LETTER TO THE EDITOR

Epigenome profiling reveals aberrant DNA methylation signature in GATA2 deficiency

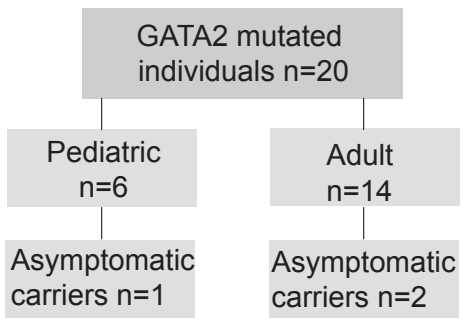
Oskar Marin-Bejar,[1] Damia Romero-Moya,[1] Javier Rodriguez-Ubreva,[2] Maximiliano Distefano,[3] Francesca Lessi,[4] Paolo Aretini,[4] Alessandro Liquori,[5,6] Julio Castaño,[7] Emilia Kozyra,[8] Lili Kotmayer,[9] Clara Bueno,[10] José Cervera,[5,6,11] José Carlos Rodriguez-Gallego,[12,13,14] Josep F. Nomdedeu,[15] Laura Murillo- Sanjuán,[16] Cristina Díaz de Heredia,[16] Antonio Pérez-Martinez,[17,18,19] Félix López-Cadenas,[20,21] Carolina Martínez-Laperche,[22] Nieves Dorado-Herrero,[22] Francisco M. Marco,[23] Felipe Prósper,[6,24,25] Pablo Menendez,[6,10,26,27] David Valcárcel,[28] Esteban Ballestar,[2,29] Csaba Bödör,[9] Anna Bigas,[6,30,31] Albert Catalá,[3,32] Marcin W. Wlodarski[33] and Alessandra Giorgetti[1,4,34]

[1]Regenerative Medicine Program, Bellvitge Institute for Biomedical Research (IDIBELL) and Program for Clinical Translation of Regenerative Medicine in Catalonia (P-CMRC), Barcelona, Spain; [2]Epigenetics and Immune Disease Group, Josep Carreras Research Institute (IJC), Barcelona, Spain; [3]Department of Hematology and Oncology, Institut de Recerca Sant Joan de Déu, Hospital Sant Joan de Déu, Barcelona, Spain; [4]Fondazione Pisana Per la Scienza ONLUS (FPS), San Giuliano Terme, Italy; [5]Hematology Research Group, Instituto de Investigación Sanitaria La Fe, Valencia, Spain; [6]Centro de Investigación Biomédica en Red de Oncología (CIBERONC), Instituto de Salud Carlos III, Madrid, Spain; [7]Advanced and Cell Therapy Services, Banc de Sang i Teixits, Barcelona, Spain; [8]Division of Pediatric Hematology and Oncology, Department of Pediatrics and Adolescent Medicine, Medical Center, Faculties of Medicine and Biology, University of Freiburg, Freiburg, Germany; [9]HCEMM-SE Molecular Oncohematology Research Group, Department of Pathology and Experimental Cancer Research, Semmelweis University, Budapest, Hungary; [10]Josep Carreras Leukemia Research Institute, Department of Biomedicine, School of Medicine, University of Barcelona, Barcelona, Spain; [11]Genetics Unit, Hospital Universitario y Politécnico La Fe, Valencia, Spain; [12]Department of Immunology, University Hospital of Gran Canaria Dr. Negrin, Canarian Health System, Las Palmas de Gran Canaria, Spain; [13]Department of Clinical Sciences, University Fernando Pessoa Canarias, Las Palmas de Gran Canaria, Spain; [14]Department of Medical and Surgical Sciences, School of Medicine, University of Las Palmas de Gran Canaria, Las Palmas de Gran Canaria, Spain; [15]Servei d´Hematologia Hospital de la Santa Creu i Sant Pau, Universitat Autònoma de Barcelona, IIB Sant Pau/Josep Carreras Leukemia Research Institute (IJC), Barcelona, Spain; [16]Pediatric Hematology and Oncology Division, Hospital Universitari Vall d'Hebron, Vall d'Hebron Institut de Recerca, Barcelona, Spain; [17]Pediatric Department, Universidad Autonoma de Madrid, Madrid, Spain; [18]Hospital La Paz Institute for Health Research, Madrid, Spain; [19]Pediatric Hemato-Oncology Department, University Hospital La Paz, Madrid, Spain; [20]Servicio de Hematología Hospital Clínico Universitario de Salamanca, Salamanca, Spain; [21]Instituto Biosanitario de Salamanca (IBSAL), Salamanca, Spain; [22]Servicio de Hematología, Hospital General Universitario Gregorio Marañón, Instituto de Investigación Sanitaria Gregorio Marañón, Madrid, Spain; [23]Immunology Department, Dr. Balmis General University Hospital, Institute for Health and Biomedical Research (ISABIAL), Alicante, Spain; [24]Area de Hemato-Oncología, CIMA Universidad de Navarra, Instituto de Investigación Sanitaria de Navarra (IDISNA), Pamplona, Spain; [25]Servicio de Hematologia, CCUN, Clínica Universidad de Navarra, Universidad de Navarra, Pamplona, Spain; [26]Red Española de Terapias Avanzadas (TERAV) - Instituto de Salud Carlos III, Madrid, Spain; [27]Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain; [28]Hematology Department, Vall d´Hebron University Hospital; Experimental Hematology, Vall d'Hebron Institute of Oncology (VHIO), Department of Medicine, Universitat Autònoma de Barcelona, Barcelona, Spain; [29]Epigenetics in Inflammatory and Metabolic Diseases Laboratory, Health Science Center (HSC), East China Normal University (ECNU), Shanghai, China; [30]Programa de Investigación en Cáncer, IMIM (Hospital del Mar Medical Research Institute), Barcelona, Spain; [31]Josep Carreras Leukemia Research Institute (IJC), Barcelona, Spain; [32]Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Instituto de Salud Carlos III, Madrid, Spain; [33]Department of Hematology, St. Jude Children's Research Hospital, Memphis, TN, USA and [34]Department of Pathology and Experimental Therapeutics, Faculty of Medicine and Health Sciences, Barcelona University, Barcelona, Spain

