrogated or not interpretable in WGS. To date, detection of germline variants for most hematopoietic and HM genes including TF genes have been confined to the coding regions with exceptions such as GATA2 intron 4 enhancer (86), ANKRD26 5'UTR (89) and TERT promoter (90).

ETV6 germline aberration due to rearrangements, fusions, mutations, or deletions resulting in monoallelic expression of ETV6, contributes to several types of myeloid and lymphoid malignant susceptibility, with approximately two-thirds being B-ALL (85, 91). ETV6 is involved in over 30 translocation partnerships in leukemia and MDS including fusions with PDGFRB, AML1, MN1, JAK2, ASC2, ABL2, BTL, ARNT, MDS2, PER1 and an ETV6-RUNX1 fusion seen in over 22% of childhood B-ALL (79, 85). Subsequent to the translocation in B-ALL cases, the wildtype ETV6 allele is often mutated or deleted implicating its tumor-suppressive function (79, 92). Ironically, other ETS TFs involved in hematopoiesis (e.g., ERG, FEV) are also seen in non-

TABLE 4 Strong candidate inclusions to BMF/HM predisposition gene lists.

TF gene symbol	TF Family	DNA Binding Motif	Reported Human phenotype (Hemato-logical)		MGI Mouse Phe-notype (Hemato-logical)	Role in hema-topoiesis (downstream differentiation)	Known Interact-ing TF partners (STRING)	Target genes (TFs)
			OMIM	MONDO	Phenotypes			ENCODE and ChEA Consen-sus TFs from ChIP-X
GATA3	GATA	WGATAR	N/A	N/A	Abnormal T cell physiology and differentiation Abnormal definitive hematopoiesis Abnormal erythrocyte morphology	T Cells, HSC, NK cells, B cells, CLP	TAL1, IL5 IL33, IL33, STAT6, TBX21, FOXP3, IL13, SMAD3, ESR1, FOXA1	GATA1, SALL4, KLF4, MYC, E2F1, REST, TRIM28, SPI1, STAT3
SPI1 (PU.1)	ETS	GGAA	Decreased/ absent circulating B cells Impaired B-cell development Lymphopenia Neutropenia	N/A	Transient neutropenia Abnormal granulocyte, neutrophil, macrophage, blood cell and B/T cell differentiation Increased hematopoietic stem cell proliferation/ abnormal definitive hematopoiesis Abnormal common myeloid progenitor cell morphology Abnormal erythropoiesis, leukopoiesis	Monocyte, B cells, GMP, CLP, CMP, Granulocyte	IRF4, GATA1, CSF3R, EP300, JUN, TBP, GATA2, IRF8, CEBPE, CEBPA	SPI1, ELF1, FLI1, GABPA, TAF1, MAX, CREB1, UBTF, ATF2
ERG	ETS	GGA(A/T)	N/A	AML	Abnormal embryonic erythropoiesis Thrombocytopenia Decreased leukocyte HSC cell number Pancytopenia	No differentiation - specific role	TMPRSS2, AR, SPI1, RUNX2, CBFB	SPI1, TAF1, YY1, MAX, NFYB, UBTF, BRCA1, E2F1, MYC

Candidate inclusions to BMF/HM predisposition gene lists, ranked strong due to known role in hematopoiesis, previous disease association, pathogenicity predictor scores, and gene ontology. TF demonstrates similarities to known BMF/HM predisposition genes in reported human/mouse phenotypes, roles in downstream hematopoietic differentiation (15, 18–21) and known interacting TF partners and target genes. Known interacting partners collated through STRING (12) and target genes through ChEA from Chip-X (13).
NA, not applicable.

hematological malignant fusions (ERG/TMPRSS2, ERG-EWS, FEV-EWS) (i.e., prostate cancer) as well as hematological malignant fusions (e.g., TLS/FUS-ERG and FLI1-EWS seen in myeloid leukemia and Ewing's Sarcoma, respectively) (93). Germline ETV6 pathogenic variants phenocopy RUNX1 germline pathogenic variants in terms of platelet defects, heightened HM predisposition and their association with a poorer overall survival but, unlike RUNX1, clonal hematopoiesis has not been reported in ETV6 carriers (94, 95). Experimental studies including RNA sequencing indicate that despite patients harboring germline ETV6 or RUNX1 pathogenic variants having similar clinical phenotypes, distinctive molecular mechanisms occur to generate haploinsufficiency (94).

The majority of clinically reported ETV6 and FLI1 loss-of-function (LOF) pathogenic variants (ClinVar) cluster in the highly conserved ETS DNA binding domain (85). In the normal population, both ETV6 and FLI1 loss-of-function variants are not well tolerated (gnomAD, LOEUF 0.12 and 0.09, respectively) (29), consistent with them being pathogenic. (Table 2, Figure 3) Notably, because of the high conservation of their ETS domains, in regions of

similarity, genetic changes may impact on structure and hence function such that variants demonstrated to be pathogenic in one ETS TF are likely also to be pathogenic at the corresponding amino acid position in another. Interestingly, pathogenic variants at 4 amino acids in FLI1 (ClinVar; p.R324W, p.R337Q/W, p.R340H/P, p.Y343C) have identical amino acids in wildtype ERG, ETV2, FEV and ETS1; identification of similar changes at the corresponding amino acids in these or other ETS TFs would be predicted to have similar deleterious effects.

Functional evaluation of germline ETV6 and FLI1 variants from leukemia and thrombocytopenia patients suggests that the substitution of conserved residues disrupts the protein's general function. Using in vitro biochemical assays, it was found that damaging ETV6 variants resided within, or caused truncation of, the ETS domain, exhibiting significant impairment of transcription repression, a decreased ability to bind DNA and a loss of nuclear localization (76, 79, 85). Consistent throughout literature, pathogenic FLI1 variants also repress transcriptional activity in validated FLI1 target genes including GP9, exhibiting predominantly cytoplasmic localization and in contrast, result in

large, fused platelets that have electron-dense α-granules, characteristic of the Paris-Trousseau syndrome (78, 80, 81).

The oncogenic potential of ETS TFs PU.1, ERG and ETS2 have been well characterized, making them convincing candidate genes for BMF and/or HM predisposition (Table 4, Figure 3); however, none are yet an established disease gene nor hold strong or definitive (ClinGen) gene-disease relationships. The importance of these TFs in normal hematopoiesis has been well established, both in lymphoid and myeloid lineages, however, studies have suggested these same TFs may also have a potential role in leukemogenesis. Cohort studies have identified the overexpression of ERG and ETS2 as a biomarker correlated with an adverse clinical outcome in AML patients, with ETS2 hypothesized to induce apoptosis (in the presence of p53) and ERG crucial for leukemic maintenance (96–100). Conversely, creating a model of hypomorphic PU.1 established a correlation between low PU.1 expression and AML (101, 102). This suggests that increases or decreases in the critical threshold levels of various ETS TFs may contribute to AML pathogenesis.

Like many other HM TF predisposition genes, the mechanisms of germline ETV6 and FLI1-mediated leukemogenesis remains poorly understood. Although, their direct associations with other important hematopoietic TFs such as with GATA2 and RUNX1 (Table 1) may highlight a mechanism by which missense variants disrupt the expression pattern of TFs within the normal hematopoietic network (12, 13, 96)

The IKAROS family of BMF/HM predisposition TFs

The IKAROS protein family (IKZF) are master mediators of cell differentiation and function *via* DNA binding (ZF1-4) and dimerization domains (ZF5-6) which orchestrate transcriptional repression and/or activation of a large number of genes. IKZF1 and IKZF3 encode for IKAROS and AIOLOS proteins, and amongst other hematopoietic roles, are critical regulators of lymphoid development and differentiation. Germline variants in IKZF1 and IKZF3 result in a broad range of phenotypes including hematological ones which include immunodeficiency disorders (103) and most commonly lymphoid leukemias (104, 105). While dysfunctional IKZF2 is implicated in B-ALL and T-cell lymphoma, only IKZF1, IKZF3 and IKZF5 are formally recognized as BMF and/or HM predisposition genes (104, 106, 107). Interestingly, all three Ikaros BMF/HM predisposition genes are not highly expressed in hematopoietic cells in normal hematopoiesis or AML, indicating no correlation between expression patterns and predisposition to BMF or malignancy. (Figure 2) The potential of causal germline genetic variants to be discovered in other Ikaros family members (*i.e.*, IKZF2 and IKZF4) remains plausible, with high constraint scores (average pLI; 0.93, Table 2, Figure 3), despite IKZF4 poorly or not expressed during hematopoiesis (56).

Observed reoccurrence, familial segregation, and functional evaluation of IKZF1 variants enabled its identification as a HM predisposition gene. Somatic IKZF1 genetic aberrations have adverse effects on clinical outcomes (overall survival and relapse-free survival) and several molecular mechanisms have been implicated as the mode of IKAROS deficiency in both autoimmune deficiency syndromes and HM manifestation (108, 109). These include missense variants at DNA contact residues that are still able to dimerize with wildtype IKAROS but are unable to bind target DNA (103, 110), haploinsufficiency with deletion of one IKZF1 allele (103), differential expression in gene networks involved in cancer, cell signaling, apoptosis, and hematopoiesis (111) and transcription, dimerization, subcellular localization, and cell adhesion LOF (107, 108). Interestingly, somatic mutations in IKZF1 are recurrently detected in IKZF1 germline carriers with B-cell precursor ALL, suggesting that cells carrying the germline variant favor acquisition of a second (i.e., biallelic) IZKF1 transforming event (112). Curiously, the core consensus DNA binding motif for IKZF TFs is "GGAA" which is similar to that of ETS TFs, although the surrounding nucleotides confer greater specificity for specific IKZF and ETS TFs. Hence, at some binding sites there may be competition or synergism between some or many of these TFs depending on their expression levels and DNA promoter/enhancer contexts. Interestingly, in T-ALL cells, IKZF1 binding sites across the genome are most closely associated with FLI1 (similar GGAA consensus) and RUNX1 binding sites (113). Consistent with this, germline RUNX1 variants can also predispose to T-ALL as well as other lymphoid malignancies.

IKZF proteins form homodimers and heterodimers *via* their two C-terminal ZFs to initiate nucleosome remodeling and enable their pleiotropic roles in various hematopoietic cells. It is this phenomenon that sparks the possibility that aberration in any IKZF gene may contribute to leukemogenesis by perturbing the transcriptional networks of IKZF genes normal functions including lymphoid maturation, tumor suppression, cell-cycle regulation, kinase signaling and chromatin modification (108). Early reports associated germline IKZF1 deletions with poorer treatment responses and unfavorable outcomes (114–116), while more recent studies found the converse to be true, with beneficial responses to induction therapy (117).

PAX5, MECOM, GATA1 and STAT3 BMF/HM predisposition TF genes

PAX5, MECOM and GATA1 are also critical contributors to cooperative transcriptional networks involved in hematopoiesis and included on germline targeted sequencing panels for BMF and/or HM susceptibility. These TFs demonstrate extremely low tolerance to LOF, with pLI scores all above 0.95, domain constraints scored intolerant or higher (30), and limited LOF seen in databases reflecting the normal population (gnomAD) (Table 2, Figure 3) (29). The oncogenic potential of these TFs is established by impactful somatic lesions, malignant fusions, and embryonic lethality.

PAX5 is a master regulator of lymphopoiesis, being required for normal B cell development and differentiation including B-lineage commitment, maintenance of B cell identity and VHDJH recombination. Somatic PAX5 aberrations are evident in ~30% of sporadic B-ALL cases (118) with mechanisms including copy number alterations (often deletion including -9/-9p), translocations

generating fusion proteins that retain the DNA binding paired domain and nuclear localization signal of PAX5 (results in a fusion TF that acts as a dominant negative), intragenic amplification (including direct head-to-tail concatenation of exons 2 to 5 resulting in an extra 4-5 copies of the DNA binding and octapeptide domains), disruption to normal alternative splicing and isoform levels, and point mutations (119). The majority of these are believed to decrease PAX5 expression and/or transcriptional activity causing both repression and activation of downstream target genes (120). In almost all B-ALL cases with point mutations, there is a concomitant loss of the wildtype allele. However, complete loss of PAX5 activity is not seen in B-ALL suggesting that a residual amount of activity is required for the disease phenotype.

PAX5 has few known germline aberrations. The p.G183S missense variant, reported in three unrelated families and 2 sporadic cases, is located in the octapeptide domain, and strongly linked to pre-B cell leukemogenesis with impaired capacity for transcriptional activation and/or repression (118, 121–124). PAX5 variants follow autosomal dominant transmission with variable penetrance, suggesting it is likely that additional secondary or tertiary predisposing elements occur in these families, particularly somatic loss of the wildtype (WT) PAX5 allele due to aberrations of chromosome 9p (118, 121). Incomplete penetrance seen in PAX5 mutated families is likely explained by the synergistic effects of these secondary or tertiary events, including environmental factors since, in heterozygous Pax5 mice, infection exposure mediated HM development (123). Interestingly, only one other PAX5 germline missense variant (p.R38H) has been linked with leukemic predisposition (B-ALL). Curiously, 2 of 3 germline (p.R38H) cases acquired a PAX5 (p.R140L) mutation (125) which is in keeping with somatic p.R38H and p.R140L mutations co-occurring in a biallelic fashion in 10/11 B-ALL patients harboring a p.R140L mutation (125). Functionally, p.R38H is not able to regulate PAX5 target genes nor trigger B-cell differentiation, maintaining PAX5 cell growth properties only partially, without overt dominant-negative effect on the normal PAX5 function (125, 126). Presumably, the combination of these 2 concurrent variants provides PAX5 activity at levels highly conducive to HM. Somewhat like germline RUNX1 and CEBPA-driven HM, and even more so, in germline PAX5-driven B-ALL, the WT allele of PAX5 is almost always deleted or mutated in some way to abrogate PAX5 activity from this allele, in addition to the perturbation due to the inherited germline variant allele (119).

MECOM is a ZF TF gene which through differential splicing encodes three protein isoforms: MDS1, EVI1 and fusion protein MDS1-EVI1 (127). Typically, EVI1 has repressor functions while MDS1-EVI1 acts as a transcriptional activator. MECOM is a stemness gene, required for long-term HSC survival and is essential for regulating both embryonic and adult HSCs, by directly regulating GATA2. Germline MECOM variants cause heterogenous BMF syndromes and only recently was identified as a predisposition gene to inherited HM (128). Haploinsufficiency of MECOM, leads to loss of HSC within the first few months of life, *via* defects in stem cell self-renewal proliferation and repopulation capacity, causing severe neonatal BMF (129, 130). The location of

pathogenic germline variants is somewhat associated with the phenotypic variability observed in patients. Germline variants causing haploinsufficiency including nonsense, frameshift variants and *de novo* microdeletion (covering entire locus or just the MDS1 coding regions) causes congenital thrombocytopenia but not skeletal abnormalities (131) whereas missense variants in the ZF2 domain affecting the MDS1-EVI1 and EVI1 transcripts cause limb defects and thrombocytopenia including radioulnar synostosis with amegakaryocytic thrombocytopenia 2 (RUSAT2; OMIM #616738). These missense variants cause partial LOF or gain-of-function (GOF) (132). Missense mutations in MECOM typically cluster within the C-terminal ZF domain which contains an ETS-like motif, affecting the protein folding stability or DNA-binding amino acid residues. Transactivation studies have demonstrated varying effects of these missense variants on transcriptional regulation with a possible dominant-negative effect on TF AP-1 (Jun/Fos) signaling, and partial LOF of TGF-beta signaling (132). The C-terminal ZF domain also includes an oligomerization domain required for homodimerization with EVI1 and interaction with RUNX1.

Chromosomal rearrangements targeting MECOM are relatively common in sporadic myeloid leukemias. These include inversion inv(3)(q21.3q26.2) and translocation t(3;3)(q21.3;q26.2) which result in EVI1 overexpression (133). Translocation t[3;12] leading to RUNX1-EVI1 fusion and t[3;12] resulting in ETV6-EVI1 fusion both result in EVI1 overexpression. Interestingly, and not surprisingly, both the inv(3)(q21.3q26.2) and t(3;3)(q21.3;q26.2) rearrangements converge on the GATA2 gene, resulting in bringing the GATA2 enhancer close to the MECOM gene which leads to ectopic MECOM expression and GATA2 haploinsufficiency (134). The RUNX1-EVI1 fusion preserves the DNA-binding RUNT domain of RUNX1, with loss of the transactivation domain. Expression of RUNX1-EVI1 in progenitor cells committed to the hematopoietic lineage results in activation of a pan-lineage hematopoietic gene expression program, and a loss of regulation of vascular differentiation programs and partial cell-cycle arrest, likely affecting the terminal differentiation potential of multipotent progenitors (135). This fusion protein acts as a dominant-negative, suppressing the transactivation capacity of WT RUNX1 (136).

GATA1 is a ZF TF and a master regulator of erythroid development. GATA1 variants cause a range of hematological phenotypes, including X-linked thrombocytopenia with or without dyserythropoietic anemia, with several different syndromic manifestations including Diamond Blackfan Anemia, β-thalassemia and congenital erythropoietic porphyria. As is observed with other TFs involved in hematological disease, location within the gene and variant type is associated with clinical presentation. As GATA1 is located on the X-chromosome, loss-of-function variants are typically more severe in males with females often being asymptomatic although presentation is impacted by levels of X-chromosome inactivation skewing. (Table 1) (137–139) Two main classes of GATA1 variants exist: 1) splicing or start-loss, which lead to expression of a shorter isoform protein (i.e. GATA1s - without N-terminal transactivation domain), which can give rise to moderate to severe anemia, neutropenia and/or DBA-like phenotypes (138, 140, 141) and 2) missense variants in exons 3 and 4, most commonly in

the N-terminal ZF domain which mediates the interaction with co-factor Friend of GATA1 (FOG-1) giving rise to cytopenia-related phenotypes (137, 138, 142). These missense mutations have been shown to affect the interaction with FOG-1 or the DNA binding ability of GATA1 (139, 143). GATA1 polymorphisms have also been shown to act as genetic modifiers in disorders caused by variants in GATA1-dependent genes (144). Acquired GATA1 mutations resulting in expression of GATA1s are common in Down Syndrome patients with trisomy 21, causing transient abnormal myelopoiesis that may spontaneously resolve or progress to AML. Interestingly, two families with germline GATA1 variants resulting in expression of GATA1s had several family members who developed acute megakaryoblastic leukemia with acquired trisomy 2, suggesting the order of mutation acquisition is not important (145).

STAT3 is a TF critical for appropriate cell proliferation, inflammation, differentiation, survival, and an important mediator of the innate and adaptive immune response (146, 147). STAT3 is the most common member of the STAT protein family to be mutated in hematopoietic cancers. GOF and LOF germline variants have been identified and are observed to give rise to immunodeficiency, autoimmune and cancer phenotypes. Germline STAT3 LOF variants are responsible for autosomal-dominant hyper–immunoglobulin E syndrome while GOF/activating variants are causal for early-onset multiorgan autoimmunity (148) including lymphoproliferation and pediatric large granular lymphocytic leukemia (LGL) including neutropenia, thrombocytopenia (149). Germline variants in GOF STAT3 have been identified in multiple domains including the all-alpha, DNA binding, SH2, and C-terminal transactivation domain of the protein (149). The high rate of concurrent self-reactive autoimmunity and LGL in carriers of STAT3 GOF variants has led to the concept that they develop *via* similar molecular mechanisms (150–152). Other common syndromic features include interstitial lung disease, diabetes, and postnatal growth failure (153). Somatic GOF variants in STAT3, which cluster in the Src homology 2 (SH2) domain, are associated with granular lymphocytic leukemia, myelodysplastic syndrome, and aplastic anemia (150, 154). STAT3 is activated by multiple signaling pathways, beyond the canonical JAK-STAT pathway, including *via* cytokine signaling, receptor tyrosine kinase signaling, and G-protein coupled receptor signaling, to activate or repress transcription of target genes, explaining the broad functional role of STAT3 in regulating cellular functions (155, 156). Given the GOF are observed in all functional domains of the protein, it is thought that different variants in different locations of the protein will exert GOF effects *via* different steps of the signaling pathway for example, dimerization, DNA-binding, nuclear shuttling, or phosphorylation which could account for the phenotypic variability (155).

