GATA2 deficiency phenotype associated with tandem duplication of GATA2 and overexpression of GATA2-AS1

Preeti Singh, Maninder Heer, Anastasia Resteu, Aneta Mikulasova, Mojgan Reza, Laëtitia Largeaud, Stéphanie Dufrechou, Nais Prade, Rachel E. Dickinson, Jacinta Bustamante, Bénédicte Neven, Venetia Bigley, Eric Delabesse, Daniel Rico, Marlène Pasquet, and Matthew Collin

A 3-year-old girl of nonconsanguineous healthy parents presented with cervical and mediastinal lymphadenopathy due to Mycobacterium fortuitum infection. Routine blood analysis showed normal hemoglobin, neutrophils, and platelets but profound mononuclear cell deficiency (monocytes < 0.1 × 10⁹/L; B cells 78/μL; NK cells 48/μL). A 548,902-bp region containing GATA2 was sequenced by targeted capture and deep sequencing. This revealed a de novo 187-kb duplication of the entire GATA2 locus, containing a maternally inherited copy number variation deletion of 25 kb (GRCh37: esv2725896 and nsv513733). Many GATA2-associated phenotypes have been attributed to amino acid substitution, frameshift/deletion, loss of intronic enhancer function, or aberrant splicing. Gene deletion has been described, but other structural variation has not been reported in the germline configuration. In this case, duplication of the GATA2 locus was paradoxically associated with skewed diminished expression of GATA2 messenger RNA and loss of GATA2 protein. Chimeric RNA fusion transcripts were not detected. A possible mechanism involves increased transcription of the anti-sense long noncoding RNA GATA2-AS1 (RP11-472.220), which was increased several fold. This case further highlights that evaluation of the allele count is essential in any case of suspected GATA2-related syndrome.

Introduction

GATA2 is expressed in early hematopoietic progenitors and is required for stem cell longevity and lympho-myeloid differentiation. Germline heterozygous mutation of GATA2 is associated with progressive loss of mononuclear cell production, immunodeficiency, myelodysplasia, familial acute myeloid leukemia, lymphedema, and deafness. Hundreds of cases associated with sequence variation in the GATA2 gene causing amino acid substitution, frameshift/deletion, loss of intronic enhancer function, or aberrant splicing have been described. Structural variation is less common, although several cases of germline gene deletion have been described, and displacement of the distal enhancer by inv(3)(q21;q26) or t(3;3)(q21;q26) is well documented in high-risk leukemia. Here, we describe a patient presenting with typical features of mycobacterial infection associated with severe...
monocytopenia, B-cell and NK cell lymphopenia, and myelodysplasia who had a de novo 187-kb duplication of the entire GATA2 locus associated with skewed diminished expression of GATA2, loss of GATA2 protein, and increased transcription of the antisense long noncoding RNA GATA2-AS1 (RP11-472.220).

Case description

A 3-year old girl of nonconsanguineous healthy parents presented with cervical and mediastinal lymphadenopathy due to Mycobacterium fortuitum infection (2010). Her parents reported that she had hearing difficulties. Routine blood analysis showed normal hemoglobin, neutrophils, and platelets but profound mononuclear cell deficiency (monocytes < 0.1 x 10^9/L; B cells 78/μL; and NK cells 48/μL). Immunoglobulins were normal (IgG, 11.3 g/dL; IgA, 0.95 g/dL, IgM, 1.72 g/dL), but mitotic responses of whole blood were diminished. She was treated with ciprofloxacin, amikacin, sulfamethoxazole, and linezolid but did not respond and required surgical intervention for airway compression due to mediastinal lymphadenopathy. Sensorineural deafness was documented after antibiotic therapy. She was subsequently treated with interferon-γ and her condition slowly improved. However, by the age of 8 years she developed a chronic respiratory illness with persistent cough, fatigue, nodular and ground glass infiltrates, and reduced diffusion capacity. Lung biopsy showed peribronchiolar infiltrates, but no evidence of infection or pulmonary alveolar proteinosis. Bone marrow examination revealed multi-lineage dysplasia with trisomy 8 in 5 of 22 metaphases. She underwent successful allogeneic bone marrow transplantation using a matched unrelated donor at age 9 years. Four years later she has normal cell counts and remains free of infections and respiratory compromise. Genetic screening was negative for mutations of IFNGR1, IFNGR2, STAT1, IL12B, IL12RB1, ISG15, and IRF8. No mutations of GATA2 were detected in coding regions or enhancers, but MLPA suggested the existence of a whole gene duplication. Fms-like tyrosine kinase-3 (Flt3) ligand was markedly elevated (7940 pg/mL).

