GATA2 deficiency and MDS/AML: Experimental strategies for disease modelling and future therapeutic prospects

Lili Kotmayer1 | Damia Romero-Moya2 | Oskar Marin-Bejar2 | Emilia Kozyra3,4 | Albert Català5,6 | Anna Bigas7,8 | Marcin W. Wlodarski3,9 | Csaba Bödör1 | Alessandra Giorgetti2,10,11

1HCEMM-SE Molecular Oncohematology Research Group, 1st Department of Pathology and Experimental Cancer Research, Semmelweis University, Budapest, Hungary
2Regenerative Medicine Program, Institut d’Investigació Biomèdica de Bellvitge (IDIBELL), Barcelona, Spain
3Division of Pediatric Hematology and Oncology, Department of Pediatrics and Adolescent Medicine, Medical Center, Faculty of Medicine, University of Freiburg, Freiburg, Germany
4Faculty of Biology, University of Freiburg, Freiburg, Germany
5Department of Hematology and Oncology, Institut de Recerca Sant Joan de Déu, Hospital Sant Joan de Deu, Barcelona, Spain
6Biomedical Network Research Centre on Rare Diseases, Instituto de Salud Carlos III, Madrid, Spain
7Institut Hospital del Mar d’Investigacions Mèdiques, CIBERONC, Hospital del Mar, Barcelona, Spain
8Josep Carreras Research Institute (IJC), Badalona, Barcelona, Spain
9Department of Hematology, St. Jude Children’s Research Hospital, Memphis, Tennessee, USA
10Fondazione Pisana Per la Scienza ONLUS (FPS), San Giuliano Terme, Italy
11Department of Pathology and Experimental Therapeutics, Faculty of Medicine and Health Sciences, Barcelona University, Barcelona, Spain

Correspondence
Alessandra Giorgetti, Bellvitge Institute for Biomedical Research (IDIBELL), Hospital Duran i Reynals-3rd floor Gran Via de l’Hospitalat, 199 08908 Hospitalet de Llobregat, Barcelona, Spain.
Email: agiorgetti@idibell.cat

Summary
The importance of predisposition to leukaemia in clinical practice is being increasingly recognized. This is emphasized by the establishment of a novel WHO disease category in 2016 called "myeloid neoplasms with germline predisposition". A major syndrome within this group is GATA2 deficiency, a heterogeneous immunodeficiency syndrome with a very high lifetime risk to develop myelodysplastic syndrome (MDS) and acute myeloid leukaemia (AML). GATA2 deficiency has been identified as the most common hereditary cause of MDS in adolescents with monosomy 7. Allogenic haematopoietic stem cell transplantation is the only curative option; however, chances of survival decrease with progression of immunodeficiency and MDS evolution. Penetrance and expressivity within families carrying GATA2 mutations is often variable, suggesting that cooperating extrinsic events are required to trigger the disease. Predictive tools are lacking, and intrafamilial heterogeneity is poorly understood; hence there is a clear unmet medical need. On behalf of the ERAPerMed GATA2 HuMo consortium, in this review we describe the genetic, clinical, and biological aspects of familial GATA2-related MDS, highlighting the importance of developing robust disease preclinical models to improve early detection and clinical decision-making of GATA2 carriers.

KEYWORDS
acute myeloid leukaemia, blood cancer, GATA2 deficiency, myelodysplastic syndromes
INTRODUCTION

Over a decade ago, the first germline pathogenic variants in the GATA2 gene with a predisposition to myelodysplastic syndrome (MDS) and acute myeloid leukaemia (AML) were reported by Hahn and colleagues. Since then, a number of germline GATA2 variants across individual cases and families with associated characteristic haematological and non-haematological phenotypes have been described. Together they define GATA2 deficiency syndrome, a rare autosomal dominant genetic disease with increased susceptibility to myeloid malignancies and other non-malignant symptoms. Clinical presentations of GATA2 deficiency syndrome may range from severe bone marrow failure (BMF), myeloid malignancies, immunodeficiencies and congenital syndromic features to a minority of asymptomatic cases due to partial or incomplete penetrance.