Biological mechanisms of BMF and/or HM predisposition TFs

TFs are key regulators of hematopoietic reprogramming, with each TF exhibiting specific temporal and spatial expression patterns during lineage differentiation. The TF sub-networks are tightly regulated, functioning in specific cell types to modulate hematopoiesis. As a result of this, and potentially with the influence of both internal and external stressors, the biological mechanisms by which distinct TF aberrations functionally influences normal hematopoiesis is highly variable. Biological models are often relied upon to elucidate the role of TFs in hematopoietic pathways, including identifying interacting partners and target genes, to gain insight into how aberrations contribute to the development of disease. It is possible that a variant may only display activity under certain physiological or pathological contexts and investigating its function in traditional biological models (e.g., iPSC cells, primary cell cultures, animal models) may not unveil salient information without knowledge of these external factors. Germline TF pathogenic variants add another layer of complexity by their non-autonomous dysregulation in hematopoietic cells, and therefore, functional consequences may only become apparent in heterocellular populations (157). Nevertheless, biological models can be very useful in many contexts. Given other comprehensive reviews on penetrant HM predisposition genes RUNX1, CEBPA, PAX5 (9, 10, 158) and the fact that the ZF and the ETS DNA binding domains are two of the most prominent functional domains associated with BMF and/or HM predisposition TFs, this section will discuss the biological mechanisms of TFs containing these functional domains.

Zinc finger BMF/HM predisposition TFs

The GATA and IKAROS families and MECOM isoforms all share structural ZF domains which function in both DNA binding, domain structure and protein: protein interactions. However, despite relatively high amino acid homology between individual members of each group, they display complex spatial, temporal and cell-type differences in expression, binding partners, role in regulation of hematopoietic networks, and underlying biological mechanisms contributing to BMF and/or HM. Gata1(-/-), Gata2 (-/-), Evi1(-/-), and Ikzf1(-/-) null mice all display embryonic lethality due to hematopoietic defects, highlighting their importance in hematopoiesis. Further, in heterozygosity, milder hematopoietic phenotypes demonstrate haploinsufficiency on prenatal definitive hematopoiesis and the function of adult HSCs (159–165).

GATA2 is integrated into a regulatory network that includes DNA-binding, interaction with numerous cofactors, genetic and epigenetic transcriptional regulation of hundreds of GATA2 target genes, and response to cellular and extracellular signals controlling its expression in positive and negative feedback loops. GATA2 directly regulates the expression of multiple target genes such as RUNX1, TAL1, SPI1, FLI1 and LMO2 (166), and different enhancers regulate specific spatial and temporal hematopoiesis. This regulation varies greatly in different biological systems and may only operate in restricted physiological and/or pathological contexts and states (e.g., progenitor cell vs mature cell or steady state vs stressed state) (167, 168). Consequently, germline variants in GATA2 may exhibit hypomorphic characteristics under some conditions, but not under others. For instance, patients with GATA2 deficiency syndrome may develop different syndromic

features under different stresses such as BMF or lymphedema with persistent infections or inflammation, myeloid malignancy with biological (e.g., cytosine deamination with aging) or chemical-induced acquisition of pathogenic mutations, hearing loss due to disruption in development of ear structures due to aminoglycoside antibiotic or stochastic events leading to urinary system malformation (169). A delicate balance of ebbs and flows of GATA2 activity is required to establish and maintain GATA2-dependent regulatory networks throughout normal and stress-induced hematopoiesis or other bodily systems (168).

The inhibition of GATA2 has been shown to be linked to biological traits of leukemogenesis in multiple systems such as human cord blood where knockdown of GATA2 markedly reduces colony forming cell growth capacity and lineage-specific colony forming cells (170). Gata2 depletion also reduces multi-lineage potential in mice, as well as HSPC cell numbers (171). Haploinsufficiency does not completely diminish hematopoietic programming but reduces the capacity of HSPCs to complete HSC maturation during the endothelial to hematopoietic transition, where Gata2(+/-) is able to initiate the process, but unable to fully execute it (172). Knockout mice exhibit near entire loss of HSCs, primitive progenitor cells and committed myeloid and erythroid progenitors (170), whilst GATA2(+/-) mice show markedly reduced HSC numbers suggesting dose-dependence. However, due to spatial and temporal expression patterns, some cells still achieve their threshold of GATA2 activity, and thus their target genes may still be appropriately regulated. Similarly, in GATA2-deficient iPSC systems, hematopoietic progenitors (HPCs), and erythroid and granulocyte progenitors are markedly reduced, despite having little effect on specification of mesodermal and endothelial lineages at pre-hematopoietic fates (173). However, like haploinsufficient mice, some patient-derived iPSCs with GATA2 variants (e.g., p.R361H) were able to retain sufficient residual GATA2 activity for normal hematopoietic development (173) as is seen in asymptomatic carriers.

GATA1 plays an indispensable role for the downstream differentiation of various hematopoietic lineages, including myeloid stem cells, megakaryocytes, erythrocytes, mast cells, platelets, and dendritic cells. (Figure 2, Table 1) GATA1s (GATA1 "short" - lacks N-terminal domain) has been extensively used as a biological tool to investigate the effect of GATA1 aberration on hematopoiesis, where LOF GATA1 initiates the arrest of both primitive and definitive erythropoiesis by inducing apoptosis in erythroid progenitors and phenotypically displays BMF syndromes (e.g., disordered terminal platelet maturation) without manifesting into leukemia (162, 174). Interestingly, mimicking the acquisition of de novo GATA1 mutations in primary human HSPCs, the same loss of GATA1 has the opposite effect on megakaryocytes, causing fetal (not neonatal or adult) megakaryocytic progenitor hyperproliferation, where megakaryocytes fail to undergo differentiation but expand dramatically, by potentially hijacking the myeloid mechanism to promote this proliferation (174–176). This abnormality has also been linked with a myelofibrotic phenotype in mice, and the (GATA1low) mice has been extensively used a biological tool for a targeted therapeutic pathway, using and Ruxolitinib and a

monoclonal antibody (RB40.34) to improve effective hematopoiesis in BM by restoring spleen architecture and reducing fibrosis in the BM (177).

The role of GATA1 in leukemogenesis (differentiation block, proliferation) has been shown primarily in trisomy 21 models where spatiotemporal intensity of GATA1s expression contributes to pathological phenotypes. For example, point mutations in trisomy 21 iPSCs show GATA1s impairs megakaryocytic differentiation, by inducing the emergence of megakaryoblasts with major platelet and a-granule formation impairments (178). Conversely, acquired trisomy 21 has been reported in a patient carrying a germline GATA1 variant causing increased GATA1s, and associating with chronic anemia and thrombocytopenia as well as transient abnormal myelopoiesis, although the latter is likely a result of the trisomy 21 (179).

Mediated through the two ZF domains, the duality of GATA1 as both an activator and repressor results in the formation of several distinct complexes involved in hematopoiesis (180), with a variety of partners including LMO2 (181), RUNX1 (182), FOG1 (183), TAL1, GFI1B, ZFP143, ETO2 (184), LDB1 (182), FLI1, EKLF, HDAC5 and PU.1 (185). Study of these interactions may provide insight into GATA1-mediated predisposition. For example, Runx1 expression in zebrafish cooperatively promotes primitive hematopoiesis with GATA1 (186), and methylation of the Gata1 locus (by recruitment of DNMT1 and the GATA1 methylation determining region) prevents GATA2-mediated GATA1 activation in HSPCs (187).

Recent studies have suggested that environmental factors may play a role in the development of GATA1 mutations, as epigenome-wide association studies have linked a metastable epiallele (VTRNA2-1) to a differentially methylated GATA1 region associated with Down Syndrome (188). Additionally, the ZF domain of GATA1 is structurally favorable for interactions with other proteins, such as ASIII, which has been shown to suppress GATA1 (and GATA2) function (189). This suppression may mediate differentiation blocks that are associated with leukemia.

MECOM mirrors the regulatory effect of GATA2 on normal hematopoiesis, demonstrating a vital role in the differentiation of HSCs directly through upstream regulation of GATA2. Inherently, Evi1(-/-) mice also show marked reduction of HSCs through simultaneous reduction of GATA2 expression seen in these models (190), and predictably, reconstituting EVI1 expression upregulates GATA2 expression and rescues HSC differentiation ability (191, 192). Heterozygous CRISPR edits of MECOM engrafted bone marrow into mice also reduced HSC production ~2 fold, but found no detectable differences in lymphoid, erythroid, megakaryocytic or monocytic lineages (192). The ZF domains enable association in DNA repair, chromatin remodeling and transcription, including with downstream TF targets (i.e., SPI1, GATA1/2, RUNX1 and CEBPA) driving the MECOM regulatory network (193, 194). Interestingly, amongst RUNX1, GATA2 and JUN, CHIP-seq data identified the most enriched motif in HSPCs is the ETS motif, which can be bound by ERG, ETV6, ETV2 and FLI1, despite highly enriched TF occupancy by FLI1, RUNX1 and GATA2 (192). Single-cell transcriptomics exemplifies the consequence of MECOM disruption (CRISPR-mediated) by

dysregulating over 700 genes including key factors expressed during hematopoiesis (192).

The function of EVI1 is critically modulated by posttranslational modifications including phosphorylation, sumoylation, ubiquitylation, and monomethylation (195). Constitutive expression of EVI1 through bone marrow transplantation in mice recapitulates phenotypes resembling MDS. (Table 5) Although none manifest into leukemia, EVI1 induces delays in HSC and erythroid differentiation which in later stages (10-12 months) lead to hematopoietic failure and death (196). The upregulation of MECOM and consequent down regulation of network targets is associated with poor prognosis in AML (192). Hence, EVI1 dysregulation may be a contributor in AML risk stratification (197) and may be a target for therapeutic intervention. For example, knockdown of EVI1 in K562 cells shows a greater sensitivity to therapeutic drugs (e.g., Imatinib) and amongst other targets (PTGS1, COX-1/2), prevents expression of a gene that affects platelet regulation (ITGA2B) (198).

A commonality in their multiple ZFs as well as their belonging to the same TF family, may imply that the mechanisms underlying IKZF-1, -2, and -3 contributions to BMF and/or HM are comparable. However, IKZF TFs typically exhibit functional redundancy, meaning that the absence of one is often compensated by the presence of another (199). This can be well articulated by the phenotypic differences in IKZF1 mice. Homozygous Ikzf1 mice containing a missense mutation in ZF3 (p.H191R) result in embryonic lethality, whilst the same heterozygous mutation has normal numbers of B-cells, but reduced numbers of B-cell precursors in the bone marrow (103). Similarly, the lack of T and B cell progenitors as well as NK cells was a result of a deletion of the first 3 ZFs. Mice with a deletion of 3 N-terminal ZFs, however, have a high incidence of T-cell leukemia. Despite these mutations seemingly resulting in the dysregulation of IKZF1, mutations retaining the ability to form heterodimers with other IKZF family proteins reflect a less severe phenotype (199). Interestingly, a mouse model harboring an IKZF3 variant (p.G158R) was able to affect the binding of both IKZF1 and IKZF3 (199). Although homozygous IKZF3 and IKZF5 knockouts have been shown to be viable, (Table 1) complete loss of IKZF1 results in embryonic lethality, associated with defects in erythroid cells and an expansion of megakaryocyte progenitors, highlighting the importance of this TF in hematopoiesis over its familial counterparts (161).

ETS BMF/HM predisposition TFs

Oncogenic transformation relies on dysregulated use of normal developmental pathways, and ETS TFs highly co-operative roles in normal hematopoiesis suggests that the impact of one ETS TF aberration may change the regulatory role of others. Interestingly, through the protein's pointed domain, ETV6 has direct interaction with another known ETS TF BMF predisposition gene, FLI1 (200). This interaction has now been shown to alter tumor growth in Ewing's Sarcoma patients, where the EWS-FLI1, an oncogenic fusion protein that drives tumor growth in 90% of patients, is constrained by a loss of ETV6 (201).

The central role of ETV6 as a 'transcriptional suppressor' makes its dysregulation unsurprising in HM. For example, co-binding of ETV6 and a protein complex mediating chromatid cohesion has been shown to lead to transcriptional repression of erythropoiesis-related genes, suggesting involvement in inhibition of erythroid differentiation in myeloid malignancies (202). Etv6(-/-) mice result in embryonic lethality with yolk sac angiogenic defects and perturbation of definitive hematopoiesis in bone marrow. (Table 2) (203) Interestingly, a mouse model of the most common ETV6 pathogenic variant outside of the ETS-domain, p.P214L (in mice Etv6P216L/wt), did not show homozygous lethality and the mice had no overt hematopoietic phenotype. In conditional knockout mice, thrombocytopenia is observed as the primary phenotype due to an increase in megakaryocytic colony forming cells implying a loss of ETV6 results in a terminal defect in megakaryocyte maturation (204). (Table 5) Modeling ETV6-associated leukemia commonly uses the ETV6-RUNX1 fusion protein. In mouse models, a low incidence of leukemia is seen, developing only after a long latency (205) and/or secondary genetic hits (206), and often results in inactive HSCs and loss of lymphoid progenitors (207).

Competitive transplant experiments revealed Etv6P216L/wt HSPCs had reduced lymphoid reconstitution potential. Specifically, Etv6P216L/wt mice showed impaired MPP4 hematopoietic progenitor populations with lymphoid potential (208). Transcriptome analysis of the MMP4 population from these mice suggested deregulation of inflammatory pathways as a consequence of the Etv6 variant, however no significant difference in cytokine or chemokines were identified in the BM. These mice also did not develop leukemia, suggesting additional models or stressors are required to achieve the full range of human disease phenotypes.

In patient-derived peripheral blood mononuclear cells, the lowering of ETV6 results in an increase in histone acetylation suggesting epigenetic changes are likely to contribute to hematopoietic dysregulation. Further, single cell RNA-seq of ETV6 variants p.P214L and p.R369Q revealed enrichment of interferon response genes across peripheral blood cell populations, and HDAC3 as an upstream master regulator (209) consistent with the ETV6/HDAC3 complex regulating the interferon response. Dysregulation of the complex in megakaryocytes impaired proplatelet formation. ETV6 variants (e.g., R369Q) in heterozygosity have also been shown to drive thrombocytopenia in iPSCs, with more, but less responsive megakaryocytes that are deficient in platelet formation, leading to fewer platelets (94).

FLI1 is the most highly expressed BMF and/or HM TF predisposition gene in both myeloid and lymphoid lineages during most stages of hematopoiesis (Figure 2), and directly regulates expression of various target genes within its broad target gene program, which may be the mechanism behind its BMF association. For example, FLI1s large transcriptional target program allows the protein to alter glycolysis, shifting the glycolytic balance to be more aerobic to enable robust cell division, and through the repression of the PKLR (pyruvate kinase) promoter, can initiate a block in erythroid differentiation (210). Despite its colocalization with GATA1/2 and RUNX1 (211),

TABLE 5 Hematological-related mouse phenotypes found in BMF/HM predisposition genes including strong candidate genes and other candidate genes.

	TF gene symbol	MGI Mouse Phenotype (Hematological)
		Phenotype
Known BMF/HM TFs	Runx1	- **abnormal** hematopoiesis-related cell morphology/development/number, hematopoietic differentiation, and spleen/thymus size/morphology - Thrombocytopenia
	Cebpa	- enlarged spleen - **abnormal** definitive hematopoiesis, myelopoiesis, granulocyte, leukocyte common myeloid progenitor and bone marrow cell morphology and myeloid leukocyte - increased erythroid progenitor, hematopoietic, neutrophil, granulocyte, monocyte cell number and decreased hematocrit
	Gata2	- **abnormal** embryonic/definitive hematopoiesis, blood cell/megakaryocyte morphology/development and stress erythropoiesis/hematopoiesis - decreased common myeloid/erythroid progenitor cell number - anemia
	Pax5	- **abnormal** lymphopoiesis, class switch recombination, B/T cell number/differentiation/physiology, immunoglobulin heavy chain V(D)J recombination and macrophage, monocyte and leukocyte cell number - decreased immunoglobulin and IgA, IgD, IgG1, IgG2a, IgG2b, IgG3, IgG, IgM levels
	Mecom/ Evi1	- **abnormal** neutrophil differentiation, common myeloid progenitor cell morphology/number, bone marrow cell morphology/development, platelet, HSC and fetal liver HPC morphology and leukocyte, HSC, precursor and B-cell numbers - thrombocytopenia
	Stat3	- **abnormal** thymocyte, T cell apoptosis/proliferation, thymocyte, neutrophil, dendritic cell, T- granulocyte, macrophage, monocyte and leukocyte cell number, T cell, macrophage differentiation/physiology and IgE, IgG1 and IgG3 levels
	Gata1	- **abnormal** embryonic erythropoiesis, definitive hematopoiesis and megakaryocyte differentiation, spleen, common myeloid progenitor cell, megakaryocyte, proerythroblast, erythrocyte, platelet, mononuclear, spleen, bone marrow cell and erythroid progenitor morphology, erythropoiesis, myelopoiesis, thrombopoiesis and extramedullary hematopoiesis, bone marrow, megakaryocyte, erythroid progenitor, eosinophil, erythrocyte, neutrophil, leukocyte, lymphocyte, osteoclast and splenocyte cell numbers, and hemocrit - anemia, anisocytosis, poikilocytosis, polychromatophilia and thrombocytopenia
	Etv6 (Tel)	- **abnormal** megakaryocyte, B cell differentiation, definitive hematopoiesis, erythropoiesis and lymphopoiesis, spleen, common myeloid progenitor, lymphocyte progenitor cell and myeloblast morphology, erythroid progenitor, erythrocyte, neutrophil cell, B cell, lymphocyte and HSC cell numbers and bone marrow cell physiology - myeloid hyperplasia and thrombocytopenia
	Fli1	- **abnormal** megakaryocyte, B cell differentiation, erythropoiesis, thrombopoiesis and hematopoiesis, thymocyte, B cell, megakaryocyte, erythroid progenitor, erythrocyte, granulocyte, NK and monocyte cell numbers, spleen and megakaryocyte, proerythroblast and leukocyte morphology - decreased B cell proliferation, hematocrit and IgG1 and IgG3 levels - anemia and thrombocytopenia
	Ikzf1	- **abnormal** B cell apoptosis, granulocyte, macrophage and B/T cell differentiation, B/T cell proliferation, thymus and spleen, definitive hematopoiesis, erythropoiesis, lymphopoiesis and myelopoiesis, common myeloid progenitor cell morphology, B/T, bone marrow, erythroid progenitor, basophil, erythrocyte, granulocyte, dendritic, NK cell, macrophage and splenocyte cell numbers, erythroid progenitor cell, eosinophil, erythroblast and NK morphology, and natural killer cell mediated cytotoxicity. - myeloid hyperplasia, extramedullary hematopoiesis and anemia - decreased hematocrit, hemoglobin content and IgG3 and IgG levels
	Ikzf3	- **abnormal** B/T cell proliferation and differentiation, spleen, neutrophil, B- cell, lymphocyte and leukocyte cell numbers and IgE, IgG1, IgG2a, IgG2b, IgG3 and IgM levels - thrombocytopenia
	Ikzf5	- **Abnormal** IgA and IgE levels - Epistaxisl, Joint hemorrhages, Persistent bleeding after trauma, Bruising susceptibility, Petechiae, Anemia of inadequate production, Poikilocytosis, Acanthocytosis, Leukocytosis and Thrombocytopenia - Increased megakaryocyte colony forming unit counts
Strong candidate BMF/HM TFs	Gata3	- increased T cell apoptosis and decreased IgE levels - **abnormal** thymus, thymocyte, T- and eosinophil cell numbers, definitive hematopoiesis, and T cell differentiation/morphology
	Spi1 (Pu.1)	- **abnormal** granulocyte, neutrophil, macrophage, and B/T cell differentiation, thymus and spleen, blood cell, bone marrow common myeloid progenitor, T-cell, proerythroblast, granulocyte, microglial, monocyte and spleen morphology, definitive hematopoiesis, erythropoiesis, leukopoiesis and myelopoiesis, B/T-, bone marrow, erythrocyte, granulocyte, neutrophil, Lymphocyte, macrophage, HSC cell numbers and HSC physiology. - increased hematopoietic stem cell proliferation, mean corpuscular volume and IgM level, and decreased thymocyte number and hematocrit - absent common myeloid progenitor cells and anemia and extramedullary hematopoiesis
	Erg	- **abnormal** embryonic erythropoiesis, spleen morphology and leukocyte, HSC and B cell numbers (decreased) - thrombocytopenia and pancytopenia

