GATA2 Deficiency: A Protean Disorder of Hematopoiesis, Lymphatics and Immunity

Michael A. Spinner1*, Lauren A. Sanchez1*, Amy P. Hsu1, Pamela A. Shaw2, Christa S. Zerbe1, Katherine R. Calvo3, Diane C. Arthur4, Wenjuan Gu5, Christine M. Gould6, Carmen C. Brewer7, Edward W. Cowen8, Alexandra F. Freeman1, Kenneth N. Olivier1, Gulbu Uzel1, Adrian M. Zelazny1, Janine R. Daub1, Christine D. Spalding1, Reginald J. Claypool1, Neelam K. Giri9, Blanche P. Alter9, Emily M. Mace10, Jordan S. Orange10, Jennifer Cuellar-Rodriguez1, Dennis D. Hickstein11, and Steven M. Holland1

* Contributed equally

1Laboratory of Clinical Infectious Diseases, National Institute of Allergy and Infectious Diseases, Bethesda, MD 20892; 2Biostatistics Research Branch, National Institute of Allergy and Infectious Diseases, Bethesda, MD 20892; 3Hematology Section, Department of Laboratory Medicine, Clinical Center, National Institutes of Health, Bethesda, MD 20892; 4Laboratory of Pathology, Center for Cancer Research, National Cancer Institute, Bethesda, MD 20892; 5Support to Biostatistics Research Branch, Clinical Research Directorate/Clinical Monitoring Research Program, Frederick National Laboratory for Cancer Research, Frederick, MD 21702; 6Pediatric Pulmonology, Department of Pediatrics, Walter Reed National Military Medical Center, Bethesda, MD 20889; 7Otolaryngology Branch, National Institute on Deafness and Other Communication Disorders, Bethesda, MD 20892; 8Dermatology Branch, National Cancer Institute, Bethesda, MD 20892; 9Clinical Genetics Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD 20892; 10Center for Human Immunobiology, Texas Children’s Hospital, Baylor College of Medicine, Houston, TX 77030; 11Experimental Transplantation and Immunology Branch, National Cancer Institute, Bethesda, MD 20892

Corresponding Author:
Steven M. Holland, M.D.
Chief, Laboratory of Clinical Infectious Diseases, NIAID, NIH
CRC B3-4141 MSC 1684
Bethesda, MD 20892-1684
301-402-7684 (v)
301-480-4508 (fax)
smh@nih.gov

Scientific categories:
HEMATopoIESIS AND STEM CELLS
IMMUNobiOLOGY
MYELOID NEOPLASIA
VASCULAR BIOLOGY
Key points:

- GATA2 deficiency has a broad phenotype encompassing immunodeficiency, MDS/AML, pulmonary disease, and vascular/lymphatic dysfunction.

- Early genetic diagnosis is critical to direct clinical management, prophylaxis, transplantation, and family screening.
Abstract

Haploinsufficiency of the hematopoietic transcription factor GATA2 underlies monocytopenia and mycobacterial infections (MonoMAC); dendritic cell, monocyte, B and NK lymphoid (DCML) deficiency; familial MDS/AML; and Emberger syndrome (primary lymphedema with MDS). A comprehensive examination of the clinical features of GATA2 deficiency is currently lacking. We reviewed the medical records of 57 patients with GATA2 deficiency evaluated at the NIH from January 1, 1992-March 1, 2013 and categorized mutations as missense, null, or regulatory to identify genotype-phenotype associations. We identified a broad spectrum of disease: hematologic (MDS 84%, AML 14%, CMML 8%), infectious (severe viral 70%, disseminated mycobacterial 53%, and invasive fungal infections 16%), pulmonary (diffusion 79% and ventilatory defects 63%, pulmonary alveolar proteinosis 18%, pulmonary arterial hypertension 9%), dermatologic (warts 53%, panniculitis 30%), neoplastic (HPV+ tumors 35%, EBV+ tumors 4%), vascular/lymphatic (venous thrombosis 25%, lymphedema 11%), sensorineural hearing loss 76%, miscarriage 33%, and hypothyroidism 14%. Viral infections and lymphedema were more common in individuals with null mutations (p=0.038, p=0.006, respectively). Monocytopenia, B, NK, and CD4 lymphocytopenia correlated with the presence of disease (p<0.001). GATA2 deficiency unites susceptibility to MDS/AML, immunodeficiency, pulmonary disease, and vascular/lymphatic dysfunction. Early genetic diagnosis is critical to direct clinical management, preventive care, and family screening.
Introduction

In 2011, heterozygous mutations in the hematopoietic transcription factor GATA2 were identified as the cause of four previously described clinical syndromes: monocytopenia and mycobacterial infections (MonoMAC) syndrome;\(^1\,^2\) dendritic cell, monocyte, B and NK lymphoid (DCML) deficiency;\(^3\,^4\) familial myelodysplastic syndromes (MDS)/acute myeloid leukemia (AML);\(^5\) and Emberger syndrome (primary lymphedema with MDS).\(^6\,^7\) Interestingly, despite highly varied reasons for patient recruitment (disseminated mycobacterial infections, dendritic cell deficiency, familial MDS/AML, and primary lymphedema), there was considerable overlap among the clinical characteristics of all four cohorts. These syndromes are now recognized as different manifestations of a single genetic disorder with protean disease manifestations. More recently, GATA2 mutations have been identified in additional clinical cohorts including pediatric neutropenia,\(^8\) aplastic anemia,\(^9\) and the original case report of human NK cell deficiency.\(^10\,^11\)

GATA2 mutations appear to cause loss of function of the mutated allele leading to haploinsufficiency.\(^12\) Germline mutations appear to arise spontaneously but are then transmitted with autosomal dominant inheritance.\(^1,^2\) The clinical hallmarks of GATA2 deficiency include immunodeficiency with marked susceptibility to human papillomaviruses (HPV) and nontuberculous mycobacteria (NTM), predisposition to MDS/AML, pulmonary alveolar proteinosis (PAP), and congenital lymphedema. The laboratory findings include severe circulating monocytopenia, dendritic cell cytopenia, and B and NK lymphocytopenia. These features overlap with other genetic and acquired syndromes. Mutations in TERT, TERC, CEBPA, and RUNX1 are also associated with familial MDS/AML.\(^13\) Severe HPV infection is also characteristic of epidermodyplasia verruciformis, DOCK8 deficiency, CXCR4 deficiency,
and idiopathic CD4 lymphocytopenia. Inherited defects in the IL-12/IFN-γ axis as well as IFN-γ autoantibodies predispose to disseminated NTM infections. GM-CSF autoantibodies and mutations in the GM-CSF receptor may lead to PAP, while FLT4 and other mutations may underlie congenital lymphedema.