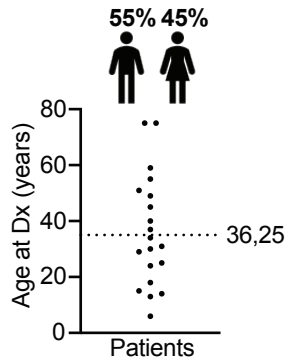
Correspondence:
A. GIORGETTI - agiorgetti@idibell.cat
O. MARIN-BEJAR - omarin@idibell.cat

https://doi.org/10.3324/haematol.2022.282305

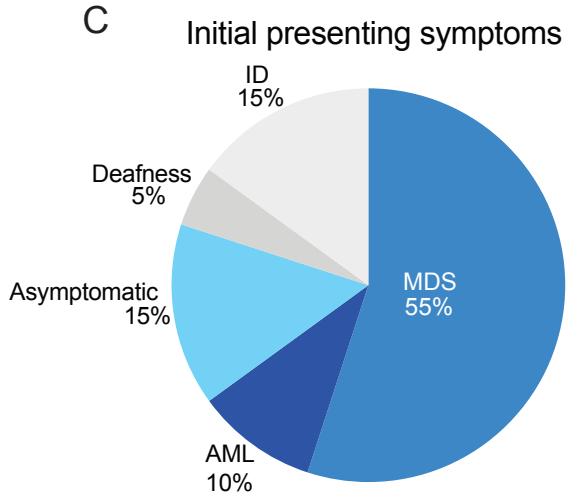
A


GATA2 mutated individuals n=20
Pediatric n=6
Adult n=14
Asymptomatic carriers n=1
Asymptomatic carriers n=2