(Continued)

TABLE 5 Continued

	TF gene symbol	MGI Mouse Phenotype (Hematological)
		Phenotype
Other candidate BMF/HM TFs	**Etv2**	- **abnormal** spleen, embryonic erythropoiesis, definitive hematopoiesis, myelopoiesis, bone marrow and myeloblast morphology/development, common myeloid progenitor, bone marrow, erythroid progenitor, leukocyte and HSC cell number and absent erythrocytes - anemia
	Elf1	- **abnormal** B cell differentiation and B cell physiology
	Fev	No Hematological phenotype reported
	Ets1	- **abnormal** B cell, thymocyte, splenocyte apoptosis, B/T cell proliferation, thymus and spleen, lymphopoiesis, B cell differentiation, B/T-, plasma, NK, leukocyte cell numbers, leukocyte, NK cell physiology and T cell activation - impaired natural killer cell mediated cytotoxicity
	Ets2	No Hematological phenotype reported
	Ikzf2	- **abnormal** regulatory T cell physiology
	Ikzf4	- decreased T cell proliferation and increased immunoglobulin level - **abnormal** spleen and B/T cell number
	Gfi1b	- **abnormal** megakaryocyte differentiation and erythropoiesis, blood, bone marrow cell morphology/development, erythroid progenitor, erythrocyte, neutrophil, lymphocyte, HSC, splenocyte cell numbers, megakaryocyte progenitor and embryonic erythrocyte cell morphology and hemoglobin content (decreased) - anemia and thrombocytopenia
	Tal1 (SCL)	- **abnormal** embryonic hematopoiesis and erythropoiesis, thymus and spleen, common myeloid progenitor, bone marrow, leukocyte, and HSC cell numbers and bone marrow, megakaryocyte, erythroid progenitor and proerythroblast cell morphology - anemia, anisocytosis and thrombocytopenia - decreased hematocrit and hemoglobin content, mean corpuscular hemoglobin concentration and mean corpuscular volume
	Lyl1	- **abnormal** B/T cell number/differentiation and hematopoietic stem cell physiology - thrombocytopenia
	Lmo2	- **abnormal** embryonic hematopoiesis, thymus and spleen, common myeloid progenitor and megakaryocyte progenitor cell morphology, erythrocyte, granulocyte and monocyte cell number - decreased hemoglobin content, increased mean corpuscular volume - thrombocytopenia
	Meis1	- decreased common myeloid progenitor, megakaryocyte, platelet cell number - **abnormal** bone marrow and megakaryocyte progenitor cell morphology - increased hematopoietic stem cell number
	Tcf12	- thymus hypoplasia - **abnormal** thymocyte B/T cell number and morphology - arrested T cell differentiation
	Gfi1	- **abnormal** granulocyte, monocyte, NK and T-cell differentiation, T cell proliferation, thymus, thymocyte, T/B-cell neutrophil, granulocyte, monocyte and HSC number, definitive hematopoiesis, lymphopoiesis and myelopoiesis, neutrophil and monocyte morphology and NK cell and T-helper physiology
	Fos	- increased macrophage apoptosis - **abnormal** osteoclast differentiation, thymus and spleen, thymocyte, T/B-cell, granulocyte, splenocyte and osteoclast cell number, lymphopoiesis and macrophage physiology
	Hoxa9	- **abnormal** common myeloid progenitor cell morphology, bone marrow and HSC cell number and spleen
	Fosb	No Hematological phenotype reported
	Foxc2	No Hematological phenotype reported
	Eklf (Klf1)	- **abnormal** spleen, erythropoiesis, hematopoiesis and erythroblastosis, erythrocyte morphology and cell number - decreased hematocrit, hemoglobin content, mean corpuscular hemoglobin - anemia, anisocytosis, poikilocytosis and polychromatophilia
	Nfkb1	- **abnormal** spleen, B/T cell proliferation, bone marrow cell morphology/development, erythrocyte, B/T-, NK, myeloid dendritic, osteoclast and leukocyte cell number, lymphocyte, B/T cell and macrophage physiology, IgA, IgE, IgG1, IgG3, IgG, IgM and immunoglobulin levels and CD4-positive, alpha-beta T cell physiology - myeloid hyperplasia and decreased mean corpuscular hemoglobin concentration

Mouse phenotypes collated through Mouse MGI (25–27).
NA, not applicable.

FLI1 has also been shown to cooperate with RUNX1 by co-binding to ETS-RUNX motifs, to restrain transcription factors that aid T-cell differentiation (212). Importantly, as a key regulator of T-cell differentiation programs, FLI1 deficiency does not diminish the T-cell population, but aids in safeguarding its transcriptional and epigenetic commitment (212).

Fli1 mouse models utilizing the deletion of its carboxy-terminal regulatory (CTA) domain showed impairment of megakaryocytic development and platelet number and function (213). Work showing a synergistic relation between FLI1 and GATA1 to regulate megakaryocytic genes (214), was later confirmed by the failure of aberrant FLI1 to recruit GATA1 to several megakaryocytic promoters (213). Similarly, in human-derived iPSCs, megakaryocyte and platelet defects were shown to be predominantly a result of FLI1 deficiency (215).

Gene annotation and curation standards for TFs

As described above, core hematopoietic TFs including RUNX1, GATA2, GATA1 and MECOM play crucial roles in regulating multiple aspects of hematopoiesis and HSPC biology. Due to their central involvement, it is not surprising that genetic variation in these TFs can lead to differentiation blocks and perturbations in the regulatory networks that govern normal hematopoiesis, and thus lead to BMF and HM. While these genes are well known predisposition genes, it remains challenging to recognize other germline predisposition genes or variants in known genes especially if penetrance is low, phenotype expressivity is variable, frequency of pathogenic gene variants is very rare, or the variant exists in promoter/enhancer regions that are not characterized or linked to the causal gene.

Recurrence within families with a history of HM, BMF or a bleeding disorder, as well as identifying individuals that develop HM at a young age are strong predictors for moderate-highly penetrant germline predisposition, although there are more and more studies demonstrating that the lack of family history of a particular BMF or HM is not a great indicator that an individual is not a carrier of a germline predisposition variant (216, 217). Advancements in technology, accessibility and reduced costs associated with Next-Generation Sequencing (NGS) technologies has allowed for more routine investigation of a germline genetic cause in individuals and/or families. With increasing numbers of individuals being screened *via* NGS, undoubtedly this results in identification of novel variants and genes associated with these disorders. The American College of Medical Genetics and Genomics (ACMG) has been instrumental in developing guidelines for the interpretation of genetic variants, although these guidelines are developed for genes with a definitive role in the pathogenicity of the disease (218). Thus this can make it challenging to identify novel disease genes. Classification of variants according to ACMG criteria is based on the strength of available evidence including population frequency, computational data, functional studies, segregation data and relies on knowledge of disease mechanisms and clinical information (218). The criteria

used to classify genetic variants means it is challenging to determine novel disease genes as it often relies on multiple probands with the same phenotype, associated segregation data and functional studies to validate the gene as disease-causing. In the context of predisposition to HM, additional contributing factors are not considered including ontology, environmental factors, and somatic data. Somatic mutations are a common molecular mechanism behind oncogenicity, and so the inclusion of this data to pathogenic classification should not be overlooked. Correct classification of genetic variants is crucial in disease management and treatment.

Estimating the number of genes missed by confining variant classification to a guidelines-based approach is complex, yet their existence can be speculated. From a discovery perspective such guidelines can be extrapolated to dealing not only with germline "variant of uncertain significance" (VUS), but also with "gene of uncertain significance" (GUS, for a gene not previously known to cause a certain phenotype) and "phenotype of uncertain significance" (PUS, for a phenotype not previously associated with a gene; that is, phenotype expansion for a gene). Ultimately, a collective weight of evidence is required to transition each of these out of the "uncertain" classification.

Incorporation of somatic data in variant curation can assist in determining the pathogenicity of a novel variant; for example, somatic DDX41 mutations are often detected in individuals with myeloid malignancy with a germline DDX41 variant (80% of cases) but vary rarely seen in the absence of a pathogenic germline DDX41 variant (219, 220). Progression of disease resulting from germline predisposition of penetrant genes (GATA2, RUNX1 and CEBPA) is often accelerated by the acquisition of unique sets of somatic gene aberrations affecting disease progression, clonal architecture, and treatment response (221). Somatic RUNX1 alterations (translocations and mutations) are frequently associated with MDS and AML and are considered responsible for leukemic progression in transformation from BMF into leukemia (222–224). Somatic variant data highlights the threshold effect associated with the RUNX1 activity, since reduced or absent activity associates with severity and prognosis of BMF and/or HM. The presence of somatic RUNX1 mutations in HM development, as well as the identification of RUNX1 as a germline cause of HM, highlights the possibility that other genes recurrently observed in sporadic HM may also contribute to inherited BMF/HM. Incorporating such data into existing guidelines for future variant stratification may be beneficial.

Epidemiologic studies have established that the incidence of leukemia may be influenced by a combination of germline abnormalities and environmental risk factors including lifestyle choices (i.e., smoking and obesity) (225), environmental exposures (i.e., cytotoxic agents and electrical power) (225, 226), medical history (i.e., regular aspirin use) (227), and other chemical and biological agents (i.e., radiation, retroviruses) (228). For example, the broad range of GATA2 deficiency syndrome phenotypes and penetrance implicate the involvement of environmental stressors that increase the likelihood of developing particular phenotypes (229). The role of environmental factors in clonal hematopoiesis emergence (e.g., metabolic syndromes causing chronic inflammation and chemotoxic exposure) may also be highly

specific to the mutation that marks each stem cell clone (230). Despite these findings, and efforts to standardize variant interpretation and calling, environmental factors are currently neglected in TF/gene stratification. As knowledge of confounding environmental stressors become better known, a variant's classification may differ for different individuals within a family, for instance, dependent on environmental influences such as is already done for responses to certain medications (231, 232).

Predictive inclusions to BMF and/or HM predisposition TFs

Many TF networks have inherent redundancies that minimize the impact of disruption of any one TF to target genes' expression levels. As such, phenotypes arise due to critical gene expression changes where a particular TF has a major role that drops expression or transcriptional activity below a critical threshold leading to disease. Importantly, this threshold may only be reached in some cases due to internal or external stresses and may be the reason why we see the somatic involvement of some gene variants in disease (for example sporadic HM), but not as pathogenic germline variants. Additionally, certain genetic variants that occur somatically in the bone marrow or blood may never be seen as germline because they are embryonic lethal. A framework utilizing in-silico pathogenic predictor scores, occurrence in oncogenic gene fusions, and previous disease association could provide insight into the oncogenic potential of novel TFs as BMF and/or HM predisposition genes. By highlighting them, others may include in their variant analyses, scrutiny, and curation approaches.

In this section, we will cover criteria that may aid in determining whether novel TFs may be disease causing. Highlighting this criterion, we will use TFs from our list of potential novel candidates (Table 4). We focus on two ETS TFs, which are part of the known transcriptional cluster model for hematopoiesis, ERG and PU.1 (36), and a ZF TF, GATA3. These three candidates all regulate central biological processes in hematopoiesis and have known oncogenic potential, which make them prime candidates as novel BMF and/or HM predisposition genes.

Role in normal hematopoiesis

ERG, PU.1 and GATA3 are vital TFs in normal hematopoiesis, involved in the regulation of broad transcriptional programs in lineage differentiation of various types of mature blood cells. ERG has a unique and multifaceted role within the normal HSC compartment, as it is particularly critical for B lymphopoiesis, progenitor self-renewal and HSC function to sustain definitive hematopoiesis, often by preventing HSC exhaustion (83, 233–237). Despite this, hematopoietic specification and the initiation of definitive hematopoiesis is not ERG-dependent (236). PU.1 is another 'pioneer' TF, heavily reliant on its large transcriptional program and alternative mechanisms (SWI/SNF, and c-Jun) (234, 235) to assist hematopoiesis, particularly myeloid and B cell differentiation (238, 239). GATA family of TFs (GATA1-3) are vital determinants of multilineage hematopoiesis, and so far,

germline variants in GATA1 and GATA2 have an established role in HM development, particularly GATA2 as one of the most characterized predisposition genes. Given the crucial role of the GATA TF family in both normal and leukemic hematopoiesis, it is imperative that much like its familial counterparts, germline mutations in GATA3 will likely also have pathogenic consequences. GATA3 is essential for the development, maintenance, survival and proliferation of early T-cell progenitors and HSC emergence, despite its expressional absence in most hematopoietic cells. (Figure 2).

Disease phenotypes and malignancy associations

The central involvement of ERG, PU.1 and GATA3 in hematopoietic lineage differentiation and their embryonic lethality occurring concomitantly with the onset of definitive hematopoiesis makes them primal candidates for hematopoietic predisposition genes (240–242). Both ETS TFs, ERG and PU.1, have been linked to hematological-related phenotypes in humans, and all three TFs have exhibited hematologic-related phenotypes in mice, further strengthening their candidacy (Table 4). Somatic LOF mutations in all three TFs have been observed in cancer with GATA3 exhibiting the highest frequency of LOF mutations (over 700).

ERG is constitutively expressed and exerts multiple hematopoietic/non-hematopoietic homeostatic functions, though its pathological dysregulation outside of this homeostatic range has previously associated the protein with HM amongst other diseases. ERG's overexpression has been identified as a biomarker correlated with adverse AML clinical outcomes, and its deregulation (i.e., intragenic deletions) found in ALL patients, has been the result of favorable outcomes (70, 96–100, 243, 244). This suggests that like GATA2 germline mutations, ERG may not uniformly disrupt expression and function in all cellular contexts. Traditional biological models (i.e., knock-out, knock-in and knockdown mouse models) have also implicated ERG in BMF and/or HM phenotypes. The characterization of a (mouse) germline ERG variant (p.S305P) residing in BMF (thrombocytopenia) also alludes to the possibility of ERG as a low penetrant leukemic predisposition gene (83, 233, 245). In other diseases, ERG-dependent transcription has been shown to modulate cardiovascular disease (246), and germline LOF ERG variants have been reported to lead to lymphedema (247), see also in ~10-15% of germline GATA2 cases (74). Perhaps the strongest example of oncogenic potential is ERG's involvement in chromosomal translocations. Gene fusions between androgen-regulated genes and ERG (e.g., TMPRSS2-ERG) occur in ~50% of prostate cancers, with the alteration resulting in the presence of ERG overexpression in both early and late-stage prostate cancer (248, 249). This overexpression drives oncogenic effects such as increased cell growth, increased expression of neurotransmitter receptors and promotes tumor development, making the fusion a standard biomarker for diagnosis and stratification (249, 250). Hematopoietic-associated fusions involving ERG also extend to EWSR1-ERG in 5-10% of Ewing's sarcoma, and ELF4-ERG and FUS-ERG in acute myeloid leukemia (96, 251, 252).

Despite its heavy-weighted importance in blood formation, large HM-related phenotypic link in mice, and embryonic lethality seen in homozygous mice, (Table 4) pathogenic germline SPI1 (encodes PU.1) variants are also yet to be found. The dysregulation of PU.1 and its consequent role in HM-related pathology, however, has been well studied. Specifically, PU.1 protein levels have been linked to the inhibition of cell division, cell cycle and leukemogenesis. PU.1 overexpression leads to differentiation blocks and thus acute erythroleukemia (101), but the reduction of PU.1 expression (and consequently gene network) has been shown to aid in leukemic transformation by many mechanisms including TET2 deficiency (253, 254), differentiation blocks and cellular expansion involving synergistic combination of PRC2 and HDAC1 (255), and at the post-transcriptional level, sustained expression of miR-155 (256). PU.1's well established role in leukemogenesis makes it an attractive target for therapeutic intervention; for example, the inhibition of miR-155 to inhibit cell growth, controlled by dysregulated PU.1, has been proposed as a potential pathway to impact outcomes in HM (257).

Although PU.1 has a well-established role in leukemogenesis, the prevalence of somatic mutations remains relatively low. Only 0.32% of somatic point mutations in PU.1 are reported in hematopoietic and lymphoid tissues in the COSMIC database, (Table 2) and one recurrent somatic PU.1 mutation shown to be associated with a poor prognosis of Waldenstrom macroglobulinemia, underscoring the rarity of PU.1 somatic mutations. The variant (p.Q226E) has been shown to modify DNA binding specificity and transactivation capacity on ETS-like binding sites (258). *De novo* aberrations in PU.1 have also been linked to autosomal dominant "Agammaglobulinemia, including lymphopenia, neutropenia and impaired B-cell development" phenotypes. Patient cells show loss of PU.1 protein expression, which is consistent with a LOF and haploinsufficiency mechanism (259). These observations suggest a possibility of reduced penetrance and variable expressivity for HM-related phenotypes.

GATA3's pivotal role in developmental T lymphopoiesis, particularly in late stages of T-cell differentiation (242), may assist in predicting mutational burden, although its role in the progression of lymphoid-related leukemias is largely unexplored. Dysregulated GATA3 expression has previously been implicated in a subgroup of T-ALL patients, where both high and low GATA3 expression resulted in changes of target gene expression clusters (260). This dysregulation was initially reported to have no significant changes in clinical outcomes (261); however, a subsequent study revealed that GATA3 dysregulation was associated with poorer survival and adverse prognostic implications (262). In biological models, GATA3 aberration has been shown to disrupt signaling cascades, one of which, a GATA3-dependent cytokine (i.e., IL-13), has been shown to promote the growth and survival of malignant T-cells (260). A GATA3 mutation in zebrafish embryos (p.R276Q) was also hypothesized to collectively affect T-cell proliferation and differentiation, eventually contributing to the pathogenesis of T-ALL, highlighting an emerging role for GATA3 in HM (263).