Despite the highly variable manifestations of GATA2 deficiency, associated phenotypes are now recognized as a single monogenic syndrome, introduced in the revised 2016 WHO classification of haematopoietic tumours as the novel category of “myeloid neoplasms with germline GATA2 mutations”. GATA2 deficiency syndrome encompasses cellular immunodeficiency disorders, namely MonoMac (mono-cytopenia and mycobacterial infections) and Emberger syndrome (myelodysplasia and lymphoedema), dendritic cell, monocyte, B- and NK lymphoid (DCML) deficiency, and myeloid malignancies, predominantly familial and primary paediatric MDS and AML. Germline GATA2 mutations have also been identified and linked to additional clinical manifestations such as chronic neutropenia, hypoplastic BMF, defects of megakaryopoiesis, B and NK cell deficiency, congenital auditory and neurological disorders (predominantly sensorineural deafness and ADHD) and recurrent miscarriages.

Despite the frequent and significant overlap among these clinical features and temporal changes in the clinical presentations over the disease course, no common pathognomonic GATA2 phenotype has been defined to date. It suggests that different unknown underlying co-operating factors (“stressor events”) in specific cellular contexts and processes are required for each GATA2-related clinical manifestation, as recently proposed by Homan and colleagues. Amalgamation of an impaired GATA2 function and the accumulation of stochastic biological or environmental stressors may (substantially) dysregulate a subset of GATA2-dependent processes, rattling normal cellular function. This mechanism has been proposed for haematopoietic stem cell (HSC) exhaustion and BMF due to recurrent infections, myeloid malignancies due to acquired somatic alterations, such as ASXL1 mutations or monosomy 7 and presumably lymphoedema due to mechanical or inflammatory stresses of lymphatic vessels and valves during development.

The ERAPerMed GATA2 HuMo consortium is an international collaborative network aiming for the precise understanding of GATA2 deficiency syndrome and subsequent myeloid transformation by combining clinical and genomic approaches with in vitro and in vivo models. In this review, we provide an overview of the molecular background of GATA2 deficiency with an insight into clinical implications, and current and future challenges in the management of patients with GATA2 predispositions.

ROLE OF GATA2 TRANSCRIPTION FACTOR IN HAEMATOPOIETIC AND NONHAEMATOPOIETIC TISSUES

GATA2 is a member of the GATA transcription factor family and is a master regulator of haematopoiesis and lymphatic angiogenesis by regulating target genes through occupying WGATAR DNA structure motifs with two zinc finger domains (ZF1 and ZF2). ZF1 controls protein–protein interactions, while ZF2 is critical for regulating transcription through protein-DNA binding. GATA2 has a major role in the early embryonic haematopoietic specification through endothelial haematopoietic transition (EHT). During postnatal haematopoiesis GATA2 is expressed at high levels in HSCs and early haematopoietic progenitors, and promotes myeloid-erythroid progenitor cell development (Figure 1A).

Furthermore, GATA2 is expressed in the lymphatic endothelium, instigating fetal lymphatic vessel valve development and normal valve function in adults, and its mechanosensitive expression in endothelial cell compartment also implicates GATA2 in angiogenesis and vascular integrity. Several single nucleotide polymorphisms (SNPs) of GATA2 have been associated with coronary artery disease implying its putative role in arterial development; however, underlying molecular mechanisms involving GATA2 are still poorly understood. Additionally, HSPCs and endothelial components, high GATA2 protein levels were observed during fetal neurodevelopment, as well as in selected types of differentiated cells, including neurons, androgen receptor-expressing and endocrine cells and macrophages.

EXISTING AND NEW EXPERIMENTAL MODELS TO STUDY GATA2 FUNCTION

The role of GATA2 gene as master regulator of haematopoietic system has been defined in 1994, when Orkin and colleagues showed that Gata2-KO mice died at 10.5 days post-coitum (dpc), due to the collapse of primitive and definitive haematopoiesis. Notably, mouse Gata2-null endothelial cells failed to produce HSCs because of impaired EHT. Formal demonstration of GATA2 function within HSC development, proliferation and survival in adult haematopoiesis has been gleaned using Gata2 +/- mice. Heterozygous Gata2 +/- mice survive and show a reduced number of HSCs both in the embryo and within adult BM, and Gata2+/- HSCs are qualitatively defective.
addition, in vivo studies at single cell level revealed a GATA2 dose-dependent mechanism regulating HSC identity.\textsuperscript{53–55} Furthermore, Gata2 haploinsufficiency also reduces granulocytes-macrophage function but no other myeloid committed progenitors.\textsuperscript{56} Of note, most recently, others and us have demonstrated a primary role of GATA2 in promoting EHT also in a human background using engineered pluripotent stem cells.\textsuperscript{28,30–35,57} Importantly, GATA2 is highly expressed in HSCs and its expression is tightly regulated by conserved cis-regulatory elements and other signalling pathways. Results of mouse studies based on targeted ablation of the far upstream Gata2–77 enhancer showed that it abrogates the multilineage differentiation potential of fetal liver progenitor cells without affecting HSC emergence,\textsuperscript{38} which is diminished by the intronic +9.5 enhancer deletion\textsuperscript{35,39,58,59} (Figure 2). Furthermore, Notch signals tightly regulate the precise level of expression of GATA2 required for generation of HSCs in the embryonic aorta.\textsuperscript{60,61}