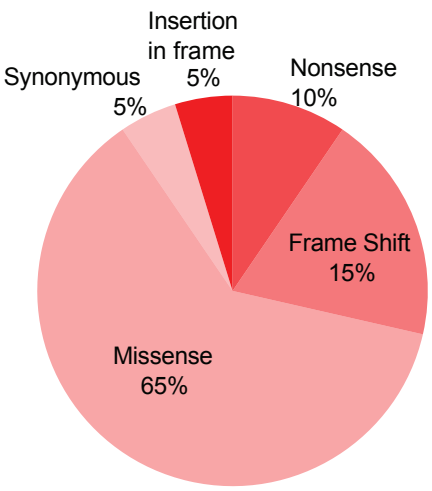
B


55% 45%
Age at Dx (years)
80
60
40
20
0
36,25
Patients

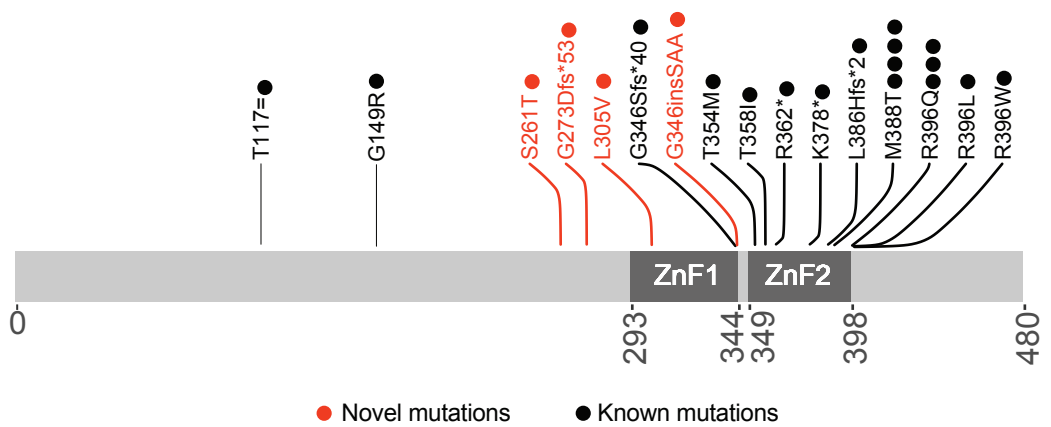
C


Initial presenting symptoms
ID 15%
Deafness 5%
Asymptomatic 15%
MDS 55%
AML 10%

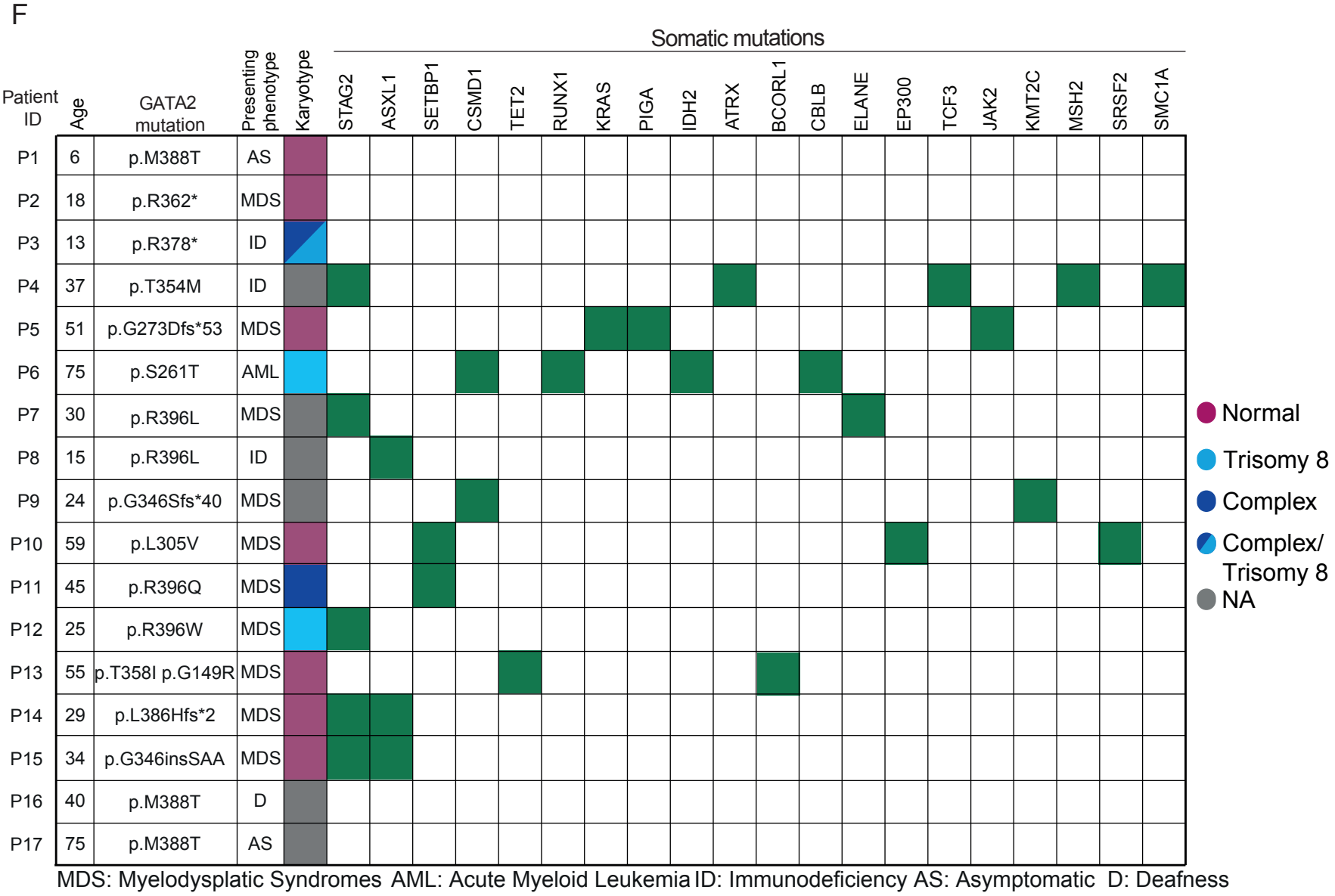
D


Insertion in frame 5%
Synonymous 5%
Nonsense 10%
Frame Shift 15%
Missense 65%

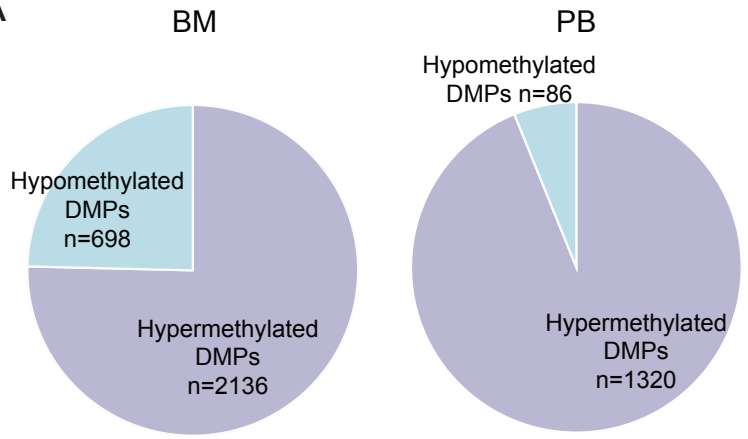
E


T117=
G149R
S261T
G273Dfs*53
L305V
G346Sfs*40
G346insSAA
T354M
T358I
R362*
K378*
L386Hfs*2
M388T
R396Q
R396L
R396W
ZnF1
ZnF2
0
293
344
349
398
480
Novel mutations
Known mutations

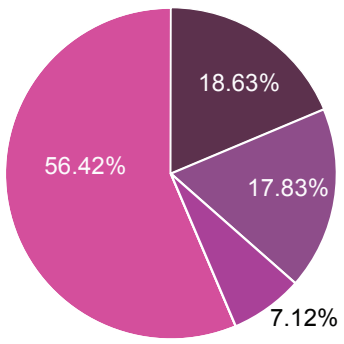
F

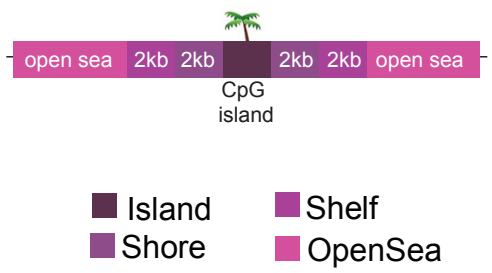

Somatic mutations
Patient ID
Age
GATA2 mutation
Presenting phenotype
Karyotype
STAG2
ASXL1
SETBP1
CSMD1
TET2
RUNX1
KRAS
PIGA
IDH2
ATRX
BCORL1
CBLB
ELANE
EP300
TCF3
JAK2
KMT2C
MSH2
SRSF2
SMC1A
P1 6 p.M388T AS
P2 18 p.R362* MDS
P3 13 p.R378* ID
P4 37 p.T354M ID
P5 51 p.G273Dfs*53 MDS
P6 75 p.S261T AML
P7 30 p.R396L MDS
P8 15 p.R396L ID
P9 24 p.G346Sfs*40 MDS
P10 59 p.L305V MDS
P11 45 p.R396Q MDS
P12 25 p.R396W MDS
P13 55 p.T358I p.G149R MDS
P14 29 p.L386Hfs*2 MDS
P15 34 p.G346insSAA MDS
P16 40 p.M388T D
P17 75 p.M388T AS
Normal
Trisomy 8
Complex
Complex/ Trisomy 8
NA
MDS: Myelodysplatic Syndromes AML: Acute Myeloid Leukemia ID: Immunodeficiency AS: Asymptomatic D: Deafness