Methods

Patient material was obtained with informed consent through the French GATA2-like Project and assigned to Newcastle Biobank by material transfer agreement (Newcastle and North Tyneside 1 Research Ethics Committee Reference 17/NE/0361). Whole-blood DNA of the patient, both parents, and 2 unrelated healthy controls was extracted using a QIAamp DNA Mini-Kit (QIAGEN). A 548 902-bp region containing GATA2, flanked by EEFSEC and RPN1, was obtained by custom capture with Nimblegen baits for deep sequencing on Illumina platforms (98% coverage ≥ 1000 fold) (Figure 1A). Sequence was aligned to GRCh37 using Burrows-Wheeler Aligner. Structural variation was analyzed with ONCOCNV v6.8, using the parents and healthy controls as baseline and by manual inspection of discordant read pairs and soft-clipped reads in Integrative Genomics Viewer. GATA2 protein was detected by western blot using rabbit polyclonal antibody (Abcam; ab153820). Allelic expression was assessed in cDNA amplified with CCCGGGACTGGTGCTCTTTC/GGCAGAAACCCTGTGGGTCC (rs1573858); CCCTCTAGAGCCCTGTAGTTCCT/TGAGCGGGGT TGAGCATAGTAG (rs2335237). Stranded messenger RNA sequencing was performed with Illumina NextSeq to obtain 150-bp paired-end reads at 134 to 155 million reads per sample on libraries prepared with

![Figure 1. Region of chromosome 3 showing copy number variation in the patient. (A) The captured region (3:127849971-128398873) is indicated by dashed lines, the duplication detected (3:128190688-128404309) is shown in green, and the deletion (3:128379023-128303996) is shown in blue. (B) The normalized read count mapping to the captured region. (C) The predicted arrangement of alleles in the parents and patient. The asterisk (*) indicates the position of GATA2, and the broken line denotes the site of the deletion.]
a SmartSeq Kit 2 (Takara) from RNA extracted with a QIAamp RNA Micro Kit (QIAGEN). Reads were aligned with STAR17 and examined for fusions with Arriba.18

Results and discussion

The patient had a de novo tandem duplication of 187 kb spanning GATA2 and RPN1. The duplication, which has not been previously reported, contained a deletion of 25 kb 5’ of RPN1. The deletion was inherited from the mother (Figure 1) and is a relatively common copy number variation reported in 4% of Europeans (GRCh37: esv2725896 and nsv513733).19 The prevalence of deletions in this region indicates that it is susceptible to structural variation. However, the deletion alone has no reported phenotype, even though it removes an alternative transcription start site for RPN1 at 128 400 kb, associated with CTCF and H3K27ac binding peaks. A second copy of GATA2 was joined to this region by the tandem duplication. It was not obvious why duplication of the GATA2 region should result in a classical GATA2-deficiency phenotype. GATA2 was not detected in patient peripheral blood mononuclear cells (PBMCs) or healthy control PBMCs by western blot. However, a consistent reduction was observed in the patient using primary fibroblast lines.
in which GATA2 is easily detected (Figure 2A). Analysis of heterozygous single-nucleotide polymorphisms indicated skewed expression toward the maternal allele, relative to genomic sequence chromatograms, in PBMCs and fibroblast cDNA (Figure 2B), and overall expression was diminished in 3 of 4 amplicons tested (Figure 2C).