Zebrafish is another valuable model that have been extensively used to study the function of GATA2 in haematopoiesis. A genome duplication event in early teleosts generated two Gata2 paralogs in the zebrafish genome: Gata2a and Gata2b.\textsuperscript{62} Previous studies have indicated that Gata2a is implicated in lymphovascular development with some expression in HSPCs in adult haematopoiesis,\textsuperscript{63–65} whereas Gata2b is mainly expressed in HE cells that give rise to HSCs in adult haematopoiesis.\textsuperscript{66} In addition, Gata2a mutant for the conserved enhancer i4, corresponding to the described +9.5 enhancer in the mouse Gata2 locus, revealed that Gata2a is required upstream of Gata2b, regulating Runx1 and Gata2b expression in the haemogenic endothelium\textsuperscript{67} (Figure 2). Recent transcriptomic analyses have shown that Gata2b loss impaired progression of the myeloid transcriptional programme while increasing the lymphoid programme in haematopoietic progenitors.\textsuperscript{58,69} Altogether these finding suggest that mammalian Gata2 functions are divided among gata2a and gata2b in zebrafish.

Overall, these findings have contributed greatly to our understanding of the role of GATA2 in embryonic and adult haematopoiesis. However, Gata2 +/- mice do not develop MDS and AML, probably due to the short lifespan of mice, the different external stressor factors or the absence of concurrent somatic lesions as observed in GATA2-related MDS patients. Furthermore, there is still lack of a knock-in mouse
model, which allows the overexpression of GATA2 mutant protein, mimicking the haematological manifestations observed in human GATA2 carriers. Therefore, there is a clear need for the development of better faithful mouse models of GATA2 deficiency, including patient-derived xenograft (PDX), to mimic better this complex disease. Nowadays, there are no PDX studies focusing on patients carrying germline GATA2 mutations. Generally, the establishment of human preclinical models by xenotransplantation of primary MDS cells in immunodeficient mice represent a bottleneck that hampers research of the mechanism understanding this disease. During the last decade, several laboratories have tested different approaches including intratibial injection in NSG (NOD-Scid-IL-2Rnull) or NSG-S (NSG/IL3/GM-CSF/hSCF), cotransplantation with patient-derived or normal mesenchymal stem cell (MSC). However, these methods showed low efficiency and poor engraftment rate, with a skew towards the lymphoid lineage. The biological reasoning behind the poor engraftment of MDS primary cells remains unclear and might be related to both the specific requirements for environmental factors and/or the intrinsic biological characteristic of individual samples. This limitation might be overcome by the generation of a humanized BM-MSC-derived ossicles (ectopic bone organoid composed of human bone, stroma, and haematopoietic tissue) that provides a human microenvironment. A more consistent outcome was achieved by using MISTRG mice, a humanized strains expressing human M-CSF, IL-3, GM-CSF, SIRP alpha and thrombopoietin at physiological level. The selection of the most suitable mouse model can significantly affect the outcome of the research findings. The pivotal question that remains unanswered is which mouse model should be used to study GATA2 deficiency. Another limitation for the development of mouse disease models is the genetic heterogeneity and the low prevalence of GATA2 deficiency that limits the access to patient primary cells. A
possible solution could be the generation of GATA2 deficiency preclinical mouse models using engineered human haematopoietic CD34+ cells (both from cord blood and mobilized peripheral blood). The ongoing improvement of CRISPR/Cas9 genomic engineering technologies enable the manipulation of genes in primary CD34+ cells, facilitating the development of preclinical models to study the molecular mechanisms and the clonal evolution of haematological diseases.26–78 (Figure 2). Furthermore, those models would be fundamental tools to identify genetic subtype that are sensitive or resistant to therapeutic agents. In parallel, the induced pluripotent stem cells (hiPSCs)-based approach represents an alternative cellular platform for mechanistic studies and drug screening.79–82 Recently, we reported the generation of human iPSC carrying two of the most recurrent GATA2 mutations in paediatric MDS patients.83 The availability of differentiated haematopoietic progenitor cells from iPSCs carrying GATA2 mutations will allow a comprehensive investigation of effects of these mutations as well as of secondary oncogenic somatic alterations (Figure 2).