Dysregulation of GATA3 has also been associated with other diseases, including breast cancer, where reduced GATA3 expression was linked with a poorer prognosis and unfavorable tumor phenotype (264), and where mutations in exon 6 of GATA3 were detected in >50% of tested patients. Interestingly, the study suggested a difference in GATA3 mutation type may have different outcomes, where intronic germline mutations correlate to better prognosis whilst protein coding variants do not (265).

Pathogenicity predictors

Scores like pLI predict the tolerance of a gene due to the sum of variants causing premature termination of a protein and are often used to prioritize candidate genes when analyzing genomic data (266). Both ERG and PU.1 are highly intolerant to LOF (0.96, 0.98, respectively, Table 2, Figure 3) as are their DNA binding ETS domains. This contrasts with other ETS TFs, such as FEV, which has a much lower pLI score (0.01, Table 3) despite having a similarly intolerant DNA binding ETS domain. Despite the rarity of somatic mutations and absence of germline mutations in BMF/HM in ERG and SPI1, nine pathogenic ClinVar variants (in FLI, ETV6 and ERF) all located within the ETS DNA binding domain, are directly aligned to corresponding amino acids in both PU.1 and ERG, which is highly suspicious of a similar disruptive effect on the associated protein's function.

GATA3 also exhibits a high pLI pathogenicity predictor score of 0.9 (Table 2, Figure 3), and like its familial counterparts, has highly intolerant (ZF1) and intolerant (ZF2) DNA binding domains. Given the number of germline ClinVar reported pathogenic variants segregating in the highly conserved ZFs of both GATA1 and GATA2, it is unsurprising that over 20 variants are directly aligned to the same amino acid in GATA3. Interestingly, GATA3 has a reported gene-disease association to both deafness and renal dysplasia, phenotypes of which have also been reported in GATA2 deficiency syndrome (53, 267).

Gene expression and interactions with known predisposition genes

Based on current reported BM and/or HM predisposition TF genes and correlation to their expression levels in various hematopoietic cell types, mRNA expression levels while informative are poorly predictive of whether variants in an expressed gene are likely to be pathogenic (Figure 2). This may be because of the difficulty in determining threshold levels of activity for each TF for normal function, and because mRNA levels don't always correlate with protein or activity levels. It is also unsurprising because single cell studies show that despite the very low or absence of GATA2 expression in myeloid progenitor cells, the preferential loss of myeloid cells (e.g., monocytes) is a feature of GATA2 deficiency syndrome (268). The impact of both temporal and spatial expression of TFs is also important in the biology leading to genetic variants being pathogenic (i.e., detrimental in a way that promotes oncogenicity either under normal physiological conditions or stressed environments), along

with interplay with cofactors and their expression profiles during hematopoietic processes.

The interconnectedness of the known TF BMF and/or HM predisposition gene network is demonstrated by reported protein-protein interactions (Supplementary 1A). Given phenotypic overlap and the biological pathways involved, it is likely novel genes would show a similar network of protein-protein interactions with this group of TFs. Using STRING analysis, it is clear that the majority of genes in our potential TF predisposition list show a strong interconnectedness with our known TF network list (Supplementary 1) (12). As an example, GATA3 shows strong links with several of the well-known predisposition TFs including RUNX1 and GATA2. Utilizing the network approach, FEV is an outlier of the network suggesting it is less likely to be an important TF in BMF/HM predisposition without known interactions with other TFs important to disease pathogenesis. The low pLI score and negligible number of somatic variants in COSMIC do not support a role for pathogenicity alongside the temporal expression of FEV during prenatal hematopoiesis (269).

ChIP Enrichment analysis (ChEA) has been utilized to analyze over 70 gene set libraries to identify consensus target genes for selected TFs (13). In cases where data for a particular TF was available from multiple experiments, a set intersection approach was employed to obtain a consensus on the top 9 interactions based on their p-value. (Table 1) Known BMF/HM TFs were among the top three interacting partners of 7 of our selected TFs (known and strong candidate BMF/HM predisposition genes), indicating that the network of genes associated with BMF/HM predisposition is highly interconnected. This also emphasizes the role of candidate inclusion TFs (GATA3, PU.1, ERG) in this gene network. Interestingly, ERG's direct regulation of RUNX1 and GATA2, and its (dose-dependent) absence, has been shown to reduce the expression of RUNX1 and GATA2, suggesting a possible mechanism of oncogenicity for ERG LOF/partial LOF variants (236).

Clonal hematopoiesis of indeterminate potential

CHIP is increasingly being investigated in hematological predisposition syndromes as a contributing factor in leukemogenesis. In the healthy population CHIP is a risk factor for HM, cardiac, and pulmonary disorders, and mortality (270), however this occurs primarily in the aging population (10% in >70 years). It is becoming apparent that early-onset CHIP occurs in TF predisposition HM syndromes including GATA2 and RUNX1 (271). What is not yet determined is whether this phenomenon of early-onset CHIP is largely attributed to TF HM predisposition genes or all HM predisposition genes. A recent study showing CHIP associated with ANKRD26 germline HM might argue against it being specific to TF predisposition genes, however this study was underpowered, warranting further investigation of larger patient numbers (95). Interestingly, germline ANKRD26 HM phenocopies RUNX1 FPD-MM, with pathogenic variants in ANKRD26

occurring in the 5'UTR and disrupting RUNX1 and FLI1 TF binding sites providing a link to altered TF biology and CHIP. [9,266] It might be anticipated that those predisposition genes, such as TFs and/or epigenetic regulators, that have an impact on a multitude of pathways during hematopoiesis are more likely to impact pathways predisposing to CHIP mutation accumulation such as DNA damage pathways, which has been implicated in RUNX1 FPD-MM. Interestingly, the majority of CHIP mutations are in genes involved in epigenetic regulation and mutations in TFs are rarely observed. The reason for this could be two-fold, firstly TF mutations, may lead to deregulation of HSC senescence and proliferation resulting in a downstream impact on cellular survival, or lead to activation of terminal differentiation of HSC which would not result in the clonal advantage required for CHIP, alternatively, the effect on gene-regulation networks and associated epigenetic changes results in progression to malignancy. An example of this is in RUNX1 FPD-MM, in unaffected individuals with germline RUNX1 variants we do not detect CHIP in RUNX1 however biallelic RUNX1 variants are the most common event associated with malignancy (229).

CHIP is associated with distinct DNA methylation (DNAm) patterns. Similarly, DNAm patterns change with both aging and in AML. CHIP has been postulated to act as a "molecular clock", with the presence of CHIP associated with an increase in age as measured by DNAm-biomarkers of aging (272).Why germline variants in TFs such as RUNX1 and GATA2, might accelerate this molecular clock still remains to be understood. Aged HSCs, as well as those with forced increases in cellular divisions, were shown to lead to alterations in DNAm (273).CHIP is likely reflecting the role of these TFs in regulating cellular processes such as mitosis, cell divisions, senescence, DNA damage and inflammatory responses, which can all lead to changes in the epigenome (274). Interestingly, the most common CHIP genes include DNMT3A and TET2, both of which are regulators of DNAm (275). It has been shown that both DNMT3A- and TET2-CHIP associated CpG DNAm patterns are enriched for ERG binding sites, while DNMT3A-CHIP sites are enriched for RUNX1 and RUNX2, GATA TF subfamily binding sites and TET2-CHIP sites enriched for ETS transcription factors (57). This is not unexpected given these are all TFs involved in hematopoiesis and implicated in leukemogenesis, thus providing a link between CHIP, DNAm changes and downstream consequences of altered HSC self-renewal and risk of leukemia development possibly mediated by these TFs.

Identification of CHIP prior to malignancy development holds great potential for future development of preventative therapies in TF HM predisposition genes. This may be through small molecule inhibitors designed to target specific epigenetic regulators, such as TET inhibitors (21) or designing drugs to target CHIP regulated pathways. An elegant study in zebrafish, demonstrated clonal dominance of HSPCs associated with loss-of-function variants in common CHIP genes asxl1,dnmt8(DNMT3A ortholog), and tp53 which was also associated with a pro-inflammatory environment to which these clones are resistant (22). This provides an opportunity to target inflammatory signaling pathways to prevent clonal outgrowth of HSPCs, such as through targeting inflammatory

Gene ontology

Gene ontology (GO) resources have long been used to characterize and identify enrichment of biological processes, molecular functions, and cellular components relevant to particular gene sets and different disease states. Ontology terms and databases can provide a baseline to organize, standardize and prioritize TFs using different biological categories important to BMF and HM development. This is evidenced when searching for GO terms enriched in our BMF and/or HM predisposition gene list (Table 1). There is a significant enrichment of biological process pertaining to hematopoiesis including stem cell proliferation and differentiation, and in particular differentiation to myeloid cellular lineages all of which are important processes for the development of BMF and HM, and the bias towards myeloid malignancies often observed with these genes (Table 1, Figure 4). GO analysis of the list of potential BMF/HM TF predisposition genes, closely correlates with our known predisposition gene list including enrichment of hematopoietic differentiation and stem cell regulation (Table 3, Figure 4). Interestingly, significant enrichment of immune response pathways and cytokine signaling are also observed. These pathways are impacted by external stimuli and stressors as well as aging which could account for why these genes have not yet been discovered as

predisposing, as it is likely these factors could act as modifiers having an impact on penetrance and clinical presentation, complicating the identification of more penetrant causative genes.

The absence of germline variation in some essential hematological TFs, could imply that a defective gene product would lead to a severe disruption in blood cell differentiation and a condition incompatible with life. This is supported by the paucity of genetic variation in the normal population (Table 2) and reported cases of miscarriages in families with pathogenic variants in multiple predisposition genes including *RUNX1*, *GATA2* and *MECOM* (Table 2) (53, 129). Furthermore, due to the role of these TFs in other organ systems, with genetic alterations leading to severe disorders (like *GATA3* mutations leading to HDR syndrome) (276), it reduces the possibility of being discovery in the context of HM predisposition. Studies focusing on stillbirth and fetal lethal conditions, detected a single case of a disease-causing variant in a novel hematopoietic TF gene (*GATA3*), terminated at gestational week 20 (Table 2) (277, 278). The limited number of associations with fetal death could be attributed to timing of genetic testing. First blood cells differentiate from extraembryonic mesoderm around the 7th day of embryonic development, and early miscarriages are often not genetically tested. The presence of truncating variants in COSMIC could suggest that LOF variants in some TF genes are intolerant to germline variation and may only be acquired, where the temporal and spatial origin of the expressed variant determining the site of malignant tissue (29, 73) Conversely, low penetrant variants in genes tolerant to variation are often not
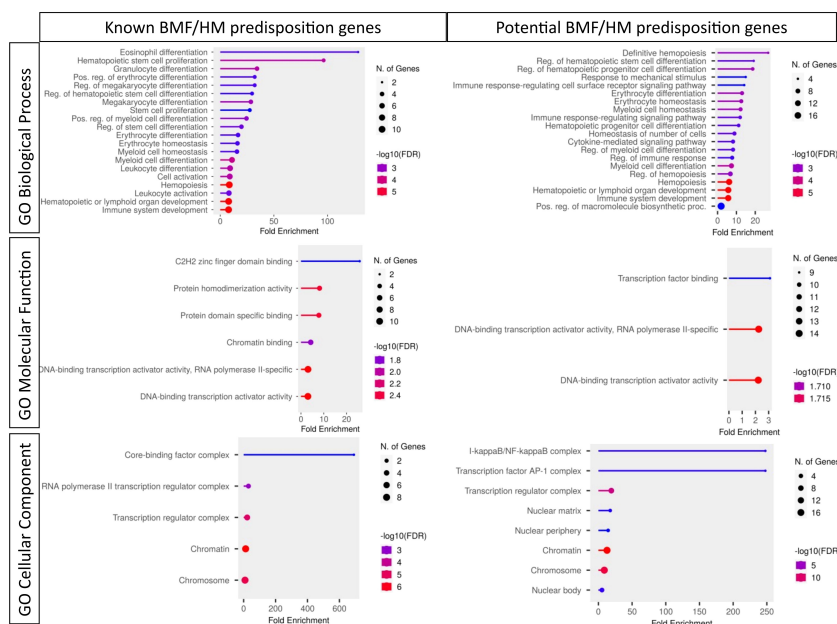
FIGURE 4
Gene Ontology of TFs implicated in hematological development and disease. Using ShinyGO0.77 analysis software (31), the known BMF/HM predisposition list in Table 1 and the potential list of BMF/HM predisposition genes in Table 4 (including *ERG*, *GATA3* and *SPI1*) were analyzed for enrichment of GO terms in biological process and Molecular functions using all known human TF genes as background (i.e., normalization). All protein coding genes were used for background in GO cellular component analysis as there was no enrichment when using TFs as background.

recognized as their individual contribution to disease is limited. However, they may contribute in combination with other genetic variants and/or under influence by environmental factors. Only with global data-sharing approaches and continued expansion of availability to routine genomic sequencing will there be the possibility of reaching the statistical power required to discover risk alleles for BMF and/or HM predisposition.

Clinical implications

Improving understanding of the genetic landscape of predisposition to BMF and/or HM is imperative to provide advice on clinical management and monitoring regimes, tailored therapies, and design targeted therapies (9). While the availability of carefully curated gene panels has led to improving the diagnostic yield of germline causes of BMF and HM, the use of such panels has caveats as novel genes such as those in the potential list of predisposition TF (Table 3) will be missed. As the cost of NGS technologies continues to decline and the access to these technologies' increases, it is inevitable that more predisposing mutations in novel TFs involved in hematopoietic pathways will be added to the list of known predisposing genes. It can be challenging for laboratories and/or hematologists not experienced in germline evaluation to identify individuals most suitable for extensive evaluation for a germline cause (270). A familial history is not always observed, confounded by phenotypic heterogeneity and highly variable penetrance associated with different types of predisposing variants in TFs which will undoubtedly impact multiple different cellular and differentiation pathways, and be impacted by common polygenic modifier variants and external stressors. To alleviate some of these difficulties, laboratories are turning to artificial intelligence machine learning technologies to stratify clinical and genetic variables, to better classify inherited versus sporadic causes of BMF/HM, triaging those individuals who are most likely to benefit from more extensive genetic investigation (271). The importance of identifying a germline cause of disease cannot be understated, with implications for surveillance, treatment protocols and clinical care. In a large study, identification of a germline cause of hematopoietic predisposition syndromes impacted the clinical management of 91% of participants (279). For example, GATA2 carriers considered for BMT/HSCT often require specific protocols including high-dose post-transplant cyclophosphamide after transplant to avoid graft-versus host disease (272). Additionally, due to the likelihood of requirement for HSCT in germline GATA2 carriers, prophylactic HSCT before transformation may be considered (273). Whereas, HSCT is often not recommended for germline *CEBPA* patients, due to good responses to chemotherapy, and the risk of morbidity and mortality associated with HSCT. However, continued molecular monitoring is recommended due to novel somatic variants identified at relapse (67). HSCT transplant is often considered as the only curative therapy for HM, with matched-related donors often considered as the optimal choice to reduce the risk of transplant rejection and graft versus host disease.

Knowledge of a germline causal variant provides the means to screen potential donors thereby reducing the risk of donor-derived leukemias (274). Continued molecular monitoring of carriers is crucial, and evidence is increasing of the unique suite of somatic mutations most likely to occur for each TF predisposition gene during disease progression. As gene/mutation specific drugs become more common-place, clinical management, treatment, and prognosis may change for affected carrier individuals. Clinically, it will be important to consider that these TF BMF/HM predisposition disorders may not be purely monogenic disorders. It is likely that multiple genetic variants (rare and common) in different genes of these highly regulated TF pathways during hematopoiesis are contributing to variable penetrance and phenotypes. For example, a germline RUNX1 family was subsequently found to carry pathogenic germline variants in DDX41 and ANKRD26 (both associated with cytopenia's and HM) in different branches of the family, and several individuals carried multiple variants (9). Future diagnosis/prognosis, management and possible therapy will likely include the use of polygenic risk scores accounting for the pathogenicity or predisposing capacity of the sum of germline variants on a case-by-case rather than familial basis.

Conclusion

TFs play a key role in normal hematopoiesis that when perturbed can give rise to a range of phenotypes including BMF and/or HM. As genomic testing becomes more widely adopted and more comprehensive in its nature, it is likely that germline predisposing variants will be identified in novel genes and non-coding regions of known hematopoietic TF genes. It is likely that these will provide different "fertile soil", predisposing to subtly new collections of phenotypes with unique penetrance's under different stressors leading to BMF, cytopenia's and/or clonal outgrowth as HM.

Author contributions

JZ, CCH, PA, AB, HS and CNH contributed to the design, ideas and structure of the manuscript. JZ, CCH and CNH wrote the review. All authors contributed to the article and approved the submitted version.

Acknowledgments

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2023.1183318/full#supplementary-material

SUPPLEMENTARY FIGURE 1
(A) Known interacting partners of HM and/or BMF predisposition TFs. The STRING database was used to map the proteins which are part of a physical complex with each known TF. In the bold blue text are the known predisposition TFs. Line thickness correlates with the strength of the association. **(B)** Interactions between TFs implicated in predisposition to BMF and HM. The STRING database was used to map known and predicted protein-protein interactions (12). Interactions include direct (physical) and indirect (functional) associations between, (i) known predisposition TFs and (ii) known and predicted predisposition TFs. Line thickness correlates with the strength of the association.