There is considerable clinical heterogeneity among patients with GATA2 deficiency. Age at disease onset ranges from early childhood to late adulthood and clinical presentations range from asymptomatic to life-threatening infections, leukemia, and respiratory failure. The genetic lesions in GATA2 are also heterogeneous, clustering into three different groups: missense mutations and in-frame deletions involving the C-terminal zinc finger; nonsense mutations, frameshifts, and large gene deletions predicted to result in null alleles; and regulatory mutations within the enhancer region of intron 5. Genotype-phenotype correlations have been sought, but sample sizes and mutation distribution have been limiting. We report 57 patients with GATA2 deficiency in whom we performed comprehensive prospective characterization of the disease and looked for genotype-phenotype associations.

Methods

**Subjects:** All patients or their guardians gave informed consent in accordance with the Declaration of Helsinki for Institutional Review Board-approved protocols at the National Institutes of Health (NIH) between January 1, 1992-March 1, 2013. Probands were recruited through natural history protocols for the study of NTM infection, primary immune deficiency, or inherited bone marrow failure syndromes. In patients with a positive family history, a reasonable effort was made to identify and screen all available blood relatives. The medical records of 57 patients with mutation-proven GATA2 deficiency were reviewed and extracted onto a
standardized form. All reported diagnostic testing was performed and reviewed internally according to standardized protocols. Data were included until chemotherapy for MDS/AML or hematopoietic stem cell transplantation (HSCT). Unless otherwise indicated, the only outcome evaluated post-transplant was survival.

**Genotyping:** Genomic DNA was isolated from peripheral blood, buccal swabs, or fibroblasts from all patients and available family members using Puregene DNA isolation kit (Qiagen). Amplification and sequencing of all exons, flanking splice sites, and regulatory regions was performed as previously described. Missense mutations were analyzed using PolyPhen2 to predict deleterious effects of amino acid changes. Mutation groups were categorized as follows: “missense” includes all missense mutations and in-frame deletions in the C-terminal zinc finger predicted to allow production of a stable mRNA that is translated into a mutated protein; “null” includes nonsense and frameshift mutations and large gene deletions predicted to result in nonsense-mediated decay and no protein product; “regulatory” includes mutations in the enhancer region of intron 5 with demonstrated reduced expression of wild-type protein. A fourth mutation group designated “uniallelic” includes phenotypically GATA2 deficient individuals with reduced or absent expression from one GATA2 allele but without defined mutations.

**Diagnostic criteria:** MDS, AML and chronic myelomonocytic leukemia (CMML) were defined according the 2008 World Health Organization criteria. The following were considered as severe viral infections: HPV or molluscum contagiosum infection with ≥10 warts/lesions or the presence of genital condyloma acuminata; herpes esophagitis or recalcitrant genital herpes infection; hemorrhagic varicella pneumonia; Epstein-Barr virus (EBV) viremia persisting greater than six months or EBV+ tumors; and disseminated cytomegalovirus.
Disseminated NTM and fungal infections were defined by biopsy-proven extrapulmonary involvement and identification in culture. PAP was defined by clinical presentation and demonstrated intra-alveolar accumulation of lipoproteinaceous material on lung biopsy. Lymphedema was considered in patients with chronic lower extremity and/or genital edema not attributable to other causes and severe enough to require treatment with compression stockings or lymphatic surgeries.

**Statistical methods:** Fisher’s exact test was used to detect associations between basic demographics and genotype. For time-to-event analyses, Kaplan-Meier curves were used to estimate survival; log rank tests and Cox regression, stratified on proband status and with robust variances, were used to detect associations between mutation groups. An exact log rank was used for one rare endpoint (lymphedema). Wald tests using robust variance estimates to account for within kindred correlation were used to evaluate mean differences between groups for continuous markers. Skewed laboratory markers were log-transformed, and counts were offset by half the lower limit of detection to avoid logarithms of zero. Statistical analyses were performed using R (Version 2.14.0. Vienna: R Development Core Team). P values (<0.05) were not adjusted for multiple comparisons.

**Results**

We identified 57 patients with GATA2 deficiency, 40 based on clinical presentation and 17 through family screening (**Table 1**). Fifty-four percent were female; 75% were Caucasian, 16% Hispanic, 7% African American, and 2% Asian (**Table S1**). Thirty-three mutations were characterized as missense, 12 null, 9 regulatory, and 3 uniallelic (**Figure 1**). There were no
associations between mutation group and gender (p=0.58) or race (p=0.39). Median age at first NIH visit was 30 years (range 4-76 years) with a median length of follow-up of 14 months (range 0-180 months). Fifty-three patients (93%) had a past or current history of at least one of the following clinical manifestations at last follow-up: severe viral or NTM infection, MDS/AML, PAP, or lymphedema. Initial clinical presentations were variable: severe viral infections (32%), disseminated NTM infections (28%), MDS/AML (21%), lymphedema (9%), and invasive fungal infections (4%). Median age at initial presentation was 20 years but was highly variable (range 5 months-78 years). Four patients (7%) had no apparent clinical manifestations as of the last follow-up (range 5-55 years).

The proportion of patients without symptoms was 50% by age 20, 25% by age 30, and 16% by age 40 (Figure 2). Survival declined soon after onset of symptoms, with 91% estimated survival by 5 years after onset, 84% by 10 years, and 67% by 20 years. Overall survival was 96% by age 20, 77% by age 40, and 45% by age 60. Twenty-one patients underwent allogeneic HSCT for MDS/AML, PAP, and/or recurrent infection (age at transplant 15-49 years). The proportion surviving without HSCT was 89% by age 20, 76% by age 30, and 53% by age 40. Post-transplant survival was 72% by 1 year, 65% by 2 years, and 54% by 4 years after HSCT.

Clinical Features

Hematologic

Hemograms and lymphocyte phenotyping from initial NIH evaluation were available for 55 and 51 patients, respectively. B lymphocytopenia (86%), NK lymphocytopenia (82%), and monocytopenia (78%) were common and profound (Figure S1). CD4 lymphocytopenia (51%)
and neutropenia (47%) were also common but generally less marked. Pancytopenia was observed in the setting of MDS and monocytosis in patients with CMML. Of 50 patients who underwent bone marrow biopsies, 42 (84%) met diagnostic criteria for MDS with most marrows demonstrating multilineage dysplasia and <5% blasts (Table S2). In contrast to de novo MDS, which is typically characterized by marrow hypercellularity, most marrows were hypocellular for age with increased reticulin fibrosis (Figure 3A). Atypical megakaryocytes were identified in 92%, even in those without overt MDS (Figure 3B-D). Forty-eight percent had abnormal bone marrow cytogenetics, including trisomy 8 in 24% and monosomy 7 in 16%. Fourteen percent evolved into AML and 8% met criteria for CMML. Increased populations of CD3+CD8+CD56+ large granular lymphocytes were detected in 16%.