A

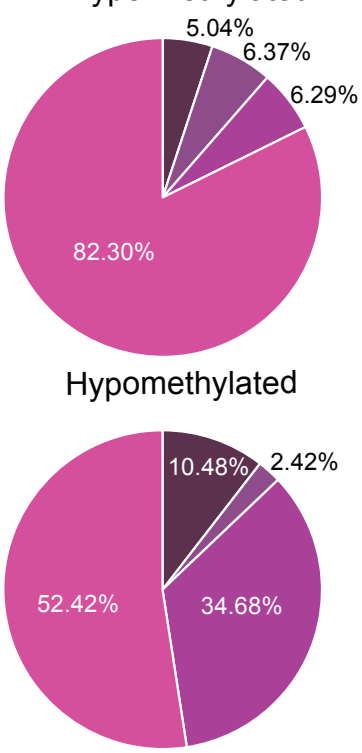


B

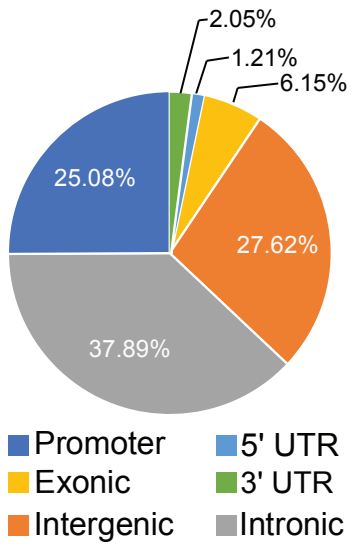




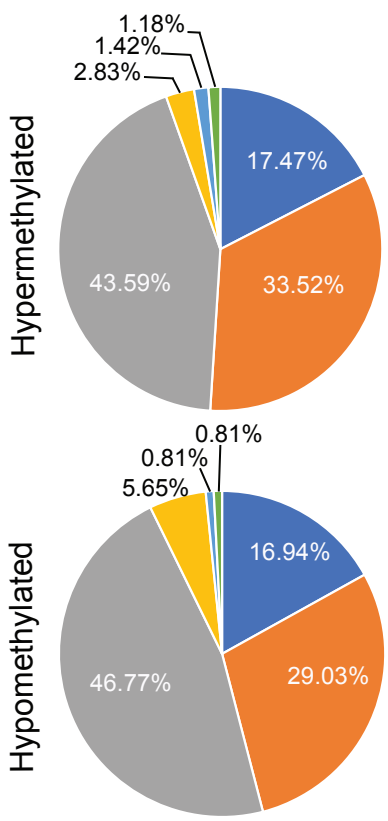
C



D



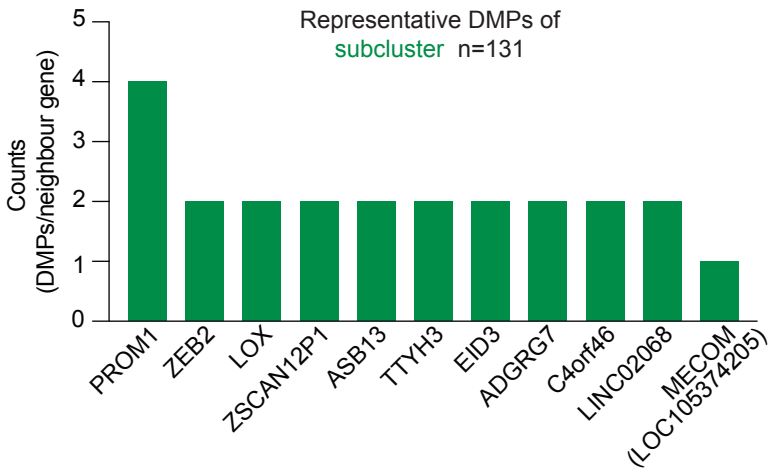
E



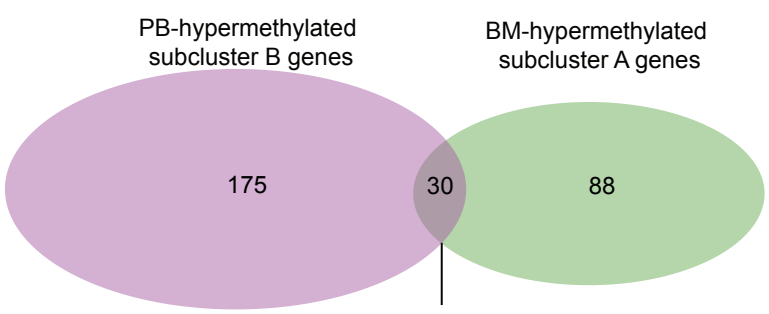
F

	Bone Marrow		Peripheral Blood	
	HyperBM	HypoBM	HyperPB	HypoPB
Exonic	0,139	0,006009	8,04E-06	1
Intergenic	<2,2 E-16	<2,2 E-16	<2,2 E-16	0,05882
Promoter	0,8227	0,009366	0,02319	0,4501
Intronic	<2,2E-16	2,00E-09	<2,2E-06	2,44E-06
3’	0,148	0,684	0,44518	0,5267
5’	0,02866	0,1807	0,55	1