References

1. Martin AR, Williams E, Foulger RE, Leigh S, Daugherty LC, Niblock O, et al. PanelApp crowdsources expert knowledge to establish consensus diagnostic gene panels. *Nat Genet* (2019) 51(11):1560–5. doi: 10.1038/s41588-019-0528-2

2. Joung J, Ma S, Tay T, Geiger-Schuller KR, Kirchgatterer PC, Verdine VK, et al. A transcription factor atlas of directed differentiation. *Cell* (2023) 186(1):209–29.e26. doi: 10.1016/j.cell.2022.11.026

3. Lambert SA, Jolma A, Campitelli LF, Das PK, Yin Y, Albu M, et al. The human transcription factors. *Cell* (2018) 175(2):598–9. doi: 10.1016/j.cell.2018.09.045

4. Chacon D, Beck D, Perera D, Wong JWH, Pimanda JE. BloodChIP: a database of comparative genome-wide transcription factor binding profiles in human blood cells. *Nucleic Acids Res* (2014) 42(Database issue):D172–7. doi: 10.1093/nar/gkt1036

5. Wang Z, Wang P, Li Y, Peng H, Zhu Y, Mohandas N, et al. Interplay between cofactors and transcription factors in hematopoiesis and hematological malignancies. *Signal Transduct Target Ther* (2021) 6(1):24. doi: 10.1038/s41392-020-00422-1

6. Luo X, Feurstein S, Mohan S, Porter CC, Jackson SA, Keel S, et al. ClinGen myeloid malignancy variant curation expert panel recommendations for germline RUNX1 variants. *Blood Adv* (2019) 3(20):2962–79. doi: 10.1182/bloodadvances.2019000644

7. Chong CE, Venugopal P, Stokes PH, Lee YK, Brautigan PJ, Yeung DTO, et al. Differential effects on gene transcription and hematopoietic differentiation correlate with GATA2 mutant disease phenotypes. *Leukemia* (2018) 32(1):194–202. doi: 10.1038/leu.2017.196

8. Lee TI, Young RA. Transcriptional regulation and its misregulation in disease. *Cell* (2013) 152(6):1237–51. doi: 10.1016/j.cell.2013.02.014

9. Homan CC, Scott HS, Brown AL. Hereditary platelet disorders associated with germline variants in RUNX1, ETV6 and ANKRD26. *Blood* (2023) 141(13):1533–43. doi: 10.1182/blood.2022017735

10. Rudelius M, Weinberg OK, Niemeyer CM, Shimamura A, Calvo KR. The international consensus classification (ICC) of hematologic neoplasms with germline predisposition, pediatric myelodysplastic syndrome, and juvenile myelomonocytic leukemia. *Virchows Arch* (2023) 482(1):113–30. doi: 10.1007/s00428-022-03447-9

11. Songdej N, Rao AK. Hematopoietic transcription factor mutations: important players in inherited platelet defects. *Blood* (2017) 129(21):2873–81. doi: 10.1182/blood-2016-11-709881

12. Szklarczyk D, Kirsch R, Koutrouli M. The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids* (2023) 51(D1):D638–46. doi: 10.1093/nar/gkac1000

13. Rouillard AD, Gundersen GW, Fernandez NF, Wang Z, Monteiro CD, McDermott MG, et al. The harmonizome: a collection of processed datasets gathered to serve and mine knowledge about genes and proteins. *Database* (2016) 2016. doi: 10.1093/database/baw100

14. Ferreira R, Ohneda K, Yamamoto M, Philipsen S. GATA1 function, a paradigm for transcription factors in hematopoiesis. *Mol Cell Biol* (2005) 25(4):1215–27. doi: 10.1128/MCB.25.4.1215-1227.2005

15. Tindemans I, Serafini N, Di Santo JP, Hendriks RW. GATA-3 function in innate and adaptive immunity. *Immunity* (2014) 41(2):191–206. doi: 10.1016/j.immuni.2014.06.006

16. Ichikawa M, Yoshimi A, Nakagawa M, Nishimoto N, Watanabe-Okochi N, Kurokawa M. A role for RUNX1 in hematopoiesis and myeloid leukemia. *Int J Hematol* (2013) 97(6):726–34. doi: 10.1007/s12185-013-1347-3

17. Noh JY. Megakaryopoiesis and platelet biology: roles of transcription factors and emerging clinical implications. *Int J Mol Sci* (2021) 22(17):9615. doi: 10.3390/ijms22179615

18. Zhu J, Emerson SG. Hematopoietic cytokines, transcription factors and lineage commitment. *Oncogene* (2002) 21(21):3295–313. doi: 10.1038/sj.onc.1205318

19. Imperato MR, Cauchy P, Obier N, Bonifer C. The RUNX1-PU.1 axis in the control of hematopoiesis. *Int J Hematol* (2015) 101(4):319–29. doi: 10.1007/s12185-015-1762-8

20. Nakajima H. Role of transcription factors in differentiation and reprogramming of hematopoietic cells. *Keio J Med* (2011) 60(2):47–55. doi: 10.2302/kjm.60.47

21. Guan Y, Tiwari AD, Phillips JG, Hasipek M, Grabowski DR, Pagliuca S, et al. A therapeutic strategy for preferential targeting of TET2 mutant and TET-dioxygenase deficient cells in myeloid neoplasms. *Blood Cancer Discovery* (2021) 2(2):146–61. doi: 10.1158/2643-3230.BCD-20-0173

22. Avagyan S, Henninger JE, Mannherz WP, Mistry M, Yoon J, Yang S, et al. Resistance to inflammation underlies enhanced fitness in clonal hematopoiesis. *Science* (2021) 374(6568):768–72. doi: 10.1126/science.aba9304

23. Homan CC, King-Smith SL, Lawrence DM, Arts P, Feng J, Andrews J, et al. The RUNX1 database (RUNX1db): establishment of an expert curated RUNX1 registry and genomics database as a public resource for familial platelet disorder with myeloid malignancy. *Haematologica* (2021) 106(11):3004–7. doi: 10.3324/haematol.2021.278762

24. Gutierrez-Rodrigues F, Munger E, Ma X, Groarke EM, Tang Y, Patel BA, et al. Differential diagnosis of bone marrow failure syndromes guided by machine learning. *Blood* (2022) 141(17):2100–13. doi: 10.1182/blood.2022017518

25. Martin ES, Ferrer A, Mangaonkar AA, Khan SP, Kohorst MA, Joshi AY, et al. Spectrum of hematological malignancies, clonal evolution and outcomes in 144 Mayo clinic patients with germline predisposition syndromes. *Am J Hematol* (2021) 96(11):1450–60. doi: 10.1002/ajh.26321

26. Parta M, NN S, Baird K, Rafei H, KR C, Hughes T, et al. Allogeneic hematopoietic stem cell transplantation for GATA2 deficiency using a busulfan-based regimen. *Biol Blood Marrow Transplant* (2018) 24(6):1250–9. doi: 10.1016/j.bbmt.2018.01.030

27. Jørgensen SF, Buechner J, Myhre AE, Galteland E, Spetalen S, Kulseth MA, et al. A nationwide study of GATA2 deficiency in Norway-the majority of patients have undergone allo-HSCT. *J Clin Immunol* (2022) 42(2):404–20. doi: 10.1007/s10875-021-01189-y

28. Williams L, Doucette K, Karp JE, Lai C. Genetics of donor cell leukemia in acute myelogenous leukemia and myelodysplastic syndrome. *Bone Marrow Transplant* (2021) 56(7):1535–49. doi: 10.1038/s41409-021-01214-z

29. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* (2020) 581(7809):434–43. doi: 10.1038/s41586-020-2308-7

30. Wiel L, Baakman C, Gilissen D, Veltman JA, Vriend G, Gilissen C. MetaDome: pathogenicity analysis of genetic variants through aggregation of homologous human protein domains. *Hum Mutat* (2019) 40(8):1030–8. doi: 10.1002/humu.23798

31. Ge SX, Jung D, Yao R. ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics* (2020) 36(8):2628–9. doi: 10.1093/bioinformatics/btz931

32. Khandanpour C, Eisfeld C, Nimmagadda SC, Raab MS, Weinhold N, Seckinger A, et al. Prevalence of the GFI1-36N SNP in multiple myeloma patients and its impact on the prognosis. *Front Oncol* (2021) 11(October):757664. doi: 10.3389/fonc.2021.757664

33. Arber DA, Orazi A, Hasserjian R, Thiele J, Borowitz MJ, Le Beau MM, et al. The 2016 revision to the world health organization classification of myeloid neoplasms and acute leukemia. *Blood* (2016) 127(20):2391–405. doi: 10.1182/blood-2016-03-643544

34. Khoury JD, Solary E, Abla O, Akkari Y, Alaggio R, Apperley JF, et al. The 5th edition of the world health organization classification of haematolymphoid tumours: myeloid and Histiocytic/Dendritic neoplasms. *Leukemia* (2022) 36(7):1703–19. doi: 10.1038/s41375-022-01613-1

35. Hamilton KV, Maese L, Marron JM, Pulsipher MA, Porter CC, Nichols KE. Stopping leukemia in its tracks: should preemptive hematopoietic stem-cell transplantation be offered to patients at increased genetic risk for acute myeloid leukemia? *J Clin Oncol* (2019) 37(24):2098–104. doi: 10.1200/JCO.19.00181

36. Wilson NK, Foster SD, Wang X, Knezevic K, Schütte J, Kaimakis P, et al. Combinatorial transcriptional control in blood stem/progenitor cells: genome-wide analysis of ten major transcriptional regulators. *Cell Stem Cell* (2010) 7(4):532–44. doi: 10.1016/j.stem.2010.07.016

37. Thoms JAI, Truong P, Subramanian S, Knezevic K, Harvey G, Huang Y, et al. Disruption of a GATA2-TAL1-ERG regulatory circuit promotes erythroid transition in healthy and leukemic stem cells. *Blood* (2021) 138(16):1441–55. doi: 10.1182/blood.2020009707

38. Thoms JAI, Beck D, Pimanda JE. Transcriptional networks in acute myeloid leukemia. *Genes Chromosomes Cancer* (2019) 58(12):859–74. doi: 10.1002/gcc.22794

39. Diffner E, Beck D, Gudgin E, Thoms JAI, Knezevic K, Pridans C, et al. Activity of a heptad of transcription factors is associated with stem cell programs and clinical outcome in acute myeloid leukemia. *Blood* (2013) 121(12):2289–300. doi: 10.1182/blood-2012-07-446120

40. Koyunlar C, de Pater E. From basic biology to patient mutational spectra of GATA2 haploinsufficiencies: what are the mechanisms, hurdles, and prospects of genome editing for treatment. *Front Genome Ed* (2020) 2:602182. doi: 10.3389/fgeed.2020.602182

41. Tarlock K, Lamble AJ, Wang YC, Gerbing RB, Ries RE, Loken MR, et al. CEBPA-bZip mutations are associated with favorable prognosis in *de novo* AML: a report from the children's oncology group. *Blood* (2021) 138(13):1137–47. doi: 10.1182/blood.2020009652

42. Fasan A, Eder C, Haferlach C, Grossmann V, Kohlmann A, Dicker F, et al. GATA2 mutations are frequent in intermediate-risk karyotype AML with biallelic CEBPA mutations and are associated with favorable prognosis. *Leukemia* (2013) 27(2):482–5. doi: 10.1038/leu.2012.174

43. Luesink M, Hollink IHIM, van der Velden VHJ, Knops RHJN, Boezeman JBM, de Haas V, et al. High GATA2 expression is a poor prognostic marker in pediatric acute myeloid leukemia. *Blood* (2012) 120(10):2064–75. doi: 10.1182/blood-2011-12-397083

44. Altıner Ş, Kutlay NY, İlhan O. Constitutional trisomy 8 mosaicism with persistent macrocytosis. *Cytogenet Genome Res* (2016) 150(1):35–9. doi: 10.1159/000452358

45. Gong Y, Deng J, Wu X. Germline mutations and blood malignancy (Review). *Oncol Rep* (2021) 45(1):49–57. doi: 10.3892/or.2020.7846

46. Hahn CN, Chong CE, Carmichael CL, Wilkins EJ, Brautigan PJ, Li XC, et al. Heritable GATA2 mutations associated with familial myelodysplastic syndrome and acute myeloid leukemia. *Nat Genet* (2011) 43(10):1012–7. doi: 10.1038/ng.913

47. Cortés-Lavaud X, Landecho MF, Maicas M, Urquiza L, Merino J, Moreno-Miralles I, et al. GATA2 germline mutations impair GATA2 transcription, causing haploinsufficiency: functional analysis of the p.Arg396Gln mutation. *J Immunol* (2015) 194(5):2190–8. doi: 10.4049/jimmunol.1401868

48. Sun R, Sun L, Xie X, Li X, Wu P, Wang L, et al. Single-cell analysis of transcription factor regulatory networks reveals molecular basis for subtype-specific dysregulation in acute myeloid leukemia. *Blood Sci* (2022) 4(2):65–75. doi: 10.1097/BS9.0000000000000113

49. Bergiers I, Andrews T, Vargel Bölükbaşı Ö, Buness A, Janosz E, Lopez-Anguita N, et al. Single-cell transcriptomics reveals a new dynamical function of transcription factors during embryonic hematopoiesis. *Elife* (2018) 7:e29312. doi: 10.7554/eLife.29312

50. Trottier AM, Godley LA. Inherited predisposition to haematopoietic malignancies: overcoming barriers and exploring opportunities. *Br J Haematol* (2021) 194(4):663–76. doi: 10.1111/bjh.17247

51. Song WJ, Sullivan MG, Legare RD, Hutchings S, Tan X, Kufrin D, et al. Haploinsufficiency of CBFA2 causes familial thrombocytopenia with propensity to develop acute myelogenous leukaemia. *Nat Genet* (1999) 23(2):166–75. doi: 10.1038/13793

52. Smith ML, Cavenagh JD, Lister TA, Fitzgibbon J. Mutation of CEBPA in familial acute myeloid leukemia. *N Engl J Med* (2004) 351(23):2403–7. doi: 10.1056/NEJMoa041331

53. Homan CC, Venugopal P. GATA2 deficiency syndrome: a decade of discovery. *Hum Mutat* (2021) 42(11):1399–421. doi: 10.22541/au.161453332.24299786/v1

54. Brown AL, Hahn CN, Scott HS. Secondary leukemia in patients with germline transcription factor mutations (RUNX1, GATA2, CEBPA). *Blood* (2020) 136(1):24–35. doi: 10.1182/blood.2019000937

55. Hayashi Y, Harada Y, Harada H. Myeloid neoplasms and clonal hematopoiesis from the RUNX1 perspective. *Leukemia* (2022) 36(5):1203–14. doi: 10.1038/s41375-022-01548-7

56. Bagger FO, Kinalis S, Rapin N. BloodSpot: a database of healthy and malignant haematopoiesis updated with purified and single cell mRNA sequencing profiles. *Nucleic Acids Res* (2019) 47(D1):D881–5. doi: 10.1093/nar/gky1076

57. Uddin MDM, Nguyen NQH, Yu B, Brody JA, Pampana A, Nakao T, et al. Clonal hematopoiesis of indeterminate potential, DNA methylation, and risk for coronary artery disease. *Nat Commun* (2022) 13(1):5350. doi: 10.1038/s41467-022-33093-3

58. de Bruijn M, Dzierzak E. Runx transcription factors in the development and function of the definitive hematopoietic system. *Blood* (2017) 129(15):2061–9. doi: 10.1182/blood-2016-12-689109

59. Mevel R, Draper JE, Lie-A-Ling M, Kouskoff V, Lacaud G. RUNX transcription factors: orchestrators of development. *Development* (2019) 146(17):dev148296. doi: 10.1242/dev.148296

60. Hass MR, Brissette D, Parameswaran S, Pujato M, Donmez O, Kottyan LC, et al. Runx1 shapes the chromatin landscape *via* a cascade of direct and indirect targets. *PloS Genet* (2021) 17(6):e1009574. doi: 10.1371/journal.pgen.1009574

61. Wehr C, Grotius K, Casadei S, Bleckmann D, Bode SFN, Frye BC, et al. A novel disease-causing synonymous exonic mutation in GATA2 affecting RNA splicing. *Blood* (2018) 132(11):1211–5. doi: 10.1182/blood-2018-03-837336

62. Kozyra EJ, Pastor VB, Lefkopoulos S, Sahoo SS, Busch H, Voss RK, et al. Synonymous GATA2 mutations result in selective loss of mutated RNA and are common in patients with GATA2 deficiency. *Leukemia* (2020) 34(10):2673–87. doi: 10.1038/s41375-020-0899-5

63. Fox LC, Tan M, Brown AL, Arts P, Thompson E, Ryland GL, et al. A synonymous GATA2 variant underlying familial myeloid malignancy with striking intrafamilial phenotypic variability. *Br J Haematol* (2020) 190(5):e297–301. doi: 10.1111/bjh.16819

64. Guidugli L, Johnson AK, Alkorta-Aranburu G, Nelakuditi V, Arndt K, Churpek JE, et al. Clinical utility of gene panel-based testing for hereditary myelodysplastic syndrome/acute leukemia predisposition syndromes. *Leukemia* (2017) 31(5):1226–9. doi: 10.1038/leu.2017.28

65. Drazer MW, Kadri S, Sukhanova M, Patil SA, West AH, Feurstein S, et al. Prognostic tumor sequencing panels frequently identify germ line variants associated with hereditary hematopoietic malignancies. *Blood Adv* (2018) 2(2):146–50. doi: 10.1182/bloodadvances.2017013037

66. Leubolt G, Redondo Monte E, Greif PA. GATA2 mutations in myeloid malignancies: two zinc fingers in many pies. *IUBMB Life* (2020) 72(1):151–8. doi: 10.1002/iub.2204

67. Tawana K, Wang J, Renneville A, Bodor C, Hills R, Loveday C, et al. Disease evolution and outcomes in familial AML with germline CEBPA mutations. *Blood* (2015) 126(10):1214–23. doi: 10.1182/blood-2015-05-647172

68. Tawana K, Fitzgibbon J. *CEBPA-associated familial acute myeloid leukemia (AML)*. Seattle: University of Washington (2021).

69. Tawana K, Rio-Machin A, Preudhomme C, Fitzgibbon J. Familial CEBPA-mutated acute myeloid leukemia. *Semin Hematol* (2017) 54(2):87–93. doi: 10.1053/j.seminhematol.2017.04.001

70. Wu D, Sunkel B, Chen Z, Liu X, Ye Z, Li Q, et al. Three-tiered role of the pioneer factor GATA2 in promoting androgen-dependent gene expression in prostate cancer. *Nucleic Acids Res* (2014) 42(6):3607–22. doi: 10.1093/nar/gkt1382

71. Li Y, Yang W, Devidas M, Winter SS, Kesserwan C, Yang W, et al. Germline RUNX1 variation and predisposition to childhood acute lymphoblastic leukemia. *J Clin Invest* (2021) 131(17):e147898. doi: 10.1172/JCI147898

72. Förster A, Decker M, Schlegelberger B, Ripperger T. Beyond pathogenic RUNX1 germline variants: the spectrum of somatic alterations in RUNX1-familial platelet disorder with predisposition to hematologic malignancies. *Cancers* (2022) 14(14):3431. doi: 10.3390/cancers14143431

73. Tate JG, Bamford S, Jubb HC, Sondka Z, Beare DM, Bindal N, et al. COSMIC: the catalogue of somatic mutations in cancer. *Nucleic Acids Res* (2019) 47(D1):D941–7. doi: 10.1093/nar/gky1015

74. Largeaud L, Collin M, Monselet N, Vergez F, Fregona V, Larcher L, et al. Somatic genetic alterations predict haematological progression in GATA2 deficiency. *Haematologica* (2023). doi: 10.3324/haematol.2022.282250

75. Porter CC, Di Paola J, Pencheva B. ETV6 thrombocytopenia and predisposition to leukemia. In: Adam MP, Everman DB, Mirzaa GM, Pagon RA, Wallace SE, Bean LJH, et al, editors. *GeneReviews*®. Seattle (WA: University of Washington, Seattle (2020).

76. Nishii R, Baskin-Doerfler R, Yang W, Oak N, Zhao X, Yang W, et al. Molecular basis of ETV6-mediated predisposition to childhood acute lymphoblastic leukemia. *Blood* (2021) 137(3):364–73. doi: 10.1182/blood.2020006164

77. Hart A, Melet F, Grossfeld P, Chien K, Jones C, Tunnacliffe A, et al. Fli-1 is required for murine vascular and megakaryocytic development and is hemizygously deleted in patients with thrombocytopenia. *Immunity* (2000) 13(2):167–77. doi: 10.1016/S1074-7613(00)00017-0

78. Stockley J, Morgan NV, Bem D, Lowe GC, Lordkipanidzé M, Dawood B, et al. Enrichment of FLI1 and RUNX1 mutations in families with excessive bleeding and platelet dense granule secretion defects. *Blood* (2013) 122(25):4090–3. doi: 10.1182/blood-2013-06-506873

79. Di Paola J, Porter CC. ETV6-related thrombocytopenia and leukemia predisposition. *Blood* (2019) 134(8):663–7. doi: 10.1182/blood.2019852418

80. Saultier P, Vidal L, Canault M, Bernot D, Falaise C, Pouymayou C, et al. Macrothrombocytopenia and dense granule deficiency associated with FLI1 variants: ultrastructural and pathogenic features. *Haematologica* (2017) 102(6):1006–16. doi: 10.3324/haematol.2016.153577

81. Stevenson WS, Rabbolini DJ, Beutler L, Chen Q, Gabrielli S, Mackay JP, et al. Paris-Trousseau thrombocytopenia is phenocopied by the autosomal recessive inheritance of a DNA-binding domain mutation in FLI1. *Blood* (2015) 126 (17):2027–30. doi: 10.1182/blood-2015-06-650887

82. Stankiewicz MJ, Crispino JD. ETS2 and ERG promote megakaryopoiesis and synergize with alterations in GATA-1 to immortalize hematopoietic progenitor cells. *Blood* (2009) 113(14):3337–47. doi: 10.1182/blood-2008-08-174813

83. Loughran SJ, Kruse EA, Hacking DF, de Graaf CA, Hyland CD, Willson TA, et al. The transcription factor erg is essential for definitive hematopoiesis and the function of adult hematopoietic stem cells. *Nat Immunol* (2008) 9(7):810–9. doi: 10.1038/ni.1617

84. Rothenberg EV, Hosokawa H, Ungerbäck J. Mechanisms of action of hematopoietic transcription factor PU.1 in initiation of T-cell development. *Front Immunol* (2019) 10:228. doi: 10.3389/fimmu.2019.00228

85. Feurstein S, Godley LA. Germline ETV6 mutations and predisposition to hematological malignancies. *Int J Hematol* (2017) 106(2):189–95. doi: 10.1007/s12185-017-2259-4

86. Hsu AP, Johnson KD, Falcone EL, Sanalkumar R, Sanchez L, Hickstein DD, et al. GATA2 haploinsufficiency caused by mutations in a conserved intronic element leads to MonoMAC syndrome. *Blood* (2013) 121(19):3830–7, S1–7. doi: 10.1182/blood-2012-08-452763

87. Beesley J, Sivakumaran H, Moradi Marjaneh M, Shi W, Hillman KM, Kaufmann S, et al. eQTL colocalization analyses identify NTN4 as a candidate breast cancer risk gene. *Am J Hum Genet* (2020) 107(4):778–87. doi: 10.1016/j.ajhg.2020.08.006

88. Moradi Marjaneh M, Beesley J, O'Mara TA, Mukhopadhyay P, Koufariotis LT, Kazakoff S, et al. Non-coding RNAs underlie genetic predisposition to breast cancer. *Genome Biol* (2020) 21(1):7. doi: 10.1186/s13059-019-1876-z

89. Mangaonkar AA, Patnaik MM. Hereditary predisposition to hematopoietic neoplasms: when bloodline matters for blood cancers. *Mayo Clin Proc* (2020) 95 (7):1482–98. doi: 10.1016/j.mayocp.2019.12.013

90. Harland M, Petljak M, Robles-Espinoza CD, Ding Z, Gruis NA, van Doorn R, et al. Germline TERT promoter mutations are rare in familial melanoma. *Fam Cancer* (2016) 15(1):139–44. doi: 10.1007/s10689-015-9841-9

91. Järviaho T, Bang B, Zachariadis V, Taylan F, Moilanen J, Möttönen M, et al. Predisposition to childhood acute lymphoblastic leukemia caused by a constitutional translocation disrupting ETV6. *Blood Adv* (2019) 3(18):2722–31. doi: 10.1182/bloodadvances.2018028795

92. Raynaud S, Cave H, Baens M, Bastard C, Cacheux V, Grosgeorge J, et al. The 12;21 translocation involving TEL and deletion of the other TEL allele: two frequently associated alterations found in childhood acute lymphoblastic leukemia. *Blood* (1996) 87(7):2891–9. doi: 10.1182/blood.V87.7.2891.bloodjournal8772891

93. Zhang H, Zhan Q, Wang X, Gao F, Yu J, Wang J, et al. TLS/FUS-ERG fusion gene in acute leukemia and myelodysplastic syndrome evolved to acute leukemia: report of six cases and a literature review. *Ann Hematol* (2022) 101(12):2583–600. doi: 10.1007/s00277-022-04979-5

94. Borst S, Nations CC, Klein JG, Pavani G, Maguire JA, Camire RM, et al. Study of inherited thrombocytopenia resulting from mutations in ETV6 or RUNX1 using a human pluripotent stem cell model. *Stem Cell Rep* (2021) 16(6):1458–67. doi: 10.1016/j.stemcr.2021.04.013

95. Drazer MW, Homan CC, Yu K, Cavalcante de Andrade Silva M, McNeely KE, Pozsgai MJ, et al. Clonal hematopoiesis in patients with ANKRD26 or ETV6 germline mutations. *Blood Adv* (2022) 6(15):4357–9. doi: 10.1182/bloodadvances.2022007211

96. Martens JHA. Acute myeloid leukemia: a central role for the ETS factor ERG. *Int J Biochem Cell Biol* (2011) 43(10):1413–6. doi: 10.1016/j.biocel.2011.05.014

97. Baldus CD, Liyanarachchi S, Mrózek K, Auer H, Tanner SM, Guimond M, et al. Acute myeloid leukemia with complex karyotypes and abnormal chromosome 21: amplification discloses overexpression of APP, ETS2, and ERG genes. *Proc Natl Acad Sci USA* (2004) 101(11):3915–20. doi: 10.1073/pnas.0400272101

98. Valk PJM, Verhaak RGW, Beijen MA, Erpelinck CAJ, Barjesteh van Waalwijk van Doorn-Khosrovani S, Boer JM, et al. Prognostically useful gene-expression profiles in acute myeloid leukemia. *N Engl J Med* (2004) 350(16):1617–28. doi: 10.1056/NEJMoa040465

99. Marcucci G, Baldus CD, Ruppert AS, Radmacher MD, Mrózek K, Whitman SP, et al. Overexpression of the ETS-related gene, ERG, predicts a worse outcome in acute myeloid leukemia with normal karyotype: a cancer and leukemia group b study. *J Clin Oncol* (2005) 23(36):9234–42. doi: 10.1200/JCO.2005.03.6137

100. Metzeler KH, Dufour A, Benthaus T, Hummel M, Sauerland MC, Heinecke A, et al. ERG expression is an independent prognostic factor and allows refined risk stratification in cytogenetically normal acute myeloid leukemia: a comprehensive analysis of ERG, MN1, and BAALC transcript levels using oligonucleotide microarrays. *J Clin Oncol* (2009) 27(30):5031–8. doi: 10.1200/JCO.2008.20.5328

101. Rimmelé P, Kosmider O, Mayeux P, Moreau-Gachelin F, Guillouf C. Spi-1/PU.1 participates in erythroleukemogenesis by inhibiting apoptosis in cooperation with

epo signaling and by blocking erythroid differentiation. *Blood* (2007) 109(7):3007–14. doi: 10.1182/blood-2006-03-006718

102. Rosenbauer F, Wagner K, Kutok JL, Iwasaki H, Le Beau MM, Okuno Y, et al. Acute myeloid leukemia induced by graded reduction of a lineage-specific transcription factor, PU.1. *Nat Genet* (2004) 36(6):624–30. doi: 10.1038/ng1361

103. Kuehn HS, Boisson B, Cunningham-Rundles C, Reichenbach J, Stray-Pedersen A, Gelfand EW, et al. Loss of b cells in patients with heterozygous mutations in IKAROS. *N Engl J Med* (2016) 374(11):1032–43. doi: 10.1056/NEJMoa1512234

104. Shahin T, Mayr D, Shoeb MR, Kuehn HS, Hoeger B, Giuliani S, et al. Identification of germline monoallelic mutations in IKZF2 in patients with immune dysregulation. *Blood Adv* (2022) 6(7):2444–51. doi: 10.1182/bloodadvances.2021006367

105. Lazarian G, Yin S, Ten Hacken E, Sewastianik T, Uduman M, Font-Tello A, et al. A hotspot mutation in transcription factor IKZF3 drives b cell neoplasia *via* transcriptional dysregulation. *Cancer Cell* (2021) 39(3):380–93.e8. doi: 10.1016/j.ccell.2021.02.003

106. Kamihara J, Shimamura A. It's ALL in the family: IKZF1 and hereditary leukemia. *Cancer Cell* (2018) 33(5):798–800. doi: 10.1016/j.ccell.2018.04.008

107. Kuehn HS, Chang J, Yamashita M, Niemela JE, Zou C, Okuyama K, et al. T And b cell abnormalities, pneumocystis pneumonia, and chronic lymphocytic leukemia associated with an AIOLOS defect in patients. *J Exp Med* (2021) 218(12):e20211118. doi: 10.1084/jem.20211118

108. Churchman ML, Qian M, Te Kronnie G, Zhang R, Yang W, Zhang H, et al. Germline genetic IKZF1 variation and predisposition to childhood acute lymphoblastic leukemia. *Cancer Cell* (2018) 33(5):937–48.e8. doi: 10.1016/j.ccell.2018.03.021

109. Shin SY, Lee H, Lee ST, Choi JR, Jung CW, Koo HH, et al. Recurrent somatic mutations and low germline predisposition mutations in Korean ALL patients. *Sci Rep* (2021) 11(1):8893. doi: 10.1038/s41598-021-88449-4

110. Yoshida N, Sakaguchi H, Muramatsu H, Okuno Y, Song C, Dovat S, et al. Germline IKAROS mutation associated with primary immunodeficiency that progressed to T-cell acute lymphoblastic leukemia. *Leukemia* (2017) 31(5):1221–3. doi: 10.1038/leu.2017.25

111. Brodie SA, Khincha PP, Giri N, Bouk AJ, Steinberg M, Dai J, et al. Pathogenic germline IKZF1 variant alters hematopoietic gene expression profiles. *Cold Spring Harb Mol Case Stud* (2021) 7(4):a6015. doi: 10.1101/mcs.a006015

112. Lopes BA, Barbosa TC, Souza BKS, Poubel CP, Pombo-de-Oliveira MS, Emerenciano M, et al. IKZF1 gene in childhood b-cell precursor acute lymphoblastic leukemia: interplay between genetic susceptibility and somatic abnormalities. *Cancer Prev Res* (2017) 10(12):738–44. doi: 10.1158/1940-6207.CAPR-17-0121

113. Sun W, Guo J, McClellan D, Poeschla A, Bareyan D, Casey MJ, et al. GFI1 cooperates with IKZF1/IKAROS to activate gene expression in T-cell acute lymphoblastic leukemia. *Mol Cancer Res* (2022) 20(4):501–14. doi: 10.1158/1541-7786.MCR-21-0352

114. Dörge P, Meissner B, Zimmermann M, Möricke A, Schrauder A, Bouquin JP, et al. IKZF1 deletion is an independent predictor of outcome in pediatric acute lymphoblastic leukemia treated according to the ALL-BFM 2000 protocol. *Haematologica* (2013) 98(3):428–32. doi: 10.3324/haematol.2011.056135

115. Sun L, Liu A, Georgopoulos K. Zinc finger-mediated protein interactions modulate ikaros activity, a molecular control of lymphocyte development. *EMBO J* (1996) 15(19):5358–69. doi: 10.1002/j.1460-2075.1996.tb00920.x

116. Churchman ML, Mullighan CG. Ikaros: exploiting and targeting the hematopoietic stem cell niche in b-progenitor acute lymphoblastic leukemia. *Exp Hematol* (2017) 46:1–8. doi: 10.1016/j.exphem.2016.11.002

117. Ayón-Pérez MF, Pimentel-Gutiérrez HJ, Durán-Avelar M de J, Vibanco-Pérez N, Pérez-Peraza VM, Pérez-González ÓA, et al. IKZF1 gene deletion in pediatric patients diagnosed with acute lymphoblastic leukemia in Mexico. *Cytogenet Genome Res* (2019) 158(1):10–6. doi: 10.1159/000499641

118. Shah S, Schrader KA, Waanders E, Timms AE, Vijai J, Miething C, et al. A recurrent germline PAX5 mutation confers susceptibility to pre-b cell acute lymphoblastic leukemia. *Nat Genet* (2013) 45(10):1226–31. doi: 10.1038/ng.2754

119. Jia Z, Gu Z. PAX5 alterations in b-cell acute lymphoblastic leukemia. *Front Oncol* (2022) 12:1023606. doi: 10.3389/fonc.2022.1023606

120. Auer F, Ingenhag D, Bhatia S, Enczmann J, Cobaleda C, Sanchez-Garcia I, et al. GEMMs addressing Pax5 loss-of-function in childhood pB-ALL. *Eur J Med Genet* (2016) 59(3):166–72. doi: 10.1016/j.ejmg.2015.11.009

121. Auer F, Rüschendorf F, Gombert M, Husemann P, Ginzel S, Izraeli S, et al. Inherited susceptibility to pre b-ALL caused by germline transmission of PAX5 c.547G>A. *Leukemia* (2014) 28(5):1136–8. doi: 10.1038/leu.2013.363

122. Hyde RK, Liu PP. Germline PAX5 mutations and b cell leukemia. *Nat Genet* (2013) 45(10):1104–5. doi: 10.1038/ng.2778

123. Escudero A, Takagi M, Auer F, Friedrich UA, Miyamoto S, Ogawa A, et al. Clinical and immunophenotypic characteristics of familial leukemia predisposition caused by PAX5 germline variants. *Leukemia* (2022) 36(9):2338–42. doi: 10.1038/s41375-022-01661-7

124. Kohlmann W, Schiffman JD. Discussing and managing hematologic germ line variants. *Blood* (2016) 128(21):2497–503. doi: 10.1182/blood-2016-06-716704

125. Duployez N, Jamrog LA, Fregona V, Hamelle C, Fenwarth L, Lejeune S, et al. Germline PAX5 mutation predisposes to familial b-cell precursor acute lymphoblastic leukemia. *Blood* (2021) 137(10):1424–8. doi: 10.1182/blood.2020005756

126. Jia Z, Hu Z, Damirchi B, Han TT, Gu Z. Characterization of PAX5 mutations in b progenitor acute lymphoblastic leukemia. *Blood* (2022) 140(Supplement 1):1001–2. doi: 10.1182/blood-2022-169975

127. Fears S, Mathieu C, Zeleznik-Le N, Huang S, Rowley JD, Nucifora G. Intergenic splicing of MDS1 and EVI1 occurs in normal tissues as well as in myeloid leukemia and produces a new member of the PR domain family. *Proc Natl Acad Sci USA* (1996) 93 (4):1642–7. doi: 10.1073/pnas.93.4.1642

128. Ripperger T, Hofmann W, Koch JC, Shirneshan K, Haase D, Wulf G, et al. MDS1 and EVI1 complex locus (MECOM): a novel candidate gene for hereditary hematological malignancies. *Haematologica* (2018) 103(2):e55–8. doi: 10.3324/haematol.2017.178723

129. van der Veken LT, Maiburg MC, Groenendaal F. *Lethal neonatal bone marrow failure syndrome with multiple congenital abnormalities, including limb defects, due to a constitutional deletion of 3' MECOM* (2018). Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5865425/.

130. Nielsen M, Vermont CL, Aten E. Deletion of the 3q26 region including the EVI1 and MDS1 genes in a neonate with congenital thrombocytopenia and subsequent aplastic anaemia. *J Med* (2012) 49(9):598–600. doi: 10.1136/jmedgenet-2012-100990

131. Germeshausen M, Ancliff P, Estrada J, Metzler M, Ponstingl E, Rütschle H, et al. MECOM-associated syndrome: a heterogeneous inherited bone marrow failure syndrome with amegakaryocytic thrombocytopenia. *Blood Adv* (2018) 2(6):586–96. doi: 10.1182/bloodadvances.2018016501

132. Niihori T, Ouchi-Uchiyama M, Sasahara Y, Kaneko T, Hashii Y, Irie M, et al. Mutations in MECOM, encoding oncoprotein EVI1, cause radioulnar synostosis with amegakaryocytic thrombocytopenia. *Am J Hum Genet* (2015) 97(6):848–54. doi: 10.1016/j.ajhg.2015.10.010

133. Baldazzi C, Luatti S, Zuffa E, Papayannidis C, Ottaviani E, Marzocchi G, et al. Complex chromosomal rearrangements leading to MECOM overexpression are recurrent in myeloid malignancies with various 3q abnormalities. *Genes Chromosomes Cancer* (2016) 55(4):375–88. doi: 10.1002/gcc.22341

134. Gröschel S, Sanders MA, Hoogenboezem R, de Wit E. A single oncogenic enhancer rearrangement causes concomitant EVI1 and GATA2 deregulation in leukemia. *Cell* (2014) 157(2):369–81. doi: 10.1016/j.cell.2014.02.019

135. Kellaway SG, Keane P, Kennett E, Bonifer C. RUNX1-EVI1 disrupts lineage determination and the cell cycle by interfering with RUNX1 and EVI1 driven gene regulatory networks. *Haematologica* (2021) 106(6):1569–80. doi: 10.3324/haematol.2019.241885

136. Senyuk V, Sinha KK, Li D, Rinaldi CR, Yanamandra S, Nucifora G. Repression of RUNX1 activity by EVI1: a new role for EVI1 in leukemogenesis. *Cancer Res* (2007) 67(12):5658–66. doi: 10.1158/0008-5472.CAN-06-3962

137. Dührsen U, Kratz CP, Flotho C, Lauenstein T, Bommer M, König E, et al. Long-term outcome of hemizygous and heterozygous carriers of a germline GATA1 (G208R) mutation. *Ann Hematol* (2011) 90(3):301–6. doi: 10.1007/s00277-010-1088-9

138. Shimizu R, Yamamoto M. GATA-related hematologic disorders. *Exp Hematol* (2016) 44(8):696–705. doi: 10.1016/j.exphem.2016.05.010

139. Freson K, Matthijs G, Thys C, Mariën P, Hoylaerts MF, Vermylen J, et al. Different substitutions at residue D218 of the X-linked transcription factor GATA1 lead to altered clinical severity of macrothrombocytopenia and anemia and are associated with variable skewed X inactivation. *Hum Mol Genet* (2002) 11(2):147–52. doi: 10.1093/hmg/11.2.147

140. Hollanda LM, Lima CSP, Cunha AF, Albuquerque DM, Vassallo J, Ozelo MC, et al. An inherited mutation leading to production of only the short isoform of GATA-1 is associated with impaired erythropoiesis. *Nat Genet* (2006) 38(7):807–12. doi: 10.1038/ng1825

141. Sankaran VG, Ghazvinian R, Do R, Thiru P, Vergilio JA, Beggs AH, et al. Exome sequencing identifies GATA1 mutations resulting in diamond-blackfan anemia. *J Clin Invest* (2012) 122(7):2439–43. doi: 10.1172/JCI63597

142. Nichols KE, Crispino JD, Poncz M, White JG, Orkin SH, Maris JM, et al. Familial dyserythropoietic anaemia and thrombocytopenia due to an inherited mutation in GATA1. *Nat Genet* (2000) 24(3):266–70. doi: 10.1038/73480

143. Yu C, Niakan KK, Matsushita M, Stamatoyannopoulos G, Orkin SH, Raskind WH. X-Linked thrombocytopenia with thalassemia from a mutation in the amino finger of GATA-1 affecting DNA binding rather than FOG-1 interaction. *Blood* (2002) 100(6):2040–5. doi: 10.1182/blood-2002-02-0387

144. Russo R, Andolfo I, Gambale A, De Rosa G, Manna F, Arillo A, et al. GATA1 erythroid-specific regulation of SEC23B expression and its implication in the pathogenesis of congenital dyserythropoietic anemia type II. *Haematologica* (2017) 102(9):e371–4. doi: 10.3324/haematol.2016.162966