Infectious

Major infections afflicted 82% of patients (Table S3). Severe viral infections were the most common (70%), particularly HPV (63%) presenting with recalcitrant warts, condylomata, and/or dysplasia. Severe herpesvirus infections were present in 35%; recurrent herpes stomatitis, esophagitis, or genital infection 16%; severe varicella 11%; persistent EBV viremia 11%; and cytomegalovirus pneumonia or disseminated disease 4%. Two patients developed severe cutaneous molluscum contagiosum infections. Disseminated NTM infections were present in 53% of patients, but this may reflect a referral bias of our institution (Figure 3E-F). NTM infections included both slow growing (Mycobacterium avium complex, M. kansasii, M. szulgai) and rapid growing mycobacteria (M. fortuitum, M. abscessus, M. chelonae). Other severe bacterial infections were observed in 49%; bacteremia 21%, skin and soft tissue infection 19%,
pneumonia 14%, and colitis 9%. *Clostridium difficile* infections were particularly severe, including 1 fatality, 1 colectomy in childhood, and 2 cases of toxic megacolon. Severe fungal infections were observed in 16%: invasive aspergillosis 9%, disseminated histoplasmosis 5%, and recalcitrant mucosal candidiasis 5%. Following HSCT, no patients had reactivation of pre-existing NTM or fungal infections, and most patients demonstrated marked clearing of HPV lesions.25

**Pulmonary**

Biopsy-proven PAP was identified in 10 adults (18%), all of whom lacked anti-GM-CSF autoantibodies and were refractory to subcutaneous and inhaled GM-CSF therapy (Figure 3G-H). Five patients with PAP (9%) developed subsequent pulmonary arterial hypertension (PAH). Following HSCT, all patients with PAP/PAH demonstrated significant improvement in pulmonary function including weaning from supplemental oxygen and medications for PAH.25 Even among patients without PAP, pulmonary function and radiographic appearance were typically abnormal (Table S4). Of 38 patients who received pulmonary function testing, 30 (79%) had diffusion defects and 24 (63%) had ventilatory defects, including obstruction (39%), restriction (11%), and mixed patterns (13%). Structural abnormalities identified on chest CT included nodules, reticular opacities, ground glass opacities, subpleural blebbing, “crazy paving”, and paraseptal emphysema. Surprisingly, despite profound circulating monocytopenia, alveolar macrophages were abundant in bronchoalveolar lavage fluids.
**Dermatologic**

Dermatologic disease was often an indicator of underlying infection or malignancy (Table S5). Recurrent warts were present in the majority of patients (53%) and were often the initial manifestation of disease. Common and flat warts were present on the extremities (39%) and/or genitalia (32%) (Figure 3I-J) and were typically refractory to cauterization, cryotherapy, topical imiquimod, 5-fluorouracil, corticosteroids, and laser therapies. Biopsy-proven panniculitis/erythema nodosum was seen in 30% and was commonly associated with underlying NTM or fungal infection (Figure 3K). Skin cancers, including basal cell carcinoma, squamous cell carcinoma (SCC), and malignant melanoma, developed in 6 patients (11%), half of whom had multiple cancers. Sweet’s syndrome was seen in 3 patients in the setting of underlying MDS/AML.

**Neoplastic**

The majority of solid tumors were related to underlying viral infections. HPV-related dysplasia was observed in 35% and ranged from squamous intraepithelial lesions/Bowenoid papulosis to invasive SCC (Table S6). Most patients with genital dysplasia required recurrent laser ablative surgeries, loop electrosurgical excision procedures, or radical surgery. HPV-related head and neck cancers were also observed, including SCC of the oral commissure and oropharynx. Unusual EBV+ mesenchymal tumors were observed in two patients: one with widely disseminated EBV+ leiomyosarcomata, and another with EBV+ spindle cell tumors of the liver (Figure 3L-M). Notably, of 18 women ≥35 years of age at last follow-up, 4 (22%)...
developed breast cancers, one of whom had a *BRCA2* mutation. Other solid tumors included adenocarcinoma of the pancreas, renal cell carcinoma, and a locally invasive desmoid tumor of the chest wall.

**Cardiovascular & Lymphatic**

Chronic lymphedema was present in 6 patients (11%) and could involve the unilateral or bilateral lower extremities and genitalia (Figure 3N). Severe cases developed in infancy or childhood, were complicated by recurrent cellulitis, and often required lymphatic surgeries. Other cases only first manifested in adolescence or adulthood and were managed conservatively with compression stockings. Venous thromboses, including deep vein thrombosis, pulmonary embolism, portal vein thrombosis, and catheter-related thrombosis, were observed in 14 patients (25%), half of whom had multiple events (Table S7). Lupus anticoagulant was transiently positive (i.e. negative on repeat testing) in 7/14 patients with thrombotic events. One patient had mutations in Factor V Leiden and MTHFR. Two patients developed culture-negative endocarditis (one post-transplant) with demonstrated mitral valve vegetations and subsequent embolic strokes (Figure 3O). Thrombotic events also developed post-transplant in two patients.

**Other Clinical Findings**

Other associated clinical findings included sensorineural hearing loss, miscarriage, and idiopathic hypothyroidism. Of 25 patients with age-adjusted audiograms, 19 (76%) had mild to severe sensorineural hearing loss typically involving the high frequency range (Table S8).
However, prior aminoglycoside exposure for disseminated NTM infections may confound this relationship. Follow-up audiograms on 9 patients found 5 with progressive deterioration. Of 43 known pregnancies in our cohort, 14 (33%) resulted in miscarriage (Table S9). Notably, all women who miscarried demonstrated marked NK lymphocytopenia. Idiopathic hypothyroidism requiring thyroid hormone replacement was observed in 14%.

**Phenotype Clustering**

Certain clinical phenotypes were frequently encountered in combination with others (Figure S2). For example, all patients with PAP had comorbid immunodeficiency with recurrent viral and NTM infections. Nine out of 10 patients with PAP also had comorbid MDS. The majority of patients with MDS also had a history of recurrent infections with viruses, NTM, or both. However, 6 patients developed MDS without associated immunodeficiency. Of 6 patients with lymphedema, 4 had recurrent warts and/or MDS characteristic of Emberger syndrome. However, 2 patients with lymphedema had no other identifiable clinical manifestations. Four patients had a history of severe viral infections with no other identifiable manifestations.