Unexpectedly, the expression of GATA2-AS1 was increased in patient PBMCs and fibroblasts by quantitative PCR and RNA sequencing (Figure 2C-D). This long noncoding RNA, also known as RP11-47222, is widely expressed and associated with MYC-induced loss of GATA2 in several cancers.20-22 The tandem duplication brings the second copy of GATA2 much closer to its distal enhancer, potentially activating transcription of the locus. It is not clear why GATA2-AS1 predominates and appears to inhibit transcription of all 3 copies of the GATA2 gene, including the paternal allele in trans (Figure 2E). The mechanism of skewed expression of the duplicated maternal allele is not clear, although suppression of a genetic mechanism potentially involving GATA2 and monoallelic expression is observed in sporadic leukemia.24 High-depth RNA sequencing of PBMCs and fibroblasts did not reveal any novel fusion transcripts (eg, chimeric RPN1-GATA2; supplemental Table 1).

The physiological function of GATA2 is critically dependent upon its level of expression, and it is known that overexpression in the progenitor compartment can disrupt hematopoiesis.25 We did not have access to bone marrow progenitors to exclude this possibility, although the phenotypic similarity with known GATA2 deficiency suggests that loss of expression is the mechanism.

In summary, this case illustrates that structural variation and epigenetic mechanisms potentially involving GATA2-AS1 can lead to constitutive loss of GATA2 expression and a classical GATA2-deficiency phenotype. The clinical presentation of such a case remains a strong argument to avoid transplantation from a related donor, given the possibility of an inherited genetic defect. The results further emphasize that it is important to evaluate the number of GATA2 alleles in suspected cases of GATA2-related syndrome.

Acknowledgments

P.S. and M.C. received funding from UKRI (MRC) project grant MR/P002005/1. M.H. was funded by an EPSRC DTP studentship. D.R. received a Wellcome Trust Seed Award in Science (206103/Z/17/Z). The authors gratefully acknowledge the Association Laurette Fuguain and Bright Red for additional contributions to funding.

Authorship

Contribution: P.S., M.R., L.L., S.D., N.P., R.E.D., and B.N. performed research; M.H., A.R., and A.M. analyzed and interpreted data; J.B. designed research and collected data; V.B. designed research and interpreted data; E.D. and D.R. designed research and analyzed and interpreted data; M.P. designed research and collected, analyzed, and interpreted data; and M.C. designed research, analyzed and interpreted data, and wrote the manuscript.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

ORCID profiles: M.H., 0000-0002-3632-7174; M.R., 0000-0001-7898-292X; L.L., 0000-0001-5341-5427; J.B., 0000-0002-3439-2482; V.B., 0000-0002-3017-2474; E.D., 0000-0002-0928-0753; D.R., 0000-0002-6561-2310; M.P., 0000-0001-6919-8383; M.C., 0000-0001-6585-9586.

Correspondence: Matthew Collin, Newcastle University, Framlington Pl, Newcastle upon Tyne NE2 4HH, United Kingdom; e-mail: matthew.collin@ncl.ac.uk.