Indeed, hiPSC model offers the possibility to study the stepwise progression of GATA2-related MDS in vitro through the sequential introduction of secondary oncogenic genetic lesions by CRISPR/Cas9 gene editing, as already shown for adult AML.84 Importantly, the iPSC-based approach allows the characterization of transcriptional programs driving specific stage malignant transition and the identification of possible prognostic markers for early therapeutic targeting (Figure 2).

**TYPES OF DISEASE-CAUSING GERMLINE GATA2 MUTATIONS**

Hereditary genetic alterations disrupting GATA2 expression and protein function are pivotal in leukemogenesis by sabotaging normal haematopoiesis leading to GATA2 deficiency and subsequent myeloid malignancies. Approximately one-third of all germline GATA2 variants are passed on through autosomal dominant inheritance (familial), while at least two-thirds of the mutations occur de novo.4,85,86 Mutational landscape of GATA2 deficiency includes a steadily increasing number of variants with approximately 500 published cases harbouring roughly 180 different familial or de novo germline mutations and partial or whole gene deletions.1,10,87–89 The majority of reported pathogenic or probably pathogenic germline variants are inactivating (null) or loss-of-function (LOF) resulting in haploinsufficiency (~60% of all GATA2 deficient cases) either through truncating (nonsense), frameshift or splicing mutations and deletions.19,22,27,88–91 These confirmed germline variants along with less common in-frame deletions and insertions are scattered throughout the coding region of GATA2.4,31

In addition to variants leading to premature translation termination prior or within ZF2, missense variants within ZF2 represent a large proportion of germline GATA2 variants (~30%) instigating LOF via disrupting protein–protein interactions21,92 and DNA binding.1,10,90,92 Monoallelic missense mutations of the ZFs have been reported mostly, albeit not exclusively, within the extended ZF2 domain with only a few confirmed cases of germline variants affecting ZF1. Intriguingly, unlike germline LOF GATA2 variants, somatic GATA2 mutations of adult AML affecting the ZFs occur with a preference for ZF1 and can exhibit gain-of-function (GOF) effects.22,93,94 Germline variants in ZF1, such as p.Ala318Thr,65 p.His313Tyr and p.Leu315Pro77 are recognized as LOF alterations and in most cases associated with paediatric MDS and accompanying symptoms of underlying GATA2 deficiency. Notably, a novel germline missense mutation in ZF1 (p.Asn317Ser) was recently reported by Rütsche and colleagues in a patient with JAK2 positive primary myelofibrosis (PM) and pancytopenia, although the role of GATA2 in the pathogenesis of PM remains unclear.95

Recurrent missense mutations in the extended ZF2 region diminishing DNA binding, including hotspot variants p.Thr354Met,1,6,10,96–99 p.Arg396Trp/Gln3–5,17,98–101 and p.Arg398Trp102 have been reported across a number of individuals and families and seem to be almost exclusively familial. Interestingly, similar to somatic GATA2 mutations p.Leu359Val82,93 and p.Arg307Trp102 the germline mutations p.Thr354Met and p.Cys373Arg were both reported to exhibit gain of binding affinity to a GATA2 partner protein (PU.1 encoded by SPI1 gene). This might suggest an additional mechanism to LOF driving leukemogenesis, namely context-dependent gain of function.20,92

Apart from missense variants, there are a few reports of synonymous GATA2 mutations in a handful of GATA2-deficient cases to date. Despite the unaltered amino acid composition, the variants result in RNA degradation and can be categorized as null alleles. The variant p.Thr117=, first described by Wehr and colleagues,103–105 along with four additional mutations, namely p.Leu217=, p.Gly327=, p.Ala341= and p.Phe472=, reported by Kozyra and colleagues,104 was found to strongly alter splicing and induce selective loss of messenger RNA.20,104,105 Although substantial phenotypic changes may support the putative role of synonymous GATA2 mutations, sequencing of GATA2 mRNA to confirm monoallelic loss and in silico analyses are essential to assess the actual impact of these seemingly innocuous, albeit potentially damaging variants on GATA2 expression for each individual case.20,104