**Phenotypic Variation Within Kindreds**

Certain phenotypes tended to cluster within families (Figure S3). For example, in kindred 40, the proband and his siblings each developed warts and MDS with trisomy 8. Their family history was also significant for multiple family members with MDS/AML and HPV-related cancer. Interestingly, however, their father carries the exact same mutation but is
clinically asymptomatic. Other families demonstrated marked clinical heterogeneity. For example, the proband in kindred 4 first presented at age 26 with disseminated *M. avium* complex and went on to develop disseminated histoplasmosis, MDS, PAP/PAH, and multiple malignancies of the skin and breast. Her sister first presented at age 19 with genital HPV infection and later developed MDS with monosomy 7 requiring HSCT. However, in contrast to their severe presentations in adolescence and early adulthood, their father was clinically silent until age 78 at which time he developed monocytosis and bone marrow changes consistent with CMML. The proband’s two children who carry the same genetic mutation are currently asymptomatic except for lower extremity lymphedema in her son.

**Genotype-Phenotype Associations**

We compared Kaplan-Meier disease-free survival curves among patients with missense, null, and regulatory mutations for the following endpoints: onset of illness, MDS/AML, severe viral infection, NTM/fungal infection, PAP, lymphedema, survival, and transplant-free survival (Figure 4). Severe viral infections were more common and had earlier-onset in those with null mutations (p=0.038). Lymphedema was only observed in those with null or regulatory mutations, but not in those with missense mutations (exact log rank p=0.006). No other endpoints differed significantly among mutation groups.
Blood Counts and Disease Severity

We developed a scoring system to estimate disease severity or progression based on the number of major complications. The following were considered as major complications given their associated morbidity and mortality: MDS/AML, severe viral infection, NTM/fungal infection, and PAP. Initial presenting B cell, NK cell, monocyte, and CD4 T cell counts were compared among individuals with 0, 1, 2, 3 or 4 of these complications at the time of clinical presentation (Figure 5). Among individuals with no major complications, median blood counts were within the normal range. However, as complications mounted, median blood counts declined significantly for all cell lines (p<0.001).

Discussion

We systematically reviewed 57 patients with GATA2 deficiency and identified a broad phenotype encompassing MDS/AML, immunodeficiency, pulmonary disease, vascular/lymphatic dysfunction, and additional clinical associations (Figure 6). Recognition of these clinical features is critical to facilitate early genetic diagnosis and to direct appropriate clinical management, preventive care, and family screening.

GATA2 belongs to a family of zinc finger transcription factors that are critical regulators of gene expression in hematopoietic cells. While GATA1 functions primarily in erythropoiesis and megakaryopoiesis and GATA3 in T cell lymphopoiesis, GATA2 is particularly critical for the genesis and function of hematopoietic stem and progenitor cells and thus all subsequent blood cell lineages. GATA2 haploinsufficiency perturbs normal hematopoietic stem cell homeostasis in murine models. The high prevalence of cytopenias and
bone marrow failure in our cohort supports a similar defect in human GATA2 haploinsufficiency. It remains unclear why monocytes, dendritic cells, B cells, and NK cells are preferentially depleted, but profound cytopenias of all these lineages appears to be highly specific for GATA2 deficiency.

The presence of neutropenia and CD4 lymphocytopenia in roughly half of our patients suggests that idiopathic neutropenia and idiopathic CD4 lymphocytopenia may be important populations to screen. Indeed, targeted screening of 14 probands from the French Severe Chronic Neutropenia Registry identified 6 with GATA2 mutations.8 Pediatric MDS represents another critical population to screen, with GATA2 mutations identified in 8/51 patients (16%) with MDS and monosomy 7 in the European Working Group of MDS and JMML in Childhood.34 In the absence of cytogenetic abnormalities or overt dysplasia, the hypoplastic marrow of GATA2 deficiency may also be initially diagnosed as aplastic anemia.9,35 Vigilant bone marrow monitoring is prudent in patients with GATA2 deficiency, as the development of excess blasts or abnormal cytogenetics may influence the timing of HSCT. Somatic mutations in ASXL1 have also been implicated in leukemic transformation and might accelerate the timing for HSCT.36,37 Importantly, family members should be screened for GATA2 mutations prior to serving as bone marrow donors for an affected proband.

Susceptibility to severe viral infections in GATA2 deficiency likely reflects the deficit and/or dysfunction of NK cells, which are critical for anti-viral immunity.11 Systemic interferon alpha improves NK cell cytotoxicity in vitro and may be beneficial for patients with refractory HPV or herpesvirus infections.11 Likewise, dysfunction of monocytes/macrophages may predispose to intramacrophagic pathogens, such as NTM and Histoplasma capsulatum. Given their marked susceptibility to HPV and NTM, we empirically recommend early HPV vaccination...
An association between PAP and hematologic malignancies has long been appreciated\textsuperscript{17}; GATA2 deficiency unifies both entities and may help explain this association. Given the abundance of alveolar macrophages in lavage fluids from our patients, PAP in GATA2 deficiency must reflect alveolar macrophage dysfunction rather than a quantitative deficit. Indeed GATA2 has been shown to regulate alveolar macrophage phagocytosis.\textsuperscript{38} We identified PAH only in the setting of PAP, likely reflecting that progressive lung disease can drive arterial dysfunction. Patients with PAP who lack anti-GM-CSF autoantibodies should be screened for GATA2 mutations, particularly those with comorbid immunodeficiency and/or MDS/AML. HSCT markedly improved pulmonary function in patients with PAP, presumably via reconstitution of the alveolar macrophage compartment.\textsuperscript{25} The high prevalence of diffusion and ventilatory defects in our cohort may reflect alveolar filling, emphysematous changes, fibrosis, or bronchiectasis secondary to recurrent pulmonary infection.

Careful dermatologic and gynecologic monitoring are prudent in view of the high prevalence of dysplastic HPV lesions and skin cancers. Panniculitis and Sweet’s syndrome should raise clinical suspicion for underlying infection or malignancy, respectively. EBV+ mesenchymal tumors, as seen in advanced HIV/AIDS, indicate a profound host defect that may also be NK cell mediated. The relationship of GATA2 deficiency to other cancers is less clear, but the development of multiple breast cancers and other solid tumors is concerning, particularly in light of the putative role of NK cells in tumor surveillance and the high frequency of somatic GATA3 mutations in human breast cancers.\textsuperscript{39,40}
GATA2 was first identified as an endothelial transcription factor and its haploinsufficiency perturbs normal vascular development.\textsuperscript{20,41-43} Thrombotic events in one quarter of patients suggest an association between GATA2 deficiency and coagulopathy, more likely due to endothelial rather than coagulation abnormalities. In the lymphatic endothelium, GATA2 is involved in the development of lymphatic valves, deficiency of which likely underlies lymphedema.\textsuperscript{19} Endocarditis may reflect GATA2 haploinsufficiency in the endocardium and/or impaired overlap with GATA4 function, which is involved in cardiogenesis.\textsuperscript{44-46} Importantly, the development of thrombotic events and endocarditis in the early post-transplant period suggests that HSCT may not immediately ameliorate cardiovascular risks.