G



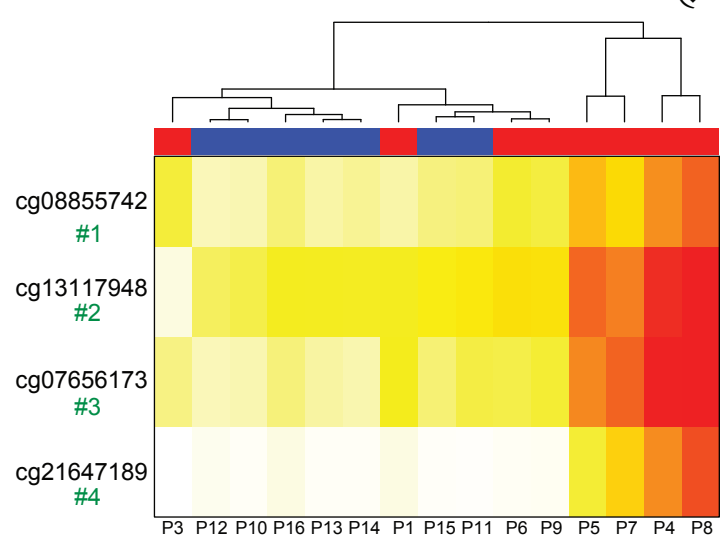
H

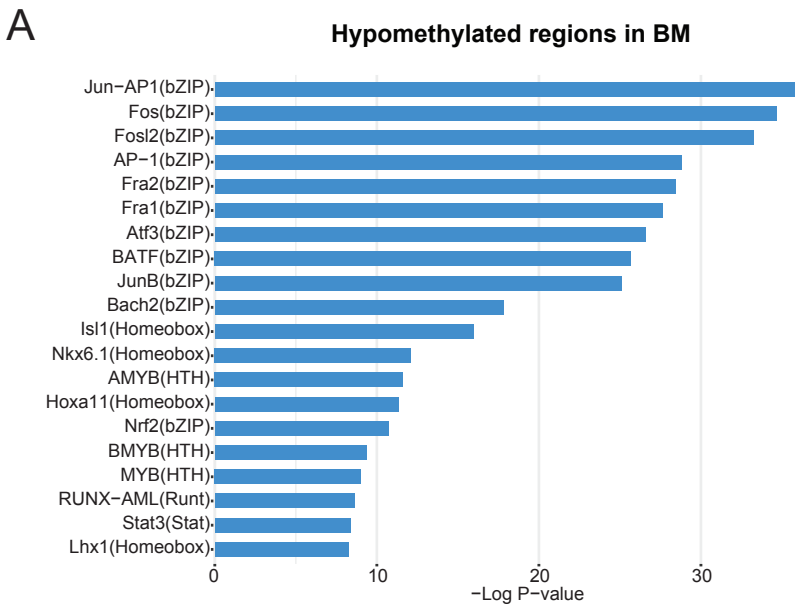


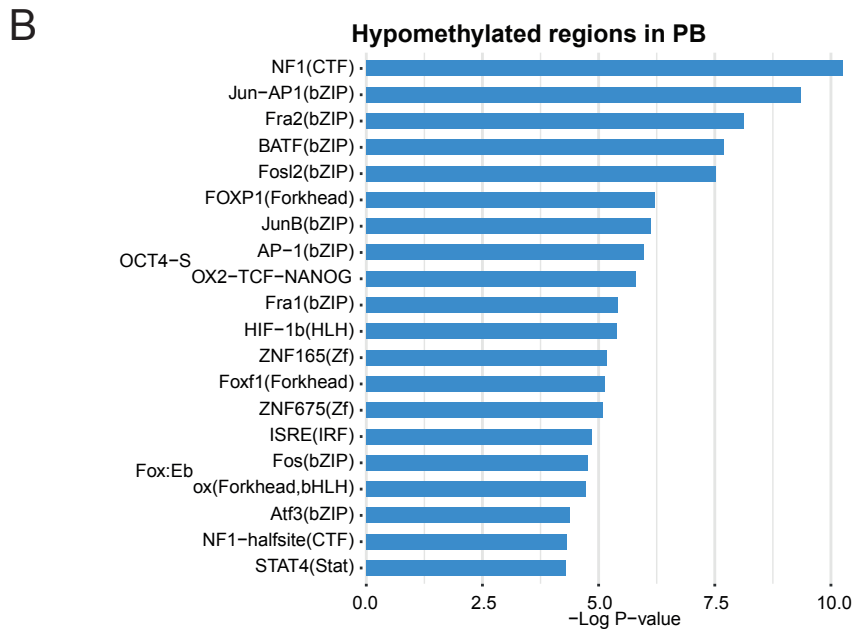
Hypermethylated genes of PB and BM subclusters