Regulatory variants (in approximately 10% of patients) constitute an important group of germline mutations affecting the GATA2 intronic +9.5 enhancer site and thus leading to haploinsufficiency via decreased GATA2 expression in HSPCs. Consequently, damaging mutations and deletions of intron 4 (NM_032638 isoform) perturb the dose-dependent transactivation activity of the Gata2 intronic enhancer disrupting haematopoietic cell development and cell differentiation fates.35,38,38,59 Interestingly, no germline variants of the distant upstream Gata2–77 enhancer site have been described to date, although current clinical platforms (panel and exome sequencing) do not account for such regulatory
regions. Future studies using whole genome sequencing have the potential to uncover undefined causes of GATA2 deficiency, that is, in hitherto unknown regulatory regions.

Rare cases of whole gene deletions,\(^{97,106,107}\) in-frame deletions\(^{1,6,9,98}\) and missense mutations located outside of relevant protein domains\(^{5,108,109}\) have also been reported, although data and functional studies on their putative impact (presumably LOF) are still absent.\(^{86}\) Recently, a novel type of germline in-frame GATA2 variants has been reported by Cavalcante de Andrade Silva and colleagues in a patient with characteristic features of GATA2 deficiency, including bilateral lower limb lymphoedema, haematological manifestations, and recurrent infections. The familial in-frame insertion of nine amino acids between ZF1 and ZF2 increased the spacing between the two ZF's abrogating target gene regulation and cell differentiation.\(^{110}\)

**CLINICAL PHENOTYPES ASSOCIATED WITH GATA2 DEFICIENCY**

Detailed description of the clinical presentations of GATA2 deficiency syndrome is beyond the scope of this review, therefore we refer to comprehensive reviews discussing GATA2-associated phenotypes\(^ {20,89}\) and previously published large cohort studies.\(^ {4,17}\)

Myeloid malignancy is the most common phenotype associated with germline GATA2 mutations, manifesting in roughly 80% of reported carriers at a median age of approximately 20 years,\(^ {1,4,5,20,89}\) (Figure 1B). GATA2 deficiency is considered the most common germline predisposition in paediatric MDS, accounting for 15% of cases with MDS with excess blasts, and 7% of overall paediatric MDS within the co-operative European Working Group of MDS in Childhood (EWOG-MDS) cohort of 426 patients published by Wlodarski and colleagues.\(^ {4}\) These numbers were validated in a larger cohort of 669 patients from the EWOG-MDS registry, where GATA2 deficiency was detected in 7% of MDS cases and was mutually exclusive with germline SAMD9/SAMD9L mutations (another driver of paediatric marrow failure and MDS) which accounted for 8%.\(^ {111}\)

Immunological phenotypes of GATA2 predisposition most commonly present as cellular immunodeficiencies with subsequent recurrent or atypical mycobacterial, viral, and fungal infections and autoimmunity. Immune dysfunction is often the initial presentation in this patient population; however, pre-existing immunodeficiency may also be observed retrospectively in more than half of the cases with myeloid malignancy and thus is strongly suggestive for GATA2-related MDS/AML.\(^ {22,89}\) Prominent clinical manifestations of GATA2 deficiency also include repeatedly observed pulmonary alveolar proteinosis (PAP), and interstitial pulmonary diseases\(^ {112}\) and high incidence of solid tumours and precancerous lesions with the majority of the cancers arising from underlying viral infections, particularly HPV and EBV.\(^ {5,6,15,100,113}\)

Apart from general symptoms including fever, fatigue and weight loss, further extra-haematological phenotypes of GATA2 deficiency encompass premature birth, developmental delays and congenital malformations,\(^ {5,17,56,114,115}\) and high incidence of primary lymphedema and lymphadenopathy in the GATA2 deficient patient population underlying the crucial role of GATA2 in lymphatic angiogenesis and lymphatic valve development.\(^ {97}\)

