Germline GATA1s generating mutations predispose to leukemia with acquired trisomy 21 and Down syndrome-like phenotype

Henrik Hasle,[1] Ronald M. Kline[2], Eigil Kjeldsen,[3] Nik F. Nik-Abdul-Rashid[4], Deepa Bhojwani[5], Jeffrey M. Verboon[6,7], Stephanie P. DiTroia[7], Katherine R. Chao[7], Klas Raaschou-Jensen[8], Josefine Palle[9], Michel Zwaan[10], Charlotte Guldborg Nyvold[11], Vijay G. Sankaran[,6,7,12], Alan B. Cantor[6,12]

Supplemental material

Family	Patient	GATA1 mutation	Sex	Disease Course	Leukemia cytogenetics
1	II-3	Wt/GATA1 c.-21 A>G	Female	Healthy with slight thrombocytopenia (134-193 x10^9/L)	
	III-2	GATA1 c.-21 A>G	Male	Lifelong macrocytic anemia; 47 years: MDS with trilineage dysplasia; treated with MUD HSCT; in remission 14 months s/p HSCT	Normal
	III-3	Wt/GATA1 c.-21 A>G	Female	Healthy with normal hematology	
	IV-2	Wt/GATA1 c.-21 A>G	Female	27 months: AMKL treated with chemotherapy 6 years: MDS progressing to AML treated with haplo HSCT; in remission 8.5 years s/p MDS/AML treatment	48,XX,der(6)del(6)(q15q22),der(7)ins(6;7)(q23q25;q32),+21,+21[24]/46,XX[1] 46,XX,-7, t(11;21)(q23;q21),+21[25]
2	I-3	unknown	Female	Teenager: Died of leukemia, subtype unknown.	N/A
	II-2	Wt/GATA c.2 T>C	Female	Lifelong macrocytic anemia	
	III-2	Wt/GATA c.2 T>C	Female	18 months: AMKL s/p chemotherapy; In remission 5 yrs post-chemotherapy; bone marrow hypoplastic with mild erythroid and megakaryocytic dysplasia	48,XX,+11,+21/48,idem,t(1;14)(q21;q32)/49,idem+8
	III-3	GATA c.2 T>C	Male	6 months: diagnosed with congenital dyserythropoietic anemia (CDA), red blood cell transfusion-dependent 6 years: underwent cord blood HSCT; Doing well 9 years s/p HSCT	
	III-4	GATA c.2 T>C	Male	2 months: developed transfusion-dependent anemia 7 months: bone marrow revealed mild erythroid and megakaryocyte dysplasia 21 months: developed AMKL; s/p chemotherapy followed by MUD HSCT due to minimal residual disease; in remission 11 years s/p HSCT	N/A

Supplemental Table 1. Clinical course summary of family members containing *GATA1* variants. AMKL, acute megakaryoblastic leukemia; CDA, congenital dyserythropoietic anemia; HSCT, hematopoietic stem cell transplant; MDS, myelodysplastic syndrome; MUD, matched unrelated donor.

Supplemental Methods

Study Approval

All family members provided written informed consent to participate in this study. The study was approved by the institutional review boards of Boston Children's Hospital and Massachusetts Institute of Technology, USA and the ethical committee of the Health Region Midt, Denmark.

Targeted GATA1 sequencing

For family 1 genomic DNA was purified from Ficoll separated blood or bone marrow (BM) or from paraffin embedded material. Sanger sequencing of *GATA1* exon 2 and adjacent regions were performed on genomic DNA using published PCR primers. Myeloid next generation sequencing (NGS) in the index patient of family 1 was performed with Sophia Genetics.