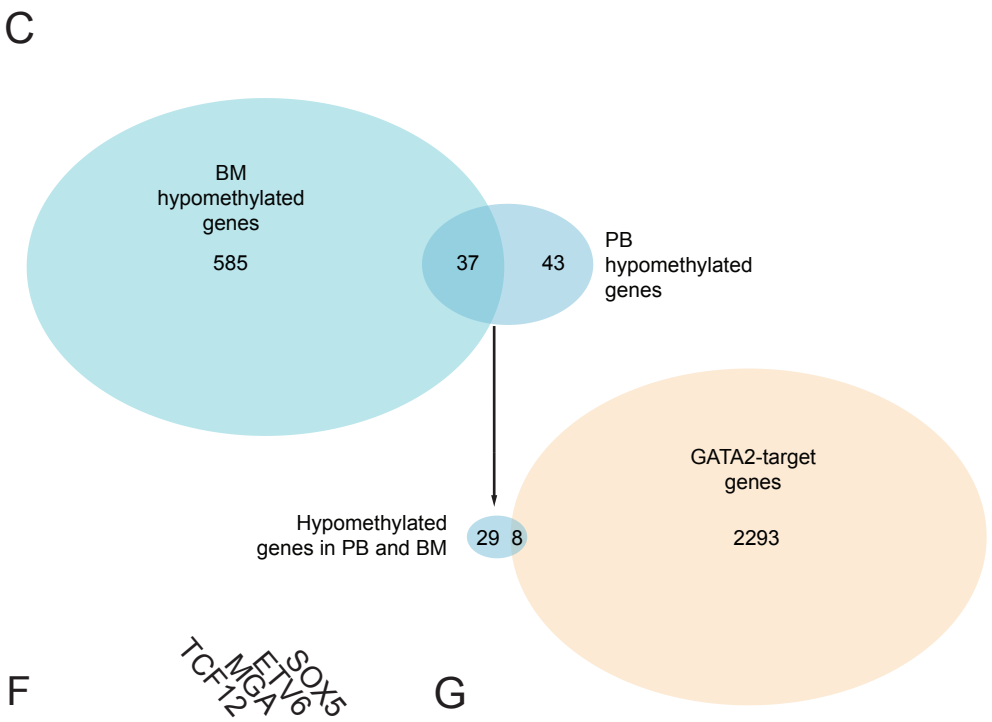
GACAT1	MIR563	PRICKLE2	LOC105374205	PLCB1-IT1	MIR4675
PIMREG	TOP1P2	MGC34796	EHF	CALCRL	ZEB2
MIR1203	CNTN4	DNAH9	CAMTA1	PIK3C2G	TTYH3
FREM1	SPP2	HPN-AS1	NXPH1	TMCC3	HTR2A
ADAMTS6	ZFPM2	PCDH9	MAS1L	ADGRG7	CCNG1

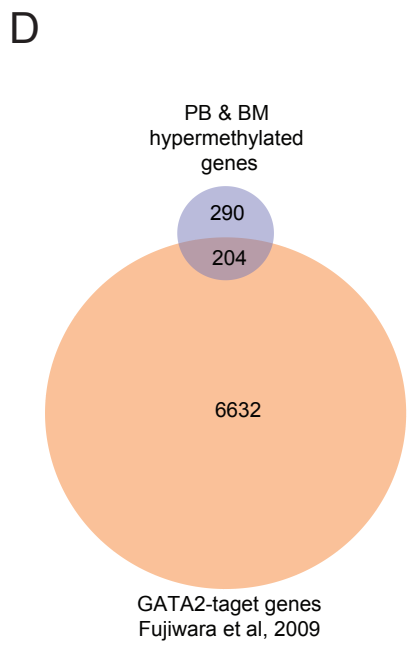
I

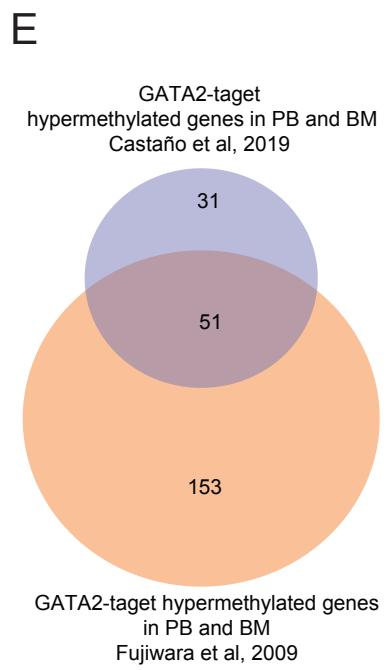


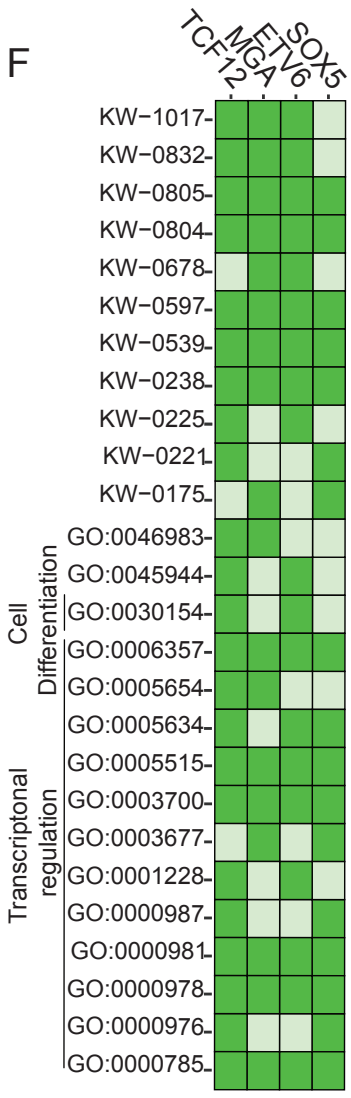

A
Hypomethylated regions in BM
Jun−AP1(bZIP)
Fos(bZIP)
Fosl2(bZIP)
AP−1(bZIP)
Fra2(bZIP)
Fra1(bZIP)
Atf3(bZIP)
BATF(bZIP)
JunB(bZIP)
Bach2(bZIP)
Isl1(Homeobox)
Nkx6.1(Homeobox)
AMYB(HTH)
Hoxa11(Homeobox)
Nrf2(bZIP)
BMYB(HTH)
MYB(HTH)
RUNX−AML(Runt)
Stat3(Stat)
Lhx1(Homeobox)
0
10
20
30
−Log P−value

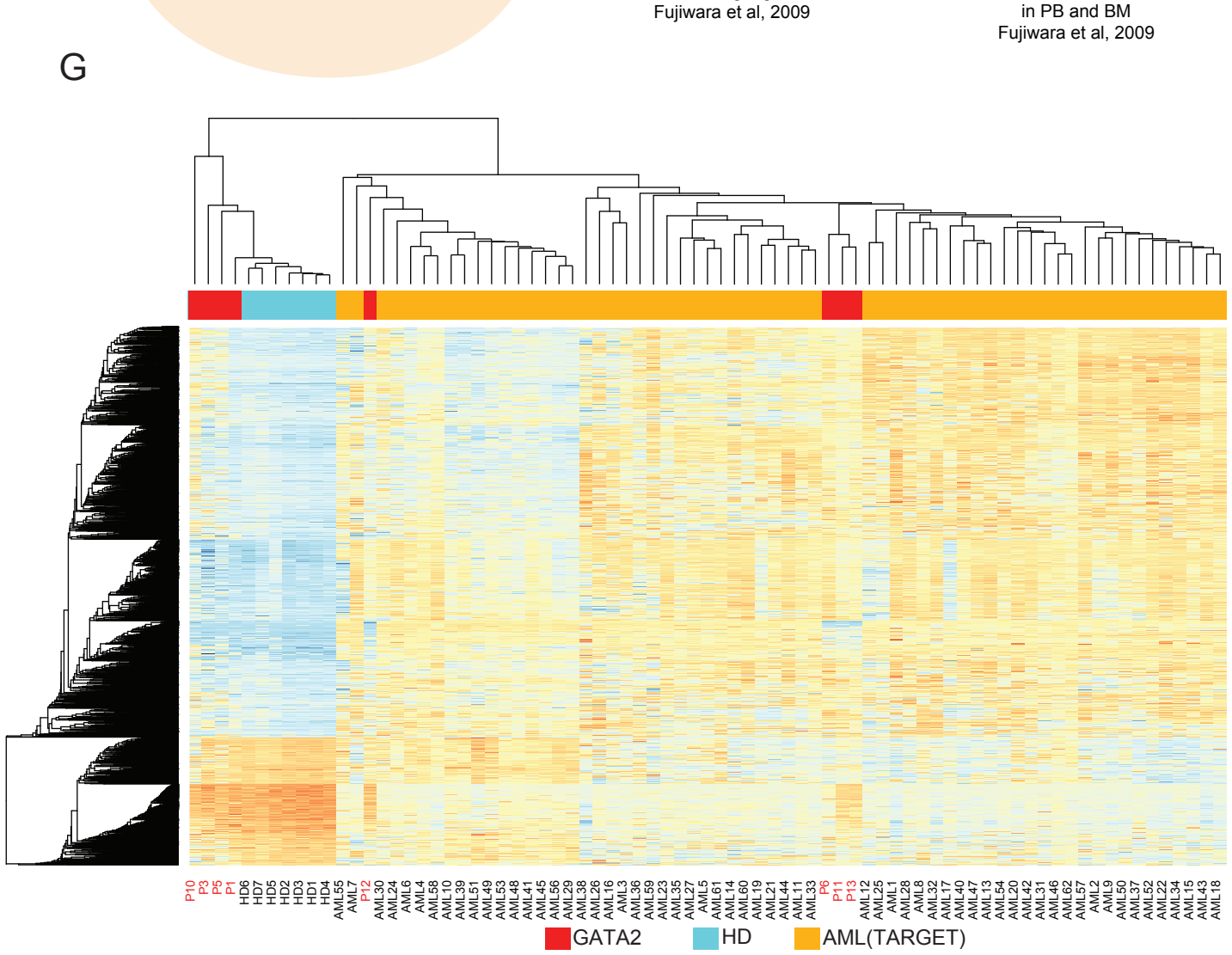

B
Hypomethylated regions in PB
NF1(CTF)
Jun−AP1(bZIP)
Fra2(bZIP)
BATF(bZIP)
Fosl2(bZIP)
FOXP1(Forkhead)
JunB(bZIP)
AP−1(bZIP)
OCT4−S
OX2−TCF−NANOG
Fra1(bZIP)
HIF−1b(HLH)
ZNF165(Zf)
Foxf1(Forkhead)
ZNF675(Zf)
ISRE(IRF)
Fos(bZIP)
Fox:Eb
ox(Forkhead,bHLH)
Atf3(bZIP)
NF1−halfsite(CTF)
STAT4(Stat)
0.0
2.5
5.0
7.5
10.0
−Log P−value


C
BM
hypomethylated
genes
585
37
43
PB
hypomethylated
genes
GATA2-target
genes
2293
Hypomethylated
genes in PB and BM
29
8


D
PB & BM
hypermethylated
genes
290
204
6632
GATA2-taget genes
Fujiwara et al, 2009


E
GATA2-taget
hypermethylated genes in PB and BM
Castaño et al, 2019
31
51
153
GATA2-taget hypermethylated genes
in PB and BM
Fujiwara et al, 2009


F
TCF12
MGA
ETV6
SOX5
KW−1017
KW−0832
KW−0805
KW−0804
KW−0678
KW−0597
KW−0539
KW−0238
KW−0225
KW−0221
KW−0175
GO:0046983
GO:0045944
GO:0030154
GO:0006357
GO:0005654
GO:0005634
GO:0005515
GO:0003700
GO:0003677
GO:0001228
GO:0000987
GO:0000981
GO:0000978
GO:0000976
GO:0000785
Cell
Differentiation
Transcriptonal
regulation


G
P10
P3
P5
P1
HD6
HD7
HD5
HD2
HD3
HD1
HD4
AML55
AML7
P12
AML30
AML24
AML6
AML4
AML58
AML10
AML39
AML51
AML49
AML53
AML48
AML41
AML45
AML56
AML29
AML38
AML26
AML16
AML3
AML36
AML59
AML23
AML35
AML27
AML5
AML61
AML14
AML60
AML19
AML21
AML44
AML11
AML33
P6
P11
P13
AML12
AML25
AML1
AML28
AML8
AML32
AML17
AML40
AML47
AML13
AML54
AML20
AML42
AML31
AML46
AML62
AML57
AML2
AML9
AML50
AML37
AML52
AML22
AML34
AML15
AML43
AML18
GATA2
HD
AML(TARGET)

Supplemental Figure 1. Patient clinical characteristics and genetic landscape.