**GENOTYPE–PHENOTYPE CORRELATIONS IN GATA2 DEFICIENCY**

Despite the rising number of reported cases with GATA2 deficiency, different types of germline GATA2 variants have not been linked strongly to substantial phenotypes, clinical outcomes or age of onset, although some phenotypes seem to cluster within affected carrier families (Figure 1B). Some analyses have shown missense GATA2 mutations to be associated with increased rates of myeloid malignancies compared to nonsense and frameshift variants\(^ {12,17,86}\) with a predominance of leukaemia in patients harbouring p.Thr354Met.\(^ {20}\) However, this association was not confirmed in a cohort of 173 European MDS patients with germline GATA2 variants, examined on behalf of the GATA2 HuMo consortium; cases with null alleles (encompassing nonsense, frameshift truncating, splice site, whole gene deletion and synonymous - RNA deleterious variants) had a higher prevalence of high-risk MDS (including leukaemia) in comparison to missense mutation carriers (unpublished observations of the EWOG-MDS study group). Recurrent or severe infections and underlying cellular immunodeficiencies have been linked to variants leading to premature translation termination and deletions;\(^ {12,86}\) although interestingly, missense mutations p.Arg398Trp and p.Arg396Gln, have also recently been associated with immunodeficiency by Homan and colleagues based on analysis of reported cases.\(^ {20}\)

Emberger syndrome has been exclusively reported in cases with regulatory, premature termination or LOF missense variants diminishing DNA binding and transactivation activity (p.Arg361Leu, p.Cys373Arg, p.Arg396Gln), implying the crucial role of haploinsufficiency in lymphedema development.\(^ {17,20,24,86,92}\) Underlying molecular mechanism for normal lymphatic vessel valve development and defects of the lymphatic vasculature in GATA2 deficiency was first proposed by Kazenwadel and colleagues.\(^ {24}\) They demonstrated that GATA2 levels in lymphatic endothelial cells are regulated via mechanic stimuli, including oscillatory fluid flow\(^ {116}\) and extracellular matrix-induced tension\(^ {42}\) driving valve morphogenesis and vessel sprouting, respectively. Stimuli-dependent expression and cell-specific transcription factor complexes of GATA2 in lymphatic endothelial cells seem to be profoundly impacted by haploinsufficiency compared to HSPCs. Based on a handful of cases, large, whole gene deletions and losses of chromosome 3q21.3 resulting in GATA2...
haploinsufficiency have been linked to dysmorphisms, developmental and neurological defects, monocytopenia and subsequent infections.\textsuperscript{12,97,97}

Regulation and transcription factors interacting with GATA2 vary greatly in selected biological systems and in the steady state versus stress, providing a possible explanation to phenotype-genotype clustering in GATA2 deficiency. Although it has been noted that substantial clinical presentations and defects of certain organs are related to distinct types of GATA2 variants, in vitro studies on GATA2 function in different cellular contexts and processes are required to provide novel insight to underlying molecular mechanisms of phenotypic complexity.\textsuperscript{20,21,86}

**SOMATIC ABERRATIONS IN MYELOID MALIGNANCIES WITH GATA2 PREDISPOSITION**

There are a number of recurrent cytogenetic aberrations and somatic mutations in a host of genes that occur commonly with GATA2 deficiency and subsequent myeloid malignancies. Monosomy 7 is the most frequently described clonal cytogenetic aberration. It has been reported in nearly half of the GATA2 deficient cases with associated myeloid malignancies\textsuperscript{20} and together with the deletion of 7q\textsuperscript{17} and unbalanced translocation der(1;7)(q10;p10)\textsuperscript{3,14,117} uniformly result in the monoallelic loss of chromosome 7q conferring poor prognosis.\textsuperscript{89,118} Among paediatric MDS with monosomy 7, 37\% of cases were found to carry germ line GATA2 mutations; this association is even more striking in adolescents with monosomy GATA2 cases where GATA2 deficiency is found in 72\% of patients.\textsuperscript{4,118,119}

Similarly, the association between der(1;7) and GATA2 deficiency was reiterated in a recent study by Kozyra and colleagues, demonstrating the majority (73\%) of primary MDS with der(1;7) have an underlying GATA2 deficiency.\textsuperscript{120} Isolated trisomy 8 is the second most common aneuploidy in GATA2 mutation carriers, occurring in approximately 20\% of the cases.\textsuperscript{1,3,6,17,20,98} Other cytogenetic events, such as trisomy 21,\textsuperscript{1,97,121} have also been encountered alone (isolated) or in combination with other aberrations; however, due to the lack of routine screening their prevalence is largely unknown to date. Contrary to MDS of adulthood, loss of 5q and complex karyotypes are generally not reported in GATA2 deficiency.\textsuperscript{11,22}