Whole exome sequencing data processing

For family 2, genomic DNA was purified from peripheral blood. Whole exome sequencing (WES) and data processing were performed by the Genomics Platform at the Broad Institute of MIT and Harvard (http://genomics.broadinstitute.org/) using Illumina Nextera exome capture (~38 Mb target) and sequenced (150 bp paired reads) to cover >80% of targets at 20x and a mean target coverage of >100x. WES sequencing data was processed using a pipeline consisting of Picard with genomic alignment using the BWA aligner (BWA-mem) to the human genome build 38. Single nucleotide variants (SNVs) and insertions/deletions (indels) were jointly called across all samples in the full CMG cohort using Genome Analysis Toolkit (GATK) HaplotypeCaller package version 3.5. Default filters were applied to SNV and indel calls using the GATK Variant Quality Score Recalibration (VQSR) approach.

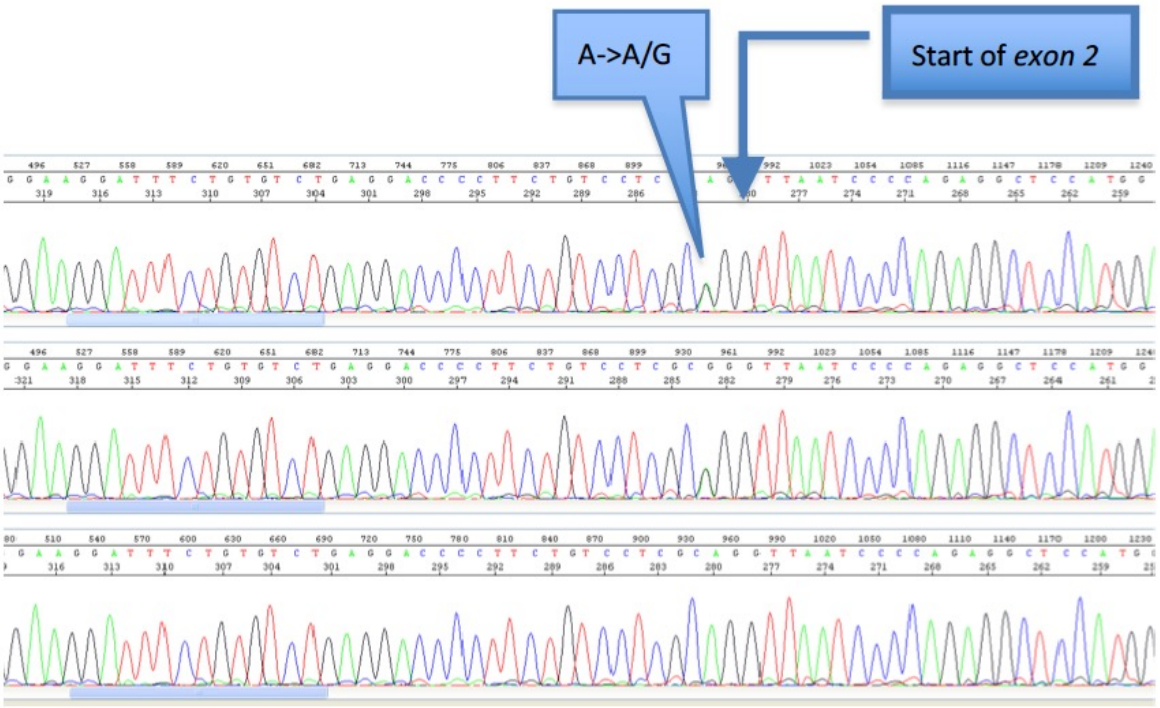
Variant annotation and identification

Called variants were annotated as previously described[1]. Two main inheritance models were explored: X-linked recessive for the anemia seen in the two affected males (III-3 and III-4), and dominant inheritance for the leukemias seen throughout the family. In both inheritance models, the *GATA1* variant chrX:48791111:T>C was flagged as the likely casual variant. The genes *ANKRD26, ETV6, GATA2, CEBPA, RUNX1, SAMD9, SAMD9L* and *DDX41* were manually inspected for potentially relevant variants, with none identified.