A) Summary of the GATA2 patient data: Pediatric (n=6) with asymptomatic carrier (n=1) and adult (n=14) with asymptomatic carriers (n=2).

B) Box plot showing age and gender distribution at the DNA sample collection.

C) Patient symptoms at diagnose. Myelodisplastic syndrome (MDS, 55%, dark blue), immunodeficiency (ID, 15%, light grey) asymptomatic (15%, light blue), acute myeloid leukemia (10%, indigo) and deafness (5%, dark grey).

D) Pie chart showing the type of the different GATA2 mutations of the Spanish cohort n=20. Type of mutations and ratio: Synonymous, 5%; insertion, 5%; nonsense, 10%; frame-shift, 15% and missense, 65%

E) Schematic of the GATA2 protein and aminoacidic mutation position. Location of germline variants of 20 individuals in our series are depicted in relation to the Zinc finger DNA binding domains. Circles represent individual cases and are colour coded by recurrent mutations (black) or novel mutations (red).

F) Aggregation of the somatic mutations identified in our cohort of patients. Abbreviations; normal karyotype (bourdeaux), chromosome 8 trisomy (light blue), complex (dark blue) and not available (NA, grey).

Supplemental Figure 2. Unsupervised hierarchical clustering and *heat map* visualization of differentially *methylated* CpG sites of GATA2-mutant patients

A) Pie chart indicating the number of differentially methylated probes (DMPs); hypomethylated (light blue) and hypermethylated (royal blue) in both, bone marrow (BM) and peripheral blood (PB).

B) Pie chart indicating the percentage of probes of the Infinium MethylationEPIC array based on CpG island distance, island (eggplant), shore (lollipop), shelf (mauve) and open sea (fandango).

C) DMPs distribution of PB samples, hypermethylated (above) and hypomethylated (below). CpG island distance, island (eggplant), shore (lollipop), shelf (mauve) and open sea (fandango).

D) Pie chart indicating the percentage of probes of the Infinium MethylationEPIC array based on gene distance, promoter (dark blue), exonic (yellow), intronic (grey), intergenic (orange), 5' UTR (light blue) and 3' UTR (green).

E) DMPs distribution of PB samples, hypermethylated (above) and hypomethylated (below). Promoter (dark blue), exonic (yellow), intronic (grey), intergenic (orange), 5' UTR (light blue) and 3' UTR (green).

F) Correlation analysis of DMP distribution of GATA2-mutant samples versus the probe distribution of the Infinium MethylEPIC array (Chi-square). Promoter (dark blue), exonic (yellow), intronic (grey), intergenic (orange), 5' UTR (light blue) and 3' UTR (green).

G) Histogram showing the proximal gene with more than 2 hypermethylated DMPs such as *PROM1, ZEB2, LOX, ZSCAN12P1, ASB13, TTYH3, EID3, ADGRG7, C4orf46, LINC02068* and *MECOM* with only one DMP.

H) Venn diagram of the hypermethylated genes in BM samples (n=131) versus subcluster of hypermethylated genes in PB samples (n=205), the genes of the intersection are 30 genes. Pvalue=7.763154e-15 (hypergeometric distribution test).

I) Heatmap showing in red on the right side of the heatmap, the patients with hypermethylated DMPs in PROM1 promoter region (P4, P5, P7 and P8).

Supplemental Figure 3. Transcription factor regulation in hypermethylated genomic regions of GATA2 patients.

A) Hypergeometric Optimization of Motif EnRichment (HOMER) analysis using hypomethylated differentially methylated probes (DMPs) in bone marrow (BM) samples. Enriched motifs found are predominantly from the bZIP family (Jun-Fos).

B) HOMER analysis using hypomethylated DMPs in PB samples. Enriched motifs found are predominantly from the bZIP.

C) Venn diagram of BM hypomethylated genes (n=622) crossed with PB hypomethylated genes (n=80) the intersection (n=37 genes) is crossed with GATA2 target genes (n=2301). The intersection shows 8 hypomethylated genes in PB and BM.

D) Venn diagram of the hypermethylated genes in both BM and PB samples (n0494) compared to the GATA2 regulated genes ChIPseq of K562 described in Fujiwara et al 2009 (n=7212) GSE18868, intersection 204 genes.

E) Venn diagram of the hypermethylated genes are targets of GATA2, Castaño *et al*, 2019 GSE107639, compared with hypermethylated genes, which are targets of GATA2 in K562, Fujiwara, *et al* 2009, GSE18868, intersection 51 out of 82 genes.

F) Gene ontology analysis of GATA2 target genes hypermethylated in both BM and PB showed enrichment in transcriptional regulation and cell differentiation pathways. The *in silico* analysis predicts TCF12, MGA, ETV6 and SOX5 as cooperative TFs of the GATA2 gene regulatory network. Colour code: green, enrichment and light green, no enrichment.

G) Unsupervised hierarchical clustering and the heat map associated with the methylation profile (according to the colour scale shown) GATA2 patients (red) integrated in the cohort of AML patient (TARGET) (orange) with the healthy donor (HD, blue) in GATA2 BM patient samples.