Recurrent somatic mutations in SETBP1, ASXL1, CBL, EZH2, KRAS/NRAS, JAK3, STAG2, RUNX1, and PTPN11 and STAG2 have been described in cases with GATA2-driven myeloid malignancies.\textsuperscript{88,96,101,106,122–131} LOF ASXL1 mutations are reported in approximately one-third of all patients with AML/MDS and similarly to SETBP1 variants, are strongly linked to monosomy 7 and thus associated with a more advanced disease state.\textsuperscript{96,98} Interestingly, single-cell level analyses reported by Pastor Loyola and colleagues revealed a non-random acquisition of these secondary events, implying monosomy 7 as an early somatic event in the MDS founding clone, followed by the acquisition of concomitant SETBP1 and ASXL1 mutations.\textsuperscript{89,132} Unlike in cases with germ line CEBPA, DDX41 and RUNXI predispositions,\textsuperscript{103–105} biallelic mutations are uncommon as somatic GATA2 mutations mainly within ZF1 seem to be rare secondary events in GATA2-driven myeloid neoplasia.\textsuperscript{20,127,133}

Similar to biological and environmental stresses, commonly seen somatically mutated genes and recurrent cytogenetic alterations are putative "stressor events" that may constitute pathogenetic triggers for clonal evolution and subsequent disease progression/evolution to AML.\textsuperscript{21} Therefore, there is a strong, and today still unmet need for comprehensive screening of concurrent somatic aberrations in GATA2 deficiency, especially in individuals and families with unexpected clinical presentations and incomplete penetrance.

**THERAPEUTIC STANDARD OF CARE**

Despite the increasing number of cases with GATA2 deficiency syndrome, currently there are no comprehensive guidelines and recommendations on the diagnosis and clinical management of individuals and families carrying germ line GATA2 mutations. Similar to other germline cancer predisposition syndromes, suspicion of GATA2 deficiency arises in cases with unusually early age of onset of myeloid malignancies, particularly MDS/AML, but positive family history may also be evocative.\textsuperscript{134} Although in some patients MDS/AML develops without pre-existing extra-haematopoietic phenotypes, others may show progressive symptoms indicative for GATA2 deficiency, such as cytopenias, generalized warts and monocytopenia, mycobacterial infections, recurrent viral (predominantly HPV) infections, and lymphedema.\textsuperscript{3,86,135} Cytogenetic testing of BM biopsy samples in cases with MDS/AML or unexplained cytopenias may be informative, as concurrent monosomy 7 and trisomy 8 can also point towards GATA2 deficiency.\textsuperscript{89}

Early diagnosis of GATA2 deficiency based on clinical presentation and/or genetic testing for germ line GATA2 mutations (with conventional bidirectional Sanger sequencing or next generation sequencing methods) is pivotal in avoiding non-curative therapies and especially immunosuppression due to underlying HSPC defects and immunodeficiency.\textsuperscript{89} Patients with high risk of evolution to advanced MDS or AML and/or high-risk karyotypes, such as monosomy 7, should undergo timely HSCT with selected conditioning regimens. Patients with stable disease (e.g. MDS with refractory cytopenia of childhood), without progressing immunodeficiencies underlying recurrent infections, transfusion-dependency and MDS with karyotypic aberrations or dysplasia may qualify for the watch and wait approach, although progressive disease can be expected over time in the majority of cases.\textsuperscript{3,11,89,136–138} General recommendations for surveillance include frequent evaluation of complete blood counts, annual BM biopsies with cytogenetic testing, analysis of lymphocyte subsets and immunoglobulins, and screening for subclinical presentations of pulmonary disorders (predominantly PAP) and (HPV- or
EBV-associated) cancers as part of yearly skin and gynaecological examinations.\(^8\) Currently, allogenic HSCT is the only curative approach for patients with GATA2 deficiency. There is consensus that in high-risk cases, the ideal point in time for pre-emptive HSCT is during the hypocellular phase of MDS prior to severe complications, such as invasive infections or evolution to MDS-EB and AML.\(^4,11,22,86,89,113,136\) Myeloablative conditioning regimens for HSCT are the preferred standard approach in paediatric MDS with GATA2 predisposition, but reduced intensity conditioning regimens have also been successfully used in some cases.\(^4,11,86,89,121,139\) Based on the results of EWOG-MDS cohort study, overall outcomes were not influenced by the presence of germline GATA2 variants with a comparably favourable 5-year overall survival of 66% in patients with MDS who had already acquired monosomy 7. Rates of infectious complications were also comparable to the control GATA2 wild-type group (66% vs. 61%),\(^4\) but despite the non-inferior results, HSCT in GATA2 deficiency remains challenging due to underlying immunodeficiencies and comorbidities threatening the outcome of HSCT.\(^86,137\) Interestingly, Hoffman and colleagues recently reported a higher incidence of thrombotic events and neurological complications in a cohort of patients with GATA2-driven paediatric MDS undergoing HSCT with myeloablative conditioning, although graft versus host disease (GVHD), graft failure and treatment related mortality rates were comparable to patients with non-GATA2-driven MDS/AML.\(^140\)

Considering highly variable phenotypes and incomplete penetrance of GATA2 deficiency, genetic testing of inherited GATA2 variants should be performed in family members to identify asymptomatic carriers.\(^20\) Although data and recommendations on early HSCT in phenotypically “silent” cases are still absent to date, baseline BM aspirate evaluation of healthy individuals carrying GATA2 variants is recommended by the EWOG-MDS group to assess GATA2-related cytogenetic aberrations.\(^86,88\) Similar to patients with stable disease, further management of asymptomatic carriers may include evaluation of complete blood counts, pulmonary function tests and surveillance for malignancy and immunodeficiencies.\(^88,141\) Notably, critical role of screening for inherited variants in relatives is also becoming generally recognized during HSCT donor selection in order to avoid donor-derived myeloid malignancies conferring dismal outcome.\(^20,142\)

**POTENTIAL NOVEL THERAPIES**

GATA2-related MDS/AML constitute a complex disorder with a profound impact on the quality of life of patients and their life expectancy. Due to the phenotypic diversity, clinical overlap, and the evolving phenotype of this disease, achieving a timely diagnosis is often difficult. This holdup has a direct impact on patient’s management. The only curative option for these patients is allogenic HSCT, usually performed in the setting of advanced MDS or severe immunodeficiency. The overall outcome after HSCT is favourable for the early stages of haematological disease but declines considerably in patients with advanced MDS. In addition, the lack of suitable donor and associated mortality and morbidity can hinder this procedure in a number of patients. Correction of patients’ HSCs by gene therapy could represent a promising alternative to allogenic HSCT, due to the proliferative advantage of corrected cells (Figure 3). This phenomenon has been recently described in one unique case, in whom spontaneous somatic genetic rescue in GATA2 gene conferred selective advantage of the corrected HSCs and...
annulled the effect of pathogenic GATA2 mutation. These findings strongly indicate that having few corrected HSCs might be sufficient to restore haematopoiesis in GATA2 carriers. It is important to highlight that the expression of GATA2 gene is tightly regulated in HSCs and during leukaemia progression. Therefore, classical lentiviral transgenic approaches do not represent the best option as gene therapy method to treat GATA2 deficiency. Advanced gene editing technology via homologous recombination (HR) has paved the way to the development of precision medicine and outcome-driven therapy for individual patients. This allows the specific correction of the mutated gene, maintaining the physiological expression of the gene and eliminating the integration of exogenous DNA material elsewhere. Moreover, gene editing based on CRISPR/Cas9 and large HR donor deliver by adeno-associated viral vector of serotype 6 (rAVV6) has already revealed promising clinical results, as recently shown in haematological diseases such as sickle cell disease and pyruvate kinase deficiency.

Based on these advances we can speculate that the optimization of a repair mechanism by large HR donor templates covering different exons could provide treatment for a substantial number of GATA2 patients allowing effective and sustainable translation in the clinical arena (Figure 3). Importantly, recurrent somatic mutations in MDS driver genes (i.e. SETBP1, ASXL1, STAG2 and RAS pathways) has been identified by us and others in GATA2-MDS patients. Therefore, a preliminary genomic screening for additional somatic mutations should be implement as standard procedure for each GATA2 deficient patient. This may serve as prognostic markers predicting the risk of relapse, and thus be crucial in guiding treatment strategy. Finally, a possible technical obstacle for clinical application of gene editing correction of GATA2 mutations is the off-target effects (OTEs) that might occur in undesired parts of the genome. Therefