Donor-derived MDS/AML in families with germline *GATA2* mutation

Pallavi Galera[1], Amy P. Hsu[2], Weixin Wang[1], Stephenie Droll[1], Rui Chen[1], Jason R. Schwartz[3], Jeffery M. Klco[3], Sally Arai[4], Luke Maese[5], Christa Zerbe[2], Mark J. Parta[6], Neal S. Young[7], Steven M. Holland[2], Dennis D. Hickstein[8], Katherine R. Calvo[1*]

SUPPLEMENTARY METHODS

Bone marrow (BM) aspirates, biopsy, and immunohistochemistry:

The BM biopsy specimens were fixed in 3.7% neutral buffered formalin with 0.5-1% zinc chloride for a minimum of 2 hours and then decalcified for 1 hour in RapidCal Immuno. The BM specimens were routinely processed and embedded using a Tissue-Tek processor, and embedded in high polymer embedding medium on a Tissue-Tek Sakura embedding instrument. Hematoxylin and eosin staining for the biopsy specimen was performed on Leica Autostainer. Wright Geimsa staining for aspirate smears was performed on Midas Giemsa staining instrument. Immunohistochemical stains (CD34; anti-CD34 (QBEnd/10) and CD61; anti-CD61 (2f2)) were performed on a Ventana Benchmark Ultra platform (Ventana Medical Systems). BM biopsies and aspirates were reviewed by at least 2 pathologists.

Flow cytometric analysis:

Flow cytometric analysis of BM aspirates was performed using FACS Canto II Analyzer (Becton Dickinson Biosciences (BD)) equipped with 3 lasers and 8 fluorescent detectors. Antibodies (and fluorochromes) against the following antigens were used: CD45(AmCyan/ V500, BD Bioscience), CD45RA (Pacific Blue, BD Horizon), CD3(PE-Cy7, BD Pharmigen), CD4 (APCH7, BD Pharmingen), CD5 (PE, BD Biosciences), CD7 (APC, Beckman Coulter), CD8 (PerCP-Cy5.5, BD Biosciences), CD10 (PE, Beckman Coulter), CD11b (APC, BD Biosciences), CD13 (PE, BD Biosciences), CD14 (APCH7, BD Biosciences), CD15 (PE-Cy7, BD

Biosciences), CD16 (FITC, BD Biosciences), CD19 (PE Cy7, BD Biosciences), CD20 (Pacific Blue, Beckman Coulter), CD23 (FITC, Beckman Coulter), CD33 (PerCP-Cy5.5, BD Biosciences), CD34 (PerCP, Cy5.5, BD Biosciences), CD36 (FITC, BD Biosciences), CD38 (APC Alexa 750, Beckman Coulter), CD56 (PC7, Beckman Coulter), CD42b (FITC, BD Biosciences), CD56 (PC7, Beckman Coulter), CD57 (FITC, BD Biosciences), CD61(FITC, BD Biosciences), CD62L (APC, BD Pharmingen), CD64 (PE, Invitrogen), CD71 (APC Alexa70, Beckman Coulter), CD117 (PerCP-Cy5.5, BD Biosciences), CD123 (APC, BD Biosciences), CD138 (PE, BD Biosciences), HLA-DR (Pacific Blue, Biolegend), and surface kappa (FITC, BD Biosciences) and lambda (PE, BD Biosciences) light chains. Cells were stained with the above-mentioned antibodies in appropriate dilutions for 15 minutes. After red blood cell lysing with BD FACS lysing solution, cells were washed with phosphate-buffered saline containing 1% albumin. Cells were fixed in a 1% paraformaldehyde solution, and 1×10^5 events were acquired using FACS Diva software (BD Biosciences). The list mode files were analyzed with FCS Express (DeNovo Software).

***GATA2* sequencing:**

Family #1: Proband and donor parent *GATA2* mutation (c.182A>G, p.Q61R) was identified by gene sequencing panel utilized for detecting mutations in genes related to germline predisposition of myeloid neoplasms.

Family #2: For the donor monozygotic twin from the second family, all exons and intronic regulatory sites of GATA2 were amplified and analyzed as described previously in Hsu A et al. Blood. Blood 2013; 121(19): 3830-7, S1-7. A *GATA2* mutation (c.1082G>A; p.R361H) was identified.

Family #3: Initial targeted sequencing and full *GATA2* gene sequencing of the specimen from the proband of the third family by next generation sequencing was inconclusive. Subsequently, a long PCR spanning exons 2-7, 5-7 and 6-7 was performed. A shorter than expected PCR product was observed and

sequencing revealed a 426 nucleotide deletion beginning 50 bases before the start of exon 6 (c.1018-50) and extending 247 nucleotides past the end of exon 6 (c.1143+247) resulting in a genomic deletion of 426 nucleotides corresponding to an in frame deletion of amino acids 340 through 381, inclusive.

Targeted sequencing and analysis for somatic mutations in genes associated with MDS/AML:

The initial AML diagnostic specimen from the proband in Family #1 was analyzed using SNaPshot method (single base extension) designed to detect mutations in *FLT3 D835, NPM1, APC, BRAF, CTNNB1, EGFR, IDH1, IDH2, KRAS, NOTCH1, NRAS, DNMT3A, PIK3CA, PTEN, TP53, ERBB2/HER2, CEBPA, C-KIT* exons 8 and 17; *FLT3* internal tandem duplication analysis was also performed.

The remaining bone marrow specimens for which there was viably frozen material available, as indicated in the manuscript text, were profiled using TruSight Myeloid Sequencing Panel (illumina, San Diego, CA), which covers the hotspots of 54 genes related to the myeloid malignancies. Genomic DNA (100 ng) from patient bone marrow cells was amplified and then sequenced using MiSeq instrument (Illumina). The sequencing files were annotated and analyzed using variant studio software (Illumina). The CADD scores were estimated using the https://cadd.gs.washington.edu/score website[1]. The criteria used for filtering the variants is as follows: read depth > 500, variant allele frequency ≥ 5%, variants are exonic, population frequency ≤ 1%, and CADD-score ≥ 10. The identified variants (*STAG2* and *SETBP1*) were further confirmed by Sanger sequencing using the primers: STAG2_Tyr331_F CCCCCATTCAGTGATGCGAT, STAG2_Tyr331_R GGGGAGGCTTCCAGAAATGT; STAG2_Arg604_F TTTCCCTAAATGCCTCACAGAA, STAG2_Arg604_R AGTCAAACTAAAGTACAGAAGACCT; and SETBP1_Gly870Ser_F AATCTCCAGCCCATCAGTGC, SETBP1_Gly870Ser_R TTCTTTGTGCTGGTGTCGGA.

Supplementary Methods References:

1. Kircher M, Witten DM, Jain P, et al. A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet. 2014 Mar;46(3):310-5.