GATA2 is required for development of the semicircular canals and perilymphatic space of the vestibular system\textsuperscript{47}, and its haploinsufficiency may lead to hearing loss. GATA2 also interacts with GATA3, haploinsufficiency of which is associated with congenital deafness.\textsuperscript{48} Although hearing loss is confounded by aminoglycoside exposure in our cohort, it was also present in several patients without prior exposure. GATA2 deficient patients are likely more susceptible to ototoxic insults, which should inform the use of aminoglycosides in these patients.

The 33\% prevalence of miscarriage in our cohort exceeds the 15-20\% rate suspected in the general population.\textsuperscript{49} Dysfunction of uterine NK cells, which regulate trophoblast invasion and uterine spiral artery angiogenesis, may be one potential mechanism.\textsuperscript{50,51} GATA2 expression may also be impaired in trophoblast giant cells, which are critical for embryonic implantation and placental function.\textsuperscript{52-54} Vascular dysfunction may also contribute to miscarriage.

Hypothyroidism affected 14\%, which may reflect GATA2-mediated transcriptional activation of the thyrotropin beta gene.\textsuperscript{55} However, thyroid dysfunction is relatively common and further studies are necessary to determine the mechanism of this association.
The marked clinical heterogeneity of GATA2 deficiency is multifactorial. Severe viral infections and lymphedema were more common in individuals with null mutations, suggesting an effect of genotype on clinical phenotype. This may explain why the great majority of patients previously reported with Emberger syndrome have nonsense and frameshift mutations or large gene deletions predicted to result in null alleles.\(^7,19\) However, in our series, even among individuals within the same family there was considerable variation in clinical phenotype, largely negating a generalized genotype-phenotype association. Contiguous gene deletion syndromes may be a somewhat separate entity, as described.\(^19\) The distinction between the roles of background genetics and environmental exposure will await work in animal models. Clearly environmental factors, such as exposure to endemic mycoses and mycobacteria, influence certain infections. ASXL1 is emerging as an important “second hit” associated with transformation from MDS to overt leukemia.\(^36,37\)

The clinical spectrum of our cohort is necessarily limited by the referral bias of our center. Many probands were recruited through natural history protocols for the study of NTM infections or primary immune deficiency. However, the addition of 17 relatives identified through family screening was an important opportunity to obtain a more complete picture of the clinical phenotype. Our cohort is also likely underrepresented for those with acute and fatal presentations, representing possible length bias and causing an optimistic bias in survival estimates.\(^56\) These potential biases can only be addressed through broader testing for GATA2 mutations and prospective study in additional cohorts.

Our pathophysiologic understanding of GATA2 deficiency is primitive but rapidly growing. GATA2 interacts closely with GATA1 in the development of the hematopoietic compartment and GATA3 in trophoblastic development.\(^26\) It underlies the development of
circulating myeloid dendritic cells but seems to permit infections or cancer only after a prolonged latency. Finally, while haploinsufficiency is clearly pathologic and leads to myeloid leukemia in a substantial number of cases, gain of function mutations are also associated with myeloid leukemias, indicating that the proper regulation of GATA2 is both critical and its range quite narrow. It is now clear that GATA2 deficiency is a protean disorder that unites susceptibility to immunodeficiency, bone marrow failure, pulmonary disease and vascular/lymphatic dysfunction. It causes cases of familial and sporadic MDS/AML, aplastic anemia, neutropenia, primary immune deficiency, and lymphedema, the full extent of which is yet to be determined. Early genetic diagnosis is essential for appropriate disease management, prophylaxis, transplantation, and family screening.
Acknowledgements:
This work was supported by the Division of Intramural Research, National Institute of Allergy and Infectious Diseases and National Cancer Institute, National Institutes of Health. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

Authorship:


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

Correspondence: Steven M. Holland, CRC B3-4141, MSC 1684, Bethesda, MD 20892-1684; email: smh@nih.gov.

Footnotes:

M.A.S. and L.A.S. contributed equally to this study.
References:


52. Ng YK, George KM, Engel JD, Linzer DI. GATA factor activity is required for the
trophoblast-specific transcriptional regulation of the mouse placental lactogen I gene.

53. Ma GT, Roth ME, Groskopf JC, et al. GATA-2 and GATA-3 regulate trophoblast-

54. Bai H, Sakurai T, Someya Y, et al. Regulation of trophoblast-specific factors by GATA2

55. Ohba K, Sasaki S, Matsushita A, et al. GATA2 mediates thyrotropin-releasing hormone-


57. Collin M, Bigley V, Haniffa M, Hambleton S. Human dendritic cell deficiency: the

myeloid transformation of chronic myeloid leukemia. *Proc Natl Acad Sci U S A.*
Table 1. Genetic and clinical features of 57 patients with GATA2 deficiency

<table>
<thead>
<tr>
<th>Patient</th>
<th>Sex</th>
<th>Mutation&lt;sup&gt;1&lt;/sup&gt;</th>
<th>Mutation Group&lt;sup&gt;2&lt;/sup&gt;</th>
<th>Age at first illness&lt;sup&gt;3&lt;/sup&gt;</th>
<th>Bone marrow pathology</th>
<th>Severe viral infections</th>
<th>Mycobacterial, fungal, and other invasive infections</th>
<th>Other clinical features</th>
<th>Age at first and last NIH visit</th>
<th>Previously cited in:</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.I.1</td>
<td>F</td>
<td>Uniallelic expression</td>
<td>-------</td>
<td>U</td>
<td>51</td>
<td>MDS, LGL</td>
<td>HPV</td>
<td>M. kansasii</td>
<td>First: 53</td>
<td>Last: 64&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td>23.I.3</td>
<td>M</td>
<td>Uniallelic expression</td>
<td>-------</td>
<td>U</td>
<td>21</td>
<td>MDS</td>
<td>MAC</td>
<td></td>
<td>First: 22</td>
<td>Last: 26&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td>29.I.1</td>
<td>F</td>
<td>Uniallelic expression</td>
<td>-------</td>
<td>U</td>
<td>45</td>
<td>MDS/AML, +8</td>
<td>HPV</td>
<td></td>
<td>First: 45</td>
<td>Last: 48&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>13.I.1&lt;sup&gt;+&lt;/sup&gt;</td>
<td>F</td>
<td>1-200_871+527del2033ins7</td>
<td>M1del290</td>
<td>N</td>
<td>40</td>
<td>HPV</td>
<td></td>
<td>Lymphedema</td>
<td>First: 60</td>
<td>Last: 60&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td>13.I.1&lt;sup&gt;+&lt;/sup&gt;</td>
<td>M</td>
<td>1-200_871+527del2033ins7</td>
<td>M1del290</td>
<td>N</td>
<td>33</td>
<td>MDS/AML, -7</td>
<td>HPV, HSV</td>
<td>Histoplasmosis, MAC, Neosartorya udagawaiae</td>
<td>First: 33</td>
<td>Last: 34&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td>8.I.1</td>
<td>M</td>
<td>243_244delAinsGC</td>
<td>G811fs</td>
<td>N</td>
<td>9</td>
<td>MDS, smoldering myeloma</td>
<td>HPV</td>
<td>MAC, soft tissue aspergillosis</td>
<td>First: 27</td>
<td>Last: 36&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>26.I.1</td>
<td>F</td>
<td>302delG</td>
<td>G101AfsX16</td>
<td>N</td>
<td>15</td>
<td>MDS/AML, CMML</td>
<td>HSV, CMV</td>
<td></td>
<td>First: 21</td>
<td>Last: 22&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td>38.I.1</td>
<td>F</td>
<td>417dupT</td>
<td>V140CfsX44</td>
<td>N</td>
<td>7</td>
<td>HPV</td>
<td></td>
<td></td>
<td>First: 24</td>
<td>Last: 24&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>27.I.1</td>
<td>F</td>
<td>586_593delG</td>
<td>G199LfsX21</td>
<td>N</td>
<td>31</td>
<td>HPV</td>
<td></td>
<td>M. kansasii</td>
<td>First: 46</td>
<td>Last: 46&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>20.I.3</td>
<td>M</td>
<td>769_778dup</td>
<td>Y260fsX24</td>
<td>N</td>
<td>13</td>
<td>MDS, -7</td>
<td>HPV, MCV</td>
<td>MAC, severe C. difficile</td>
<td>First: 13</td>
<td>Last: 17&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>22.I.1</td>
<td>F</td>
<td>941_951dup</td>
<td>A318fsX12</td>
<td>N</td>
<td>26</td>
<td>MDS</td>
<td>HPV</td>
<td>M. kansasii</td>
<td>First: 26&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Last: 27&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>39.I.1</td>
<td>F</td>
<td>988C&gt;T</td>
<td>R330X</td>
<td>N</td>
<td>6</td>
<td>MDS, +8</td>
<td>HPV, HSV, EBV</td>
<td>Salmonella enterocolitis, Neisseria meningitis</td>
<td>First: 25</td>
<td>Last: 25&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>39.I.2</td>
<td>F</td>
<td>988C&gt;T</td>
<td>R330X</td>
<td>N</td>
<td>6</td>
<td>MDS, +8</td>
<td>HPV, HSV, EBV</td>
<td>Lymphedema</td>
<td>First: 25</td>
<td>Last: 25&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>32.I.1</td>
<td>M</td>
<td>989_992dup</td>
<td>L332TfsX53</td>
<td>N</td>
<td>5 mo.</td>
<td>MDS/AML, -7 +8</td>
<td></td>
<td>Lympedema, HT</td>
<td>First: 17</td>
<td>Last: 19&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>41.I.1</td>
<td>F</td>
<td>1009C&gt;T</td>
<td>R337X</td>
<td>N</td>
<td>9</td>
<td>MDS</td>
<td>HPV, HSV</td>
<td>Lympedema, HT</td>
<td>First: 44</td>
<td>Last: 44&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>6.I.1&lt;sup&gt;+&lt;/sup&gt;</td>
<td>M</td>
<td>1017+512del28</td>
<td>-------</td>
<td>R</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td>First: 66</td>
<td>Last: 66&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>6.II.1</td>
<td>F</td>
<td>1017+512del28</td>
<td>-------</td>
<td>R</td>
<td>24</td>
<td>MDS</td>
<td>HPV</td>
<td></td>
<td>First: 27</td>
<td>Last: 35&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>4.II.1&lt;sup&gt;+&lt;/sup&gt;</td>
<td>M</td>
<td>1017+572C&gt;T</td>
<td>-------</td>
<td>R</td>
<td>78</td>
<td>CMML, LGL</td>
<td></td>
<td>Group C streptococcal bacteremia/osteomyelitis, M. tuberculosis</td>
<td>First: 72</td>
<td>Last: 79&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>4.II.1</td>
<td>F</td>
<td>1017+572C&gt;T</td>
<td>-------</td>
<td>R</td>
<td>26</td>
<td>MDS, LGL</td>
<td>HPV, CMV</td>
<td>MAC, disseminated histoplasmosis, Serratia pneumonia</td>
<td>First: 38</td>
<td>Last: 53&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>4.II.1&lt;sup&gt;+&lt;/sup&gt;</td>
<td>F</td>
<td>1017+572C&gt;T</td>
<td>-------</td>
<td>R</td>
<td>19</td>
<td>MDS, -7</td>
<td>HPV</td>
<td></td>
<td>First: 30</td>
<td>Last: 42&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>4.III.1&lt;sup&gt;+&lt;/sup&gt;</td>
<td>M</td>
<td>1017+572C&gt;T</td>
<td>-------</td>
<td>R</td>
<td>30</td>
<td></td>
<td></td>
<td></td>
<td>First: 23</td>
<td>Last: 30&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>4.III.2&lt;sup&gt;+&lt;/sup&gt;</td>
<td>F</td>
<td>1017+572C&gt;T</td>
<td>-------</td>
<td>R</td>
<td>28</td>
<td></td>
<td></td>
<td></td>
<td>First: 21</td>
<td>Last: 28&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>11.II.1</td>
<td>F</td>
<td>1017+572C&gt;T</td>
<td>-------</td>
<td>R</td>
<td>9</td>
<td>HPV, VZV</td>
<td>Recurrent mucosal</td>
<td></td>
<td>First: 30</td>