145. Hasle H, Kline RM, Kjeldsen E, Nik-Abdul-Rashid NF, Bhojwani D, Verboon JM, et al. Germline GATA1s-generating mutations predispose to leukemia with acquired trisomy 21 and down syndrome-like phenotype. *Blood* (2022) 139 (21):3159–65. doi: 10.1182/blood.2021011463

146. Hillmer EJ, Zhang H, Li HS, Watowich SS. STAT3 signaling in immunity. *Cytokine Growth Factor Rev* (2016) 31:1–15. doi: 10.1016/j.cytogfr.2016.05.001

147. Haddad E. STAT3: too much may be worse than not enough! *Blood* (2015) 125 (4):583–4. doi: 10.1182/blood-2014-11-610592

148. Mackie J, Ma CS, Tangye SG, Guerin A. The ups and downs of STAT3 function: too much, too little and human immune dysregulation. *Clin Exp Immunol* (2023) 212(2):107–16. doi: 10.1093/cei/uxad007

149. Milner JD, Vogel TP, Forbes L, Ma CA, Stray-Pedersen A, Niemela JE, et al. Early-onset lymphoproliferation and autoimmunity caused by germline STAT3 gain-of-function mutations. *Blood* (2015) 125(4):591–9. doi: 10.1182/blood-2014-09-602763

150. Koskela HLM, Eldfors S, Ellonen P, van Adrichem AJ, Kuusanmäki H, Andersson EI, et al. Somatic STAT3 mutations in large granular lymphocytic leukemia. *N Engl J Med* (2012) 366(20):1905–13. doi: 10.1056/NEJMoa1114885

151. Singh M, Jackson KJL, Wang JJ, Schofield P, Field MA, Koppstein D, et al. Lymphoma driver mutations in the pathogenic evolution of an iconic human autoantibody. *Cell* (2020) 180(5):878–94.e19. doi: 10.1016/j.cell.2020.01.029

152. Masle-Farquhar E, Jackson KJL, Peters TJ, Al-Eryani G, Singh M, Payne KJ, et al. STAT3 gain-of-function mutations connect leukemia with autoimmune disease by pathological NKG2Dhi CD8+ T cell dysregulation and accumulation. *Immunity* (2022) 55(12):2386–404.e8. doi: 10.1016/j.immuni.2022.11.001

153. Fabre A, Marchal S, Barlogis V, Mari B, Barbry P, Rohrlich PS, et al. Clinical aspects of STAT3 gain-of-Function germline mutations: a systematic review. *J Allergy Clin Immunol Pract* (2019) 7(6):1958–69.e9. doi: 10.1016/j.jaip.2019.02.018

154. Jerez A, Clemente MJ, Makishima H, Rajala H, Gómez-Seguí I, Olson T, et al. STAT3 mutations indicate the presence of subclinical T-cell clones in a subset of aplastic anemia and myelodysplastic syndrome patients. *Blood* (2013) 122:2453–9. doi: 10.1182/blood-2013-04-494930

155. Faletti L, Ehl S, Heeg M. Germline STAT3 gain-of-function mutations in primary immunodeficiency: impact on the cellular and clinical phenotype. *BioMed J* (2021) 44(4):412–21. doi: 10.1016/j.bj.2021.03.003

156. Shahmarvand N, Nagy A, Shahryari J, Ohgami RS. Mutations in the signal transducer and activator of transcription family of genes in cancer. *Cancer Sci* (2018) 109(4):926–33. doi: 10.1111/cas.13525

157. Soukup AA, Matson DR, Liu P, Johnson KD, Bresnick EH. Conditionally pathogenic genetic variants of a hematopoietic disease-suppressing enhancer. *Sci Adv* (2021) 7(50):eabk3521. doi: 10.1126/sciadv.abk3521

158. Wagener R, Elitzur S, Brozou T, Borkhardt A. Functional damaging germline variants in ETV6, IKZF1, PAX5 and RUNX1 predisposing to b-cell precursor acute lymphoblastic leukemia. *Eur J Med Genet* (2023) 66(4):104725. doi: 10.1016/j.ejmg.2023.104725

159. Gao X, Johnson KD, Chang YI, Boyer ME, Dewey CN, Zhang J, et al. Gata2 cis-element is required for hematopoietic stem cell generation in the mammalian embryo. *J Exp Med* (2013) 210(13):2833–42. doi: 10.1084/jem.20130733

160. Goyama S, Yamamoto G, Shimabe M, Sato T, Ichikawa M, Ogawa S, et al. Evi-1 is a critical regulator for hematopoietic stem cells and transformed leukemic cells. *Cell Stem Cell* (2008) 3(2):207–20. doi: 10.1016/j.stem.2008.06.002

161. Malinge S, Thiollier C, Chlon TM, Doré LC, Diebold L, Bluteau O, et al. Ikaros inhibits megakaryopoiesis through functional interaction with GATA-1 and NOTCH signaling. *Blood* (2013) 121(13):2440–51. doi: 10.1182/blood-2012-08-450627

162. Crispino JD, Horwitz MS. GATA factor mutations in hematologic disease. *Blood* (2017) 129(15):2103–10. doi: 10.1182/blood-2016-09-687889

163. Saleque S, Cameron S, Orkin SH. The zinc-finger proto-oncogene gfi-1b is essential for development of the erythroid and megakaryocytic lineages. *Genes Dev* (2002) 16(3):301–6. doi: 10.1101/gad.959102

164. Fujiwara Y, Chang AN, Williams AM, Orkin SH. Functional overlap of GATA-1 and GATA-2 in primitive hematopoietic development. *Blood* (2004) 103(2):583–5. doi: 10.1182/blood-2003-08-2870

165. Boutboul D, Kuehn HS, Van de Wyngaert Z, Niemela JE, Callebaut I, Stoddard J, et al. Dominant-negative IKZF1 mutations cause a T, b, and myeloid cell combined immunodeficiency. *J Clin Invest* (2018) 128(7):3071–87. doi: 10.1172/JCI98164

166. Fabozzi F, Strocchio L, Mastronuzzi A, Merli P. GATA2 and marrow failure. *Best Pract Res Clin Haematol* (2021) 34(2):101278. doi: 10.1016/j.beha.2021.101278

167. Bresnick EH, Jung MM, Katsumura KR. Human GATA2 mutations and hematologic disease: how many paths to pathogenesis? *Blood Adv* (2020) 4 (18):4584–92. doi: 10.1182/bloodadvances.2020002953

168. Soukup AA, Bresnick EH. GATA2 +9.5 enhancer: from principles of hematopoiesis to genetic diagnosis in precision medicine. *Curr Opin Hematol* (2020) 27(3):163–71. doi: 10.1097/MOH.0000000000000576

169. Spinner MA, Sanchez LA, Hsu AP, Shaw PA, Zerbe CS, Calvo KR, et al. GATA2 deficiency: a protean disorder of hematopoiesis, lymphatics, and immunity. *Blood* (2014) 123(6):809–21. doi: 10.1182/blood-2013-07-515528

170. Menendez-Gonzalez JB, Vukovic M, Abdelfattah A, Saleh L, Almotiri A, Thomas LA, et al. Gata2 as a crucial regulator of stem cells in adult hematopoiesis and acute myeloid leukemia. *Stem Cell Rep* (2019) 13(2):291–306. doi: 10.1016/j.stemcr.2019.07.005

171. Ling KW, Ottersbach K, van Hamburg JP, Oziemlak A, Tsai FY, Orkin SH, et al. GATA-2 plays two functionally distinct roles during the ontogeny of hematopoietic stem cells. *J Exp Med* (2004) 200(7):871–82. doi: 10.1084/jem.20031556

172. Koyunlar C, Gioacchino E, Vadgama D, de Looper HWJ, Zink J, Ter Borg M, et al. Gata2-regulated Gfi1b expression controls endothelial programming during endothelial-to-hematopoietic transition. *Blood Adv* (2023) 7(10):2082–93. doi: 10.1182/bloodadvances.2022008019

173. Jung M, Cordes S, Zou J, Yu SJ, Guitart X, Hong SG, et al. GATA2 deficiency and human hematopoietic development modeled using induced pluripotent stem cells. *Blood Adv* (2018) 2(23):3553–65. doi: 10.1182/bloodadvances.2018017137

174. Ling T, Zhang K, Yang J, Gurbuxani S, Crispino JD. Gata1s mutant mice display persistent defects in the erythroid lineage. *Blood Adv* (2022). doi: 10.1182/bloodadvances.2022008124

175. Gialesaki S, Mahnken AK, Schmid L, Labuhn M, Bhayadia R, Heckl D, et al. GATA1s exerts developmental stage-specific effects in human hematopoiesis. *Haematologica* (2018) 103(8):e336–40. doi: 10.3324/haematol.2018.191338

176. Matsuo S, Nishinaka-Arai Y, Kazuki Y, Oshimura M, Nakahata T, Niwa A, et al. Pluripotent stem cell model of early hematopoiesis in down syndrome reveals quantitative effects of short-form GATA1 protein on lineage specification. *PloS One* (2021) 16(3):e0247595. doi: 10.1371/journal.pone.0247595

177. Verachi P, Gobbo F, Martelli F, Falchi M, di Virgilio A, Sarli G, et al. Preclinical studies on the use of a p-selectin-blocking monoclonal antibody to halt progression of myelofibrosis in the Gata1low mouse model. *Exp Hematol* (2023) 117:43–61. doi: 10.1016/j.exphem.2022.09.004

178. Arkoun B, Robert E, Boudia F, Mazzi S, Dufour V, Siret A, et al. Stepwise GATA1 and SMC3 mutations alter megakaryocyte differentiation in a down syndrome leukemia model. *J Clin Invest* (2022) 132(14):e156290. doi: 10.1172/JCI156290

179. Camargo R, Sahoo SS, Córdoba JC, Magalhães IQ. Germline GATA1 exon 2 mutation associated with chronic cytopenia and a non-down syndrome transient abnormal myelopoiesis with clonal trisomy 21. *Leukemia* (2022) 36(9):2347–50. doi: 10.1038/s41375-022-01638-6

180. Rodriguez P, Bonte E, Krijgsveld J, Kolodziej KE, Guyot B, Heck AJR, et al. GATA-1 forms distinct activating and repressive complexes in erythroid cells. *EMBO J* (2005) 24(13):2354–66. doi: 10.1038/sj.emboj.7600702

181. El Omari K, Hoosdally SJ, Tuladhar K, Karia D, Vyas P, Patient R, et al. Structure of the leukemia oncogene LMO2: implications for the assembly of a hematopoietic transcription factor complex. *Blood* (2011) 117(7):2146–56. doi: 10.1182/blood-2010-07-293357

182. Meier N, Krpic S, Rodriguez P, Strouboulis J, Monti M, Krijgsveld J, et al. Novel binding partners of Ldb1 are required for haematopoietic development. *Development* (2006) 133(24):4913–23. doi: 10.1242/dev.02656

183. Snow JW, Kim J, Currie CR, Xu J, Orkin SH. Sumoylation regulates interaction of FOG1 with c-terminal-binding protein (CTBP). *J Biol Chem* (2010) 285(36):28064–75. doi: 10.1074/jbc.M109.096909

184. Hamlett I, Draper J, Strouboulis J, Iborra F, Porcher C, Vyas P. Characterization of megakaryocyte GATA1-interacting proteins: the corepressor ETO2 and GATA1 interact to regulate terminal megakaryocyte maturation. *Blood* (2008) 112(7):2738–49. doi: 10.1182/blood-2008-03-146605

185. Lowry JA, Mackay JP. GATA-1: one protein, many partners. *Int J Biochem Cell Biol* (2006) 38(1):6–11. doi: 10.1016/j.biocel.2005.06.017

186. Hariom SK, Nelson EJR. Effects of short-term hypergravity on hematopoiesis and vasculogenesis in embryonic zebrafish. *Life Sci Space Res* (2022) 34:21–9. doi: 10.1016/j.lssr.2022.05.005

187. Yu L, Takai J, Otsuki A, Katsuoka F, Suzuki M, Katayama S, et al. Derepression of the DNA methylation machinery of the Gata1 gene triggers the differentiation cue for erythropoiesis. *Mol Cell Biol* (2017) 37(8):e00592-16. doi: 10.1128/MCB.00592-16

188. Muskens IS, Li S, Jackson T, Elliot N, Hansen HM, Myint SS, et al. The genome-wide impact of trisomy 21 on DNA methylation and its implications for hematopoiesis. *Nat Commun* (2021) 12(1):821. doi: 10.1038/s41467-021-21064-z

189. Medina S, Zhang H, Santos-Medina LV, Wan G, Bolt AM, Zhou X, et al. Arsenic impairs the lineage commitment of hematopoietic progenitor cells through the attenuation of GATA-2 DNA binding activity. *Toxicol Appl Pharmacol* (2022) 452:116193. doi: 10.1016/j.taap.2022.116193

190. Yuasa H, Oike Y, Iwama A, Nishikata I, Sugiyama D, Perkins A, et al. Oncogenic transcription factor Evi1 regulates hematopoietic stem cell proliferation through GATA-2 expression. *EMBO J* (2005) 24(11):1976–87. doi: 10.1038/sj.emboj.7600679

191. Yuan X, Wang X, Bi K, Jiang G. The role of EVI-1 in normal hematopoiesis and myeloid malignancies (Review). *Int J Oncol* (2015) 47(6):2028–36. doi: 10.3892/ijo.2015.3207

192. Voit RA, Tao L, Yu F, Cato LD, Cohen B, Fleming TJ, et al. A genetic disorder reveals a hematopoietic stem cell regulatory network co-opted in leukemia. *Nat Immunol* (2023) 24(1):69–83. doi: 10.1038/s41590-022-01370-4

193. Wilson M, Tsakraklides V, Tran M, Xiao YY, Zhang Y, Perkins AS. EVI1 interferes with myeloid maturation *via* transcriptional repression of cebpa, *via* binding to two far downstream regulatory elements. *J Biol Chem* (2016) 291(26):13591–607. doi: 10.1074/jbc.M115.708156

194. Birdwell C, Fiskus W, Kadia TM, DiNardo CD, Mill CP, Bhalla KN. EVI1 dysregulation: impact on biology and therapy of myeloid malignancies. *Blood Cancer J* (2021) 11(3):64. doi: 10.1038/s41408-021-00457-9

195. Paredes R, Doleschall N, Connors K, Geary B, Meyer S. EVI1 protein interaction dynamics: targetable for therapeutic intervention? *Exp Hematol* (2022) 107:1–8. doi: 10.1016/j.exphem.2021.12.398

196. Buonamici S, Li D, Chi Y, Zhao R, Wang X, Brace L, et al. EVI1 induces myelodysplastic syndrome in mice. *J Clin Invest* (2004) 114(5):713–9. doi: 10.1172/JCI21716

197. Nabil R, Abdellateif MS, Gamal H, Hassan NM, Badawy RH, Ghareeb M, et al. Clinical significance of EVI-1 gene expression and aberrations in patient with de-novo acute myeloid and acute lymphoid leukemia. *Leuk Res* (2023) 126:107019. doi: 10.1016/j.leukres.2023.107019

198. Vinothkumar K, Chanda S, Singh VK, Biswas S, Mohapatra S, Biswas G, et al. EVI1 upregulates PTGS1 (COX1) and decreases the action of tyrosine kinase inhibitors (TKIs) in chronic myeloid leukemia cells. *Int J Hematol* (2023) 117(1):110–20. doi: 10.1007/s12185-022-03465-y

199. Yamashita M, Kuehn HS, Okuyama K, Okada S, Inoue Y, Mitsuiki N, et al. A variant in human AIOLOS impairs adaptive immunity by interfering with IKAROS. *Nat Immunol* (2021) 22(7):893–903. doi: 10.1038/s41590-021-00951-z

200. Kwiatkowski BA, Bastian LS, Bauer TR, Tsai S, Zielinska-Kwiatkowska AG, Hickstein DD. The ets family member tel binds to the fli-1 oncoprotein and inhibits its transcriptional activity*. *J Biol Chem* (1998) 273(28):17525–30. doi: 10.1074/jbc.273.28.17525

201. Lu DY, Ellegast JM, Ross KN, Malone CF, Lin S, Mabe NW, et al. The ETS transcription factor ETV6 constrains the transcriptional activity of EWS-FLI to promote Ewing sarcoma. *Nat Cell Biol* (2023) 25(2):285–97. doi: 10.1038/s41556-022-01059-8

202. Sasca D, Yun H, Giotopoulos G, Szybinski J, Evan T, Wilson NK, et al. Cohesin-dependent regulation of gene expression during differentiation is lost in cohesin-mutated myeloid malignancies. *Blood* (2019) 134(24):2195–208. doi: 10.1182/blood.2019001553

203. Wang LC, Swat W, Fujiwara Y, Davidson L, Visvader J, Kuo F, et al. The TEL/ETV6 gene is required specifically for hematopoiesis in the bone marrow. *Genes Dev* (1998) 12(15):2392–402. doi: 10.1101/gad.12.15.2392

204. Hock H, Meade E, Medeiros S, Schindler JW, Valk PJM, Fujiwara Y, et al. Tel/Etv6 is an essential and selective regulator of adult hematopoietic stem cell survival. *Genes Dev* (2004) 18(19):2336–41. doi: 10.1101/gad.1239604

205. Bernardin F, Yang Y, Cleaves R, Zahurak M, Cheng L, Civin CI, et al. TEL-AML1, expressed from t(12;21) in human acute lymphocytic leukemia, induces acute leukemia in mice. *Cancer Res* (2002) 62(14):3904–8.

206. Li M, Jones L, Gaillard C, Binnewies M, Ochoa R, Garcia E, et al. Initially disadvantaged, TEL-AML1 cells expand and initiate leukemia in response to irradiation and cooperating mutations. *Leukemia* (2013) 27(7):1570–3. doi: 10.1038/leu.2013.15

207. Schindler JW, Van Buren D, Foudi A, Krejci O, Qin J, Orkin SH, et al. TEL-AML1 corrupts hematopoietic stem cells to persist in the bone marrow and initiate leukemia. *Cell Stem Cell* (2009) 5(1):43–53. doi: 10.1016/j.stem.2009.04.019

208. Zhou C, Uluisik R, Rowley JW, David C, Jones CL, Scharer CD, et al. Germline ETV6 mutation promotes inflammation and disrupts lymphoid development of early hematopoietic progenitors. *Exp Hematol* (2022) 112-113:24–34. doi: 10.1016/j.exphem.2022.06.002

209. Fisher MH, Kirkpatrick GD, Stevens B, Jones C, Callaghan M, Rajpurkar M, et al. ETV6 germline mutations cause HDAC3/NCOR2 mislocalization and upregulation of interferon response genes. *JCI Insight* (2020) 5(18):e140332. doi: 10.1172/jci.insight.140332

210. Sheng D, Chen B, Wang C, Xiao X, Hu A, Liu W, et al. FLI1 accelerates leukemogenesis through transcriptional regulation of pyruvate kinase-L/R and other glycolytic genes. *Med Oncol* (2022) 40(2):69. doi: 10.1007/s12032-022-01867-w

211. Tijssen MR, Cvejic A, Joshi A, Hannah RL, Ferreira R, Forrai A, et al. Genome-wide analysis of simultaneous GATA1/2, RUNX1, FLI1, and SCL binding in megakaryocytes identifies hematopoietic regulators. *Dev Cell* (2011) 20(5):597–609. doi: 10.1016/j.devcel.2011.04.008

212. Chen Z, Arai E, Khan O, Zhang Z, Ngiow SF, He Y, et al. In vivo CD8+ T cell CRISPR screening reveals control by Fli1 in infection and cancer. *Cell* (2021) 184 (5):1262–80.e22. doi: 10.1016/j.cell.2021.02.019

213. Moussa O, LaRue AC, Abangan RSJr, Williams CR, Zhang XK, Masuya M, et al. Thrombocytopenia in mice lacking the carboxy-terminal regulatory domain of the ets transcription factor Fli1. *Mol Cell Biol* (2010) 30(21):5194–206. doi: 10.1128/MCB.01112-09

214. Jackers P, Szalai G, Moussa O, Watson DK. Ets-dependent regulation of target gene expression during megakaryopoiesis. *J Biol Chem* (2004) 279(50):52183–90. doi: 10.1074/jbc.M407489200

215. Vo KK, Jarocha DJ, Lyde RB, Hayes V, Thom CS, Sullivan SK, et al. FLI1 level during megakaryopoiesis affects thrombopoiesis and platelet biology. *Blood* (2017) 129 (26):3486–94. doi: 10.1182/blood-2017-02-770958

216. Singhal D, Hahn CN, Feurstein S, Wee LYA, Moma L, Kutyna MM, et al. Targeted gene panels identify a high frequency of pathogenic germline variants in patients diagnosed with a hematological malignancy and at least one other independent cancer. *Leukemia* (2021) 35(11):3245–56. doi: 10.1038/s41375-021-01246-w

217. Feurstein S, Trottier AM, Estrada-Merly N, Pozsgai M, McNeely K, Drazer MW, et al. Germ line predisposition variants occur in myelodysplastic syndrome patients of all ages. *Blood* (2022) 140(24):2533–48. doi: 10.1182/blood.2022015790

218. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American college of medical genetics and genomics and the association for molecular pathology. *Genet Med* (2015) 17(5):405–24. doi: 10.1038/gim.2015.30

219. Makishima H, Saiki R, Nannya Y, Korotev S, Gurnari C, Takeda J, et al. Germ line DDX41 mutations define a unique subtype of myeloid neoplasms. *Blood* (2023) 141 (5):534–49. doi: 10.1182/blood.2022018221

220. Cheah JJC, Hahn CN, Hiwase DK, Scott HS, Brown AL. Myeloid neoplasms with germline DDX41 mutation. *Int J Hematol* (2017) 106(2):163–74. doi: 10.1007/s12185-017-2260-y

221. Bullinger L, Döhner K, Döhner H. Genomics of acute myeloid leukemia diagnosis and pathways. *J Clin Oncol* (2017) 35(9):934–46. doi: 10.1200/JCO.2016.71.2208

222. Xu F, Wu LY, Guo J, He Q, Zhang Z, Li X. Somatic mutations of activating signalling, transcription factor, and tumour suppressor are a precondition for leukaemia transformation in myelodysplastic syndromes. *J Cell Mol Med* (2022) 26 (23):5901–16. doi: 10.1111/jcmm.17613

223. Yokota A, Huo L, Lan F, Wu J, Huang G. The clinical, molecular, and mechanistic basis of RUNX1 mutations identified in hematological malignancies. *Mol Cells* (2020) 43(2):145–52. doi: 10.14348/molcells.2019.0252

224. Gonzales F, Barthélémy A, Peyrouze P, Fenwarth L, Preudhomme C, Duployez N, et al. Targeting RUNX1 in acute myeloid leukemia: preclinical innovations and therapeutic implications. *Expert Opin Ther Targets* (2021) 25(4):299–309. doi: 10.1080/14728222.2021.1915991

225. Björk J, Johansson B, Broberg K, Albin M. Smoking as a risk factor for myelodysplastic syndromes and acute myeloid leukemia and its relation to cytogenetic findings: a case-control study. *Leuk Res* (2009) 33(6):788–91. doi: 10.1016/j.leukres.2008.10.009

226. Maleki Behzad M, Abbasi M, Oliaei I, Ghorbani Gholiabad S, Rafieemehr H. Effects of lifestyle and environmental factors on the risk of acute myeloid leukemia: result of a hospital-based case-control study. *J Res Health Sci* (2021) 21(3):e00525. doi: 10.34172/jrhs.2021.58

227. Weiss JR, Baker JA, Baer MR, Menezes RJ, Nowell S, Moysich KB. Opposing effects of aspirin and acetaminophen use on risk of adult acute leukemia. *Leuk Res* (2006) 30(2):164–9. doi: 10.1016/j.leukres.2005.06.023

228. Wong O, Harris F, Yiying W, Hua F. A hospital-based case-control study of acute myeloid leukemia in shanghai: analysis of personal characteristics, lifestyle and environmental risk factors by subtypes of the WHO classification. *Regul Toxicol Pharmacol* (2009) 55(3):340–52. doi: 10.1016/j.yrtph.2009.08.007

229. Brown AL, Arts P, Carmichael CL, Babic M, Dobbins J, Chong CE, et al. RUNX1-mutated families show phenotype heterogeneity and a somatic mutation profile unique to germline predisposed AML. *Blood Adv* (2020) 4(6):1131–44. doi: 10.1182/bloodadvances.2019000901

230. King KY, Huang Y, Nakada D, Goodell MA. Environmental influences on clonal hematopoiesis. *Exp Hematol* (2020) 83:66–73. doi: 10.1016/j.exphem.2019.12.005

231. Jithesh PV, Abuhaliqa M, Syed N, Ahmed I, El Anbari M, Bastaki K, et al. A population study of clinically actionable genetic variation affecting drug response from the middle East. *NPJ Genom Med* (2022) 7(1):10. doi: 10.1038/s41525-022-00281-5

232. Schärfe CPI, Tremmel R, Schwab M, Kohlbacher O, Marks DS. Genetic variation in human drug-related genes. *Genome Med* (2017) 9(1):117. doi: 10.1186/s13073-017-0502-5

233. Ng AP, Coughlan HD, Hediyeh-Zadeh S, Behrens K, Johanson TM, Low MSY, et al. An erg-driven transcriptional program controls b cell lymphopoiesis. *Nat Commun* (2020) 11(1):3013. doi: 10.1038/s41467-020-16828-y

234. Zhao X, Bartholdy B, Yamamoto Y, Evans EK, Alberich-Jordà M, Staber PB, et al. PU.1-c-Jun interaction is crucial for PU.1 function in myeloid development. *Commun Biol* (2022) 5(1):961. doi: 10.1038/s42003-022-03888-7

235. Chambers C, Cermakova K, Chan YS, Kurtz K, Wohlan K, Lewis AH, et al. SWI/SNF blockade disrupts PU.1-directed enhancer programs in normal hematopoietic cells and acute myeloid leukemia. *Cancer Res* (2023) 83(7):983–96. doi: 10.1158/0008-5472.CAN-22-2129

236. Taoudi S, Bee T, Hilton A, Knezevic K, Scott J, Willson TA, et al. ERG dependence distinguishes developmental control of hematopoietic stem cell maintenance from hematopoietic specification. *Genes Dev* (2011) 25(3):251–62. doi: 10.1101/gad.2009211

237. Knudsen KJ, Rehn M, Hasemann MS, Rapin N, Bagger FO, Ohlsson E, et al. ERG promotes the maintenance of hematopoietic stem cells by restricting their differentiation. *Genes Dev* (2015) 29(18):1915–29. doi: 10.1101/gad.268409.115

238. Orkin SH. Hematopoiesis: how does it happen? *Curr Opin Cell Biol* (1995) 7 (6):870–7. doi: 10.1016/0955-0674(95)80072-7

239. Kastner P, Chan S. PU.1: a crucial and versatile player in hematopoiesis and leukemia. *Int J Biochem Cell Biol* (2008) 40(1):22–7. doi: 10.1016/j.biocel.2007.01.026

240. Xie Y, Koch ML, Zhang X, Hamblen MJ, Godinho FJ, Fujiwara Y, et al. Reduced erg dosage impairs survival of hematopoietic stem and progenitor cells. *Stem Cells* (2017) 35(7):1773–85. doi: 10.1002/stem.2627

241. Iwasaki H, Somoza C, Shigematsu H, Duprez EA, Iwasaki-Arai J, Mizuno SI, et al. Distinctive and indispensable roles of PU.1 in maintenance of hematopoietic stem cells and their differentiation. *Blood* (2005) 106(5):1590–600. doi: 10.1182/blood-2005-03-0860

242. Zaidan N, Ottersbach K. The multi-faceted role of Gata3 in developmental haematopoiesis. *Open Biol* (2018) 8(11):180152. doi: 10.1098/rsob.180152

243. Clappier E, Grardel N, Bakkus M, Rapion J, De Moerloose B, Kastner P, et al. IKZF1 deletion is an independent prognostic marker in childhood b-cell precursor acute lymphoblastic leukemia, and distinguishes patients benefiting from pulses during

maintenance therapy: results of the EORTC children's leukemia group study 58951. *Leukemia* (2015) 29(11):2154–61. doi: 10.1038/leu.2015.134

244. Zhang J, McCastlain K, Yoshihara H, Xu B, Chang Y, Churchman ML, et al. Deregulation of DUX4 and ERG in acute lymphoblastic leukemia. *Nat Genet* (2016) 48 (12):1481–9. doi: 10.1038/ng.3691

245. Peghaire C, Dufton NP, Lang M, Salles-Crawley II, Ahnström J, Kalna V, et al. The transcription factor ERG regulates a low shear stress-induced anti-thrombotic pathway in the microvasculature. *Nat Commun* (2019) 10(1):5014. doi: 10.1038/s41467-019-12897-w

246. Kalna V, Yang Y, Peghaire CR, Frudd K, Hannah R, Shah AV, et al. The transcription factor ERG regulates super-enhancers associated with an endothelial-specific gene expression program. *Circ Res* (2019) 124(9):1337–49. doi: 10.1161/CIRCRESAHA.118.313788

247. Greene D, Pirri D, Frudd K, Sackey E, Al-Owain M, Giese APJ, et al. Genetic association analysis of 269 rare diseases reveals novel aetiologies. *bioRxiv* (2022), 2022–06. doi: 10.1101/2022.06.10.22276270

248. Tomlins SA, Palanisamy N, Siddiqui J, Chinnaiyan AM, Kunju LP. Antibody-based detection of ERG rearrangements in prostate core biopsies, including diagnostically challenging cases: ERG staining in prostate core biopsies. *Arch Pathol Lab Med* (2012) 136(8):935–46. doi: 10.5858/arpa.2011-0424-OA

249. Kissick HT, On ST, Dunn LK, Sanda MG, Asara JM, Pellegrini KL, et al. The transcription factor ERG increases expression of neurotransmitter receptors on prostate cancer cells. *BMC Cancer* (2015) 15:604. doi: 10.1186/s12885-015-1612-3

250. García-Perdomo HA, Chaves MJ, Osorio JC, Sanchez A. Association between TMPRSS2:ERG fusion gene and the prostate cancer: systematic review and meta-analysis. *Cent Eur J Urol* (2018) 71(4):410–9. doi: 10.5173/ceju.2018.1752

251. Chen S, Deniz K, Sung YS, Zhang L, Dry S, Antonescu CR. Ewing Sarcoma with ERG gene rearrangements: a molecular study focusing on the prevalence of FUS-ERG and common pitfalls in detecting EWSR1-ERG fusions by FISH. *Genes Chromosomes Cancer* (2016) 55(4):340–9. doi: 10.1002/gcc.22336

252. Kim J, Park TS, Song J, Lee KA, Hong DJ, Min YH, et al. Detection of FUS-ERG chimeric transcript in two cases of acute myeloid leukemia with t(16;21)(p11.2;q22) with unusual characteristics. *Cancer Genet Cytogenet* (2009) 194(2):111–8. doi: 10.1016/j.cancergencyto.2009.06.010

253. Schleicher WE, Pietras EM. Reduced PU.1 expression collaborates with Tet2 loss to trigger myeloid leukemogenesis. *Blood Cancer Discovery* (2022) 3(5):378–81. doi: 10.1158/2643-3230.BCD-22-0100

254. Aivalioti MM, Bartholdy BA, Pradhan K, Bhagat TD, Zintiridou A, Jeong JJ, et al. PU.1-dependent enhancer inhibition separates Tet2-deficient hematopoiesis from malignant transformation. *Blood Cancer Discovery* (2022) 3(5):444–67. doi: 10.1158/2643-3230.BCD-21-0226

255. Gregoricchio S, Polit L, Esposito M, Berthelet J, Delestré L, Evanno E, et al. HDAC1 and PRC2 mediate combinatorial control in SPI1/PU.1-dependent gene repression in murine erythroleukaemia. *Nucleic Acids Res* (2022) 50(14):7938–58. doi: 10.1093/nar/gkac613

256. Burda P, Laslo P, Stopka T. The role of PU.1 and GATA-1 transcription factors during normal and leukemogenic hematopoiesis. *Leukemia* (2010) 24(7):1249–57. doi: 10.1038/leu.2010.104

257. Bašová P, Paszeková H, Minařík L, Dluhošová M, Burda P, Stopka T. Combined approach to leukemic differentiation using transcription factor PU.1-enhancing agents. *Int J Mol Sci* (2022) 23(12):6729. doi: 10.3390/ijms23126729

258. Roos-Weil D, Decaudin C, Armand M, Della-Valle V, Diop MK, Ghamlouch H, et al. A recurrent activating missense mutation in waldenström macroglobulinemia affects the DNA binding of the ETS transcription factor SPI1 and enhances proliferation. *Cancer Discovery* (2019) 9(6):796–811. doi: 10.1158/2159-8290.CD-18-0873

259. Le Coz C, Nguyen DN, Su C, Nolan BE, Albrecht AV, Xhani S, et al. Constrained chromatin accessibility in PU.1-mutated agammaglobulinemia patients. *J Exp Med* (2021) 218(7):e20201750. doi: 10.1084/jem.20201750

260. Geskin LJ, Viragova S, Stolz DB, Fuschiotti P. Interleukin-13 is overexpressed in cutaneous T-cell lymphoma cells and regulates their proliferation. *Blood* (2015) 125 (18):2798–805. doi: 10.1182/blood-2014-07-590398

261. Fransecky L, Neumann M, Heesch S, Schlee C, Ortiz-Tanchez J, Heller S, et al. Silencing of GATA3 defines a novel stem cell-like subgroup of ETP-ALL. *J Hematol Oncol* (2016) 9(1):95. doi: 10.1186/s13045-016-0324-8

262. Murga-Zamalloa C, Wilcox RA. GATA-3 in T-cell lymphoproliferative disorders. *IUBMB Life* (2020) 72(1):170–7. doi: 10.1002/iub.2130

263. Dai YT, Zhang F, Fang H, Li JF, Lu G, Jiang L, et al. Transcriptome-wide subtyping of pediatric and adult T cell acute lymphoblastic leukemia in an international study of 707 cases. *Proc Natl Acad Sci USA* (2022) 119(15):e2120787119. doi: 10.1073/pnas.2120787119

264. Reiswich V, Schmidt CE, Lennartz M, Höflmayer D, Hube-Magg C, Weidemann S, et al. GATA3 expression in human tumors: a tissue microarray study on 16,557 tumors. *Pathobiology* (2023) 17:1–14. doi: 10.1159/000527382

265. Ibrahim IH, Abdel-Aziz HG, Hassan FE, El-Sameea HS. Role of GATA3 exon 6 germline mutations in breast cancer progression in Egyptian female patients. *Exp Biol Med* (2021) 246(1):40–7. doi: 10.1177/1535370220958610

266. Ziegler A, Colin E, Goudenège D, Bonneau D. A snapshot of some pLI score pitfalls. *Hum Mutat* (2019) 40(7):839–41. doi: 10.1002/humu.23763

267. Spinner MA, Ker JP, Stoudenmire CJ, Fadare O, Mace EM, Orange JS, et al. GATA2 deficiency underlying severe blastomycosis and fatal herpes simplex virus-associated hemophagocytic lymphohistiocytosis. *J Allergy Clin Immunol* (2016) 137 (2):638–40. doi: 10.1016/j.jaci.2015.07.043

268. Wu Z, Gao S, Diamond C, Kajigaya S, Chen J, Shi R, et al. Sequencing of RNA in single cells reveals a distinct transcriptome signature of hematopoiesis in GATA2 deficiency. *Blood Adv* (2020) 4(12):2656–70. doi: 10.1182/bloodadvances.2019001352

269. Liu TH, Tang YJ, Huang Y, Wang L, Guo XL, Mi JQ, et al. Expression of the fetal hematopoiesis regulator FEV indicates leukemias of prenatal origin. *Leukemia* (2017) 31(5):1079–86. doi: 10.1038/leu.2016.313

270. Jaiswal S, Fontanillas P, Flannick J, Manning A, Grauman PV, Mar BG, et al. Age-related clonal hematopoiesis associated with adverse outcomes. *N Engl J Med* (2014) 371(26):2488–98. doi: 10.1056/NEJMoa1408617

271. West RR, Calvo KR, Embree LJ, Wang W, Tuschong LM, Bauer TR, et al. ASXL1 and STAG2 are common mutations in GATA2 deficiency patients with bone marrow disease and myelodysplastic syndrome. *Blood Adv* (2022) 6(3):793–807. doi: 10.1182/bloodadvances.2021005065

272. Nachun D, Lu AT, Bick AG, Natarajan P, Weinstock J, Szeto MD, et al. Clonal hematopoiesis associated with epigenetic aging and clinical outcomes. *Aging Cell* (2021) 20(6):e13366. doi: 10.1111/acel.13366

273. Beerman I, Bock C, Garrison BS, Smith ZD, Gu H, Meissner A, et al. Proliferation-dependent alterations of the DNA methylation landscape underlie hematopoietic stem cell aging. *Cell Stem Cell* (2013) 12(4):413–25. doi: 10.1016/j.stem.2013.01.017

274. Bell CG, Lowe R, Adams PD, Baccarelli AA, Beck S, Bell JT, et al. DNA Methylation aging clocks: challenges and recommendations. *Genome Biol* (2019) 20 (1):249. doi: 10.1186/s13059-019-1824-y

275. Abelson S, Collord G, Ng SWK, Weissbrod O, Mendelson Cohen N, Niemeyer E, et al. Prediction of acute myeloid leukaemia risk in healthy individuals. *Nature* (2018) 559(7714):400–4. doi: 10.1038/s41586-018-0317-6

276. Van Esch H, Groenen P, Nesbit MA, Schuffenhauer S, Lichtner P, Vanderlinden G, et al. GATA3 haplo-insufficiency causes human HDR syndrome. *Nature* (2000) 406(6794):419–22. doi: 10.1038/35019088

277. Stanley KE, Giordano J, Thorsten V, Buchovecky C, Thomas A, Ganapathi M, et al. Causal genetic variants in stillbirth. *N Engl J Med* (2020) 383(12):1107–16. doi: 10.1056/NEJMoa1908753

278. Byrne AB, Arts P, Ha TT, Kassahn KS, Pais LS, O'Donnell-Luria A, et al. Genomic autopsy to identify underlying causes of pregnancy loss and perinatal death. *Nat Med* (2023) 29(1):180–9. doi: 10.1038/s41591-022-02142-1

279. Vyas H, Alcheikh A, Lowe G, Stevenson WS, Morgan NV, Rabbolini DJ. Prevalence and natural history of variants in the ANKRD26 gene: a short review and update of reported cases. *Platelets* (2022) 33(8):1107–12. doi: 10.1080/09537104.2022.2071853