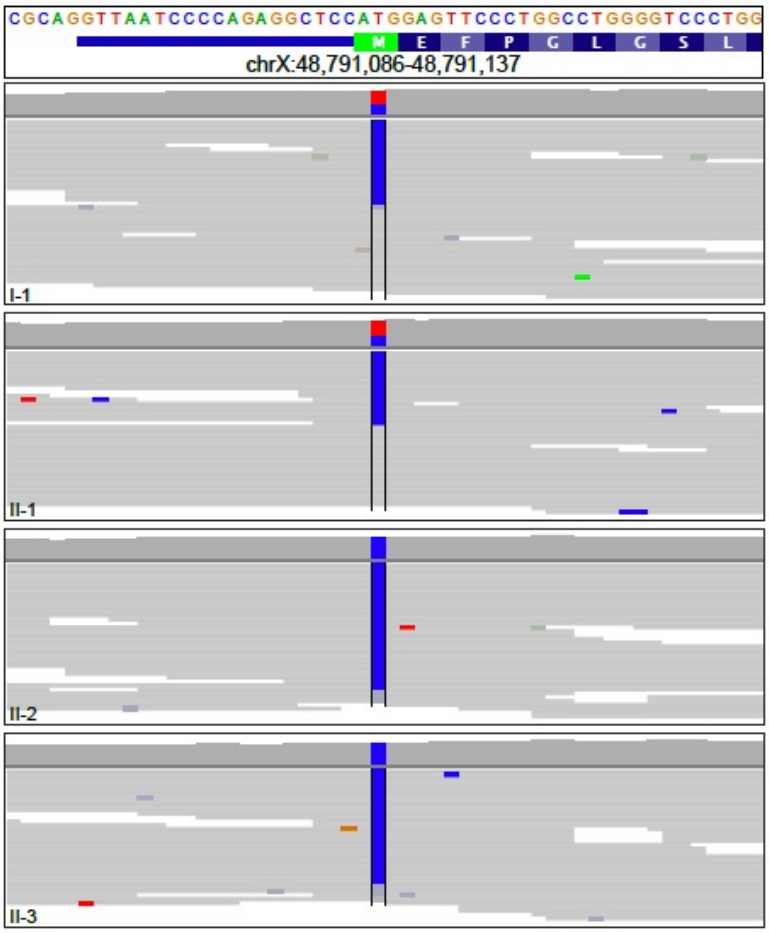
Supplemental References

1. Ulirsch JC, Verboon JM, Kazerounian S, et al. The Genetic Landscape of Diamond-Blackfan Anemia. *Am J Hum Genet* 2018; **103**(6): 930-47.

Intron-sequence and the start of *GATA1 exon 2*.



Supplemental Figure 1. Sanger sequencing electropherogram of *GATA1* showing the heterozygous genotype of paired blood (A) and bone marrow (B) from Family 1, patient IV-2. (C): A normal control showing homozygous genotype at the same genomic position. Arrows point at respectively the heterozygous position and the start of exon 2.



GATA1:p.Met1?
(chrX:48791111:T>C)

Supplemental Figure 2. Whole exome sequencing (WES) of peripheral blood of family 2. The top of the figure shows the reference DNA and translated protein sequence around the *GATA1* exon 2 translational ATG start position. The top portion of each patient's data summarizes the depth of reads per position with reads containing the reference T (shown in red) or the mutant C (shown in blue) at position chrX:48791111 (hg38). The bottom portion of each patient's data shows the individual read alignments. Positions that match the reference nucleotide are shown in gray. Nucleotides that do not match the reference sequence are are indicated in their respective color, C (blue), T (red), G (yellow), A (green). The intensity of the color indicates the relative confidence of the nucleotide call. Patient II-2 (mother of proband) had 45% reads containing c.2 T>C and 54% with the reference sequence. Patient III-3 (proband in remission) had 50% reads containing c.2 T>C and 49% with the reference sequence.