<td>Last: 31&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>Date</td>
<td>Sample Type</td>
<td>Mutation Type</td>
<td>Mutation Description</td>
<td>Disease</td>
<td>Age at Diagnosis</td>
<td>NGS Test</td>
<td>Other Tests</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>-------------</td>
<td>---------------</td>
<td>----------------------</td>
<td>----------------------------------</td>
<td>-----------------</td>
<td>----------</td>
<td>-------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25.II.1</td>
<td>M</td>
<td>1017+572C&gt;T</td>
<td>MDS</td>
<td>candidiasis</td>
<td>34</td>
<td>HPV, MAC</td>
<td>M. kansasii</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24.I.1</td>
<td>F</td>
<td>1018-1G&gt;A</td>
<td>MDS, LGL</td>
<td>HPV, VZV</td>
<td>43</td>
<td>MAC</td>
<td>M. abscessus</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.II.1</td>
<td>F</td>
<td>T354M</td>
<td>MDS/AML, LGL</td>
<td>HPV, MAC, M. abscessus</td>
<td>19</td>
<td>PAP, DVT/PE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17.I.1</td>
<td>M</td>
<td>T354M</td>
<td>MDS, +8</td>
<td>HPV, MAC</td>
<td>28</td>
<td>PAP</td>
<td>M. cheloneae</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17.II.2</td>
<td>M</td>
<td>T354M</td>
<td>--</td>
<td>MAC</td>
<td>--</td>
<td>DVT/PE</td>
<td>M. cheloneae</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19.II.1</td>
<td>M</td>
<td>T354M</td>
<td>EBV</td>
<td>M. cheloneae</td>
<td>20</td>
<td>22*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35.III.3</td>
<td>M</td>
<td>T354M</td>
<td>MDS, +8</td>
<td>Necrotizing fascitis, granulomatous lymphadenitis</td>
<td>20</td>
<td>22*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>37.I.1</td>
<td>F</td>
<td>1081C&gt;T</td>
<td>MDS</td>
<td>M. kansasii, MAC</td>
<td>26</td>
<td>MAC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12.I.1</td>
<td>M</td>
<td>1083_1094del12</td>
<td>MDS, -7</td>
<td>HPV, HSV, HCV</td>
<td>13</td>
<td>VZV, HPV</td>
<td>MAC, severe C. difficile</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>33.III.1</td>
<td>F</td>
<td>1099insG</td>
<td>D367GfsX</td>
<td>MAC, severe C. difficile</td>
<td>17</td>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>33.III.3</td>
<td>F</td>
<td>1099insG</td>
<td>D367GfsX</td>
<td>MAC, severe C. difficile</td>
<td>10</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10.I.1</td>
<td>F</td>
<td>1113C&gt;G</td>
<td>MDS/AML, -7</td>
<td>VZV, HPV</td>
<td>12</td>
<td>VZV, HPV</td>
<td>MAC, severe C. difficile</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>34.I.1*</td>
<td>F</td>
<td>1116_1130del15</td>
<td>C373del5</td>
<td>M. kansasii, severe C. difficile</td>
<td>50</td>
<td>MAC</td>
<td>Breast cancer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>34.I.1</td>
<td>M</td>
<td>1116_1130del15</td>
<td>MDS</td>
<td>MAC, severe C. difficile</td>
<td>14</td>
<td>MAC</td>
<td>Breast cancer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>34.II.2*</td>
<td>M</td>
<td>1116_1130del15</td>
<td>MDS</td>
<td>MAC, severe C. difficile</td>
<td>13</td>
<td>13</td>
<td>Breast cancer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30.I.1</td>
<td>M</td>
<td>1163T&gt;C</td>
<td>M388T</td>
<td>HPV</td>
<td>65</td>
<td>HPV</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30.II.1</td>
<td>F</td>
<td>1163T&gt;C</td>
<td>M388T</td>
<td>HPV, HSV</td>
<td>27</td>
<td>HPV</td>
<td>MAC, severe C. difficile</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30.II.4*</td>
<td>F</td>
<td>1163T&gt;C</td>
<td>M388T</td>
<td>HPV</td>
<td>30</td>
<td>HPV</td>
<td>MAC, severe C. difficile</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15.I.1</td>
<td>F</td>
<td>1186C&gt;T</td>
<td>R396W</td>
<td>HSV</td>
<td>7</td>
<td>HSV</td>
<td>MAC, E. histolytica</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>42.I.1</td>
<td>M</td>
<td>1186C&gt;T</td>
<td>R396W</td>
<td>HSV</td>
<td>18</td>
<td>HSV</td>
<td>MAC, E. histolytica</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14.I.1</td>
<td>F</td>
<td>1187G&gt;A</td>
<td>R396W</td>
<td>EBV, HSV</td>
<td>7</td>
<td>EBV</td>
<td>MAC, E. histolytica</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18.I.2</td>
<td>F</td>
<td>1187G&gt;A</td>
<td>R396W</td>
<td>EBV, HSV</td>
<td>16</td>
<td>EBV</td>
<td>MAC, E. histolytica</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40.I.1*</td>
<td>M</td>
<td>1187G&gt;A</td>
<td>R396W</td>
<td>EBV, HSV</td>
<td>18</td>
<td>EBV</td>
<td>MAC, E. histolytica</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40.I.1</td>
<td>M</td>
<td>1187G&gt;A</td>
<td>R396Q</td>
<td>M. abscessus, cardiosclerosis</td>
<td>16</td>
<td>M. abscessus, cardiosclerosis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40.II.1</td>
<td>M</td>
<td>1187G&gt;A</td>
<td>R396Q</td>
<td>HPV, VZV</td>
<td>18</td>
<td>HPV</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40.II.2*</td>
<td>M</td>
<td>1187G&gt;A</td>
<td>R396Q</td>
<td>HPV, VZV</td>
<td>16</td>
<td>HPV</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40.III.3*</td>
<td>M</td>
<td>1187G&gt;A</td>
<td>R396Q</td>
<td>HPV, VZV</td>
<td>14</td>
<td>HPV</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31.II.1</td>
<td>M</td>
<td>1187G&gt;A</td>
<td>R396Q</td>
<td>MDS</td>
<td>27</td>
<td>MDS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Granulomatous lymphadenitis, invasive aspergillosis</td>
<td>27</td>
<td>MDS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sarcoïdosis-like pulmonary process</td>
<td>30</td>
<td>32</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21.II.1</td>
<td>M</td>
<td>1192C&gt;T</td>
<td>R398W</td>
<td>M</td>
<td>32</td>
<td>MDS</td>
<td>HPV, MCV</td>
<td>MAC</td>
<td>29</td>
<td>29</td>
</tr>
<tr>
<td>21.II.1</td>
<td>M</td>
<td>1192C&gt;T</td>
<td>R398W</td>
<td>M</td>
<td>32</td>
<td>MDS</td>
<td>HPV, MCV</td>
<td>MAC</td>
<td>29</td>
<td>29</td>
</tr>
</tbody>
</table>

Abbreviations: AML = Acute myeloid leukemia, BOOP = Bronchiolitis obliterans organizing pneumonia, CMML = Chronic myelomonocytic leukemia, CMV = Cytomegalovirus, CVA = Cerebrovascular accident, DVT = Deep vein thrombosis, EBV = Epstein-Barr virus, HCV = Hepatitis C virus, HPV = Human papillomavirus, HSV = Herpes simplex virus, HT = Hypothyroidism, LGL = Large granular lymphocytosis, MAC = *Mycobacterium avium* complex, MCV = Molluscum contagiosum virus, MDS = Myelodysplastic syndromes, NTM = Nontuberculous mycobacteria, PAH = Pulmonary arterial hypertension, PAP = Pulmonary alveolar proteinosis, PE = Pulmonary embolism, PVT = Portal vein thrombosis, RCC = Renal cell carcinoma, VZV = Varicella zoster virus, -7 = Monosomy 7, +8 = Trisomy 8

1. All mutations refer to isoform NM_001145661, protein changes reference NP_116027

2. Mutation groups are defined as follows:

   M = missense, includes all missense mutations, in-frame deletions, and insertions/deletions predicted to escape nonsense mediated decay

   N = null, includes insertions/deletions and nonsense mutations predicted to result in nonsense mediated decay and no protein product

   R = regulatory, includes mutations in the intron 5 enhancer with demonstrated reduced expression of wild-type protein

   U = uniallelic, includes undefined mutations with reduced or absent transcription from one allele by cDNA analysis

3. First illness is defined as severe infection, MDS/AML, PAP, or lymphedema
Complication developed post-transplant

Deceased

Relative of proband
Figure Legend

Figure 1. GATA2 mutations. (A) Genomic organization of GATA2 showing the 2 5’-untranslated and 5 coding exons. Larger boxes represent coding regions. Insertion/deletion and nonsense mutations predicted to result in null alleles are shown above in red. Regulatory mutations in intron 5 are shown in blue. (B) Protein domains of GATA2, showing N- and C-terminal zinc fingers (ZF-1, ZF-2) and nuclear localization signal (N). (C) Missense mutations, in-frame deletions, and frameshift mutations predicted to escape nonsense mediated decay are shown in green. Superscript numerals indicate the number of independent mutations.

Figure 2. Onset of Illness, Survival and Transplantation. Kaplan-Meier curves are shown for onset of illness (first severe infection, myelodysplasia/leukemia, pulmonary alveolar proteinosis, or lymphedema), survival after onset of illness, overall survival, and transplant-free survival for 57 individuals. Cumulative incidence of transplantation is shown accounting for death as a competing event. Post-transplant survival is shown for 21 individuals who underwent transplantation. The number of individuals under follow-up is indicated under the X-axis. Confidence intervals of 95% are shaded in gray.

Figure 3. Clinical, Pathologic and Radiographic Features of GATA2 Deficiency. (A) Hypocellular marrow with trilineage hypoplasia (36-year old female), bone marrow core biopsy H&E stain. (B-D) Atypical megakaryocytes, small mononuclear (B) and dysplastic forms with separated nuclear lobes (B-D); taken from bone marrow core biopsies (B, C) and aspirate smear (D). (E) Cutaneous Mycobacterium abscessus infection and (F) skin biopsy AFB smear showing abundant mycobacteria. (G) CT scan of the chest showing the typical crazy paving pattern of pulmonary alveolar proteinosis and (H) lung biopsy H&E stain showing alveolar filling with
lipoproteinaceous material. (I) Recalcitrant periungual warts, (J) perineal condyloma, and vulvar/anal intraepithelial neoplasia. (K) Panniculitis/erythema nodosum involving the anterior shins. (L) CT scan of the abdomen with multiple hypodense liver lesions (arrows) and (M) liver biopsy showing Epstein-Barr virus (EBV)-related spindle cell tumor with in situ hybridization of EBV-encoded mRNA using EBV probe anti-fluorescein antibody and bond polymer refine detection. (N) Unilateral lymphedema of the right lower extremity. (O) MRI of the brain with embolic infarcts in the right occipital and left parietal lobes (arrows) in the setting of culture-negative endocarditis.

**Figure 4. Genotype-Phenotype Associations.** Kaplan-Meier curves are shown for patients with missense (green, N=33), null (red, N=12), and regulatory mutations (blue, N=9). Patients with uniallelic expression (N=3) are not shown. Mutation group was associated with risk of viral infection (p=0.038) and lymphedema (exact log rank test p=0.006). No other endpoints differed by mutation group. Unless otherwise noted, p-values were calculated using Cox regression stratified on proband status, with robust standard errors to adjust for kindred clusters.

**Figure 5. Blood Counts and Disease Severity.** Absolute B cell, NK cell, monocyte, and CD4+ T cell count at first NIH visit for patients with 0, 1, 2, 3, or 4 of the following major complications: myelodysplasia/leukemia; severe viral infection; disseminated mycobacterial or fungal infection; and pulmonary alveolar proteinosis. Monocyte counts and lymphocyte phenotyping were available for 55 and 51 patients, respectively. Shaded regions indicate normal blood count ranges from the NIH Department of Laboratory Medicine: absolute CD20+ B cell count 60-539 cells/μL; absolute NK cell count 126-729 cells/μL; absolute monocyte count 0.24-0.86 x10^3 cells/μL; absolute CD4+ T cell count 362-1275 cells/μL. Diamonds indicate individuals with monocytosis
in the setting of chronic myelomonocytic leukemia. Horizontal bars indicate medians. Vertical bars indicate interquartile ranges. Blood counts were significantly associated with the number of major complications (p<0.001). P-values were obtained using linear regression with robust standard errors to adjust for kindred clusters. Analysis was done for the log-transformed cell counts, with offsets of half the lower-limit of detection for B cells, NK cells, and monocytes to account for those cases in which the numbers were 0.

**Figure 6. Clinical Features of GATA2 Deficiency by Organ System.** Common clinical findings are shown by affected organ system. Primary disease features are indicated in bold. Abbreviations: AML = Acute myeloid leukemia, HSV = Herpes simplex virus, MDS = Myelodysplastic syndromes, NTM = Nontuberculous mycobacteria.
Figure 2
Figure 3
Figure 4
Figure 5
GATA2 deficiency: a protean disorder of hematopoiesis, lymphatics and immunity