References

1. Laurenti E, Doulatoev S, Zandi S, et al. The transcriptional architecture of early human hematopoiesis identifies multilevel control of lymphoid commitment. Nat Immunol. 2013;14(7):756-763.
2. Hsu AP, Sampaio EP, Khan J, et al. Mutations in GATA2 are associated with the autosomal dominant and sporadic monocytopenia and mycobacterial infection (MonosMAC) syndrome. Blood. 2011;118(10):2653-2655.
3. Dickinson RE, Griffin H, Bigley V, et al. Exome sequencing identifies GATA-2 mutation as the cause of dendritic cell, monocyte, B and NK lymphoid deficiency. Blood. 2011;118(10):2656-2658.
4. Hahn CN, Chong CE, Carmichael CL, et al. Heritable GATA2 mutations associated with familial myelodysplastic syndrome and acute myeloid leukemia. Nat Genet. 2011;43(10):1012-1017.
5. Ostergaard P, Simpson MA, Connell FC, et al. Mutations in GATA2 cause primary lymphedema associated with a predisposition to acute myeloid leukemia (Emberger syndrome). Nat Genet. 2011;43(10):929-931.
6. Spinner MA, Sanchez LA, Hsu AP, et al. GATA2 deficiency: a protean disorder of hematopoiesis, lymphatics, and immunity. Blood. 2014;123(6):809-821.
7. Dickinson RE, Milne P, Jardine L, et al. The evolution of cellular deficiency in GATA2 mutation. Blood. 2014;123(6):863-874.
8. Wlodarski MW, Hirabayashi S, Pastor V, et al; EWOG-MDS. Prevalence, clinical characteristics, and prognosis of GATA2-related myelodysplastic syndromes in children and adolescents. Blood. 2016;127(11):1387-1397, quiz 1518.
9. Donadieu J, Lamant M, Fieschi C, et al; French GATA2 study group. Natural history of GATA2 deficiency in a survey of 79 French and Belgian patients. Haematologica. 2018;103(8):1278-1287.
10. Kozyra EJ, Pastor VB, Leffkopoulos S, et al; European Working Group of MDS in Childhood (EWOG-MDS). Synonymous GATA2 mutations result in selective loss of mutated RNA and are common in patients with GATA2 deficiency. Leukemia. 2020;34(10):2673-2687.
11. Vinh DC, Palma L, Storrin J, Foulkes WD. GATA2 deficiency due to de novo complete monoallelic deletion in an adolescent with myelodysplasia. J Pediatr Hematol Oncol. 2018;40(4):e225-e228.
12. Gröschel S, Sanders MA, Hoogenboezem R, et al. A single oncogenic enhancer rearrangement causes concomitant EVI1 and GATA2 deregulation in leukemia. Cell. 2014;157(2):369-381.

13. Bustamante J. Mendelian susceptibility to mycobacterial disease: recent discoveries. Hum Genet. 2020;139(6-7):993-1000.

14. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25(14):1754-1760.

15. Boeva V, Popova T, Lienard M, et al. Multi-factor data normalization enables the detection of copy number aberrations in amplicon sequencing data. Bioinformatics. 2014;30(24):3443-3450.

16. Robinson JT, Thorvaldsdottir H, Winckler W, et al. Integrative genomics viewer [letter]. Nat Biotechnol. 2011;29(1):24-26.

17. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29(1):15-21.

18. Uhrig S, Ellermann J, Walther T, et al. Accurate and efficient detection of gene fusions from RNA sequencing data. Genome Res. 2021;31(3):448-460.

19. The 1000 Genomes Project Consortium. A global reference for human genetic variation. Nature. 2015;526(7571):68-74.

20. Kaczkowski B, Tanaka Y, Kawai H, et al; FANTOM5 Consortium. Transcriptome analysis of recurrently deregulated genes across multiple cancers identifies new pan-cancer biomarkers. Cancer Res. 2016;76(2):216-226.

21. Zhang J, Sui S, Wu H, et al. The transcriptional landscape of IncRNAs reveals the oncogenic function of LINC00511 in ER-negative breast cancer. Cell Death Dis. 2019;10(8):599.

22. Zhang L, Gao L, Shao M, Sun GY. A MYC target long non-coding RNA GATA2-AS1 regulates non-small cell lung cancer growth. Neoplasma. 2019;66(6):954-962.

23. Al Seraihi AF, Río-Machín A, Tawana K, et al. GATA2 monoallelic expression underlies reduced penetrance in inherited GATA2-mutated MDS/AML. Cancer Res. 2016;76(2):216-226.

24. Ottema S, Mulet-Lazaro R, Beverloo HB, et al. Atypical 3q26/MECOM rearrangements genocopy inv(3)/t(3;3) in acute myeloid leukemia. Blood. 2020;136(2):224-234.

25. Tipping AJ, Pina C, Castor A, et al. High GATA-2 expression inhibits human hematopoietic stem and progenitor cell function by effects on cell cycle. Blood. 2009;113(12):2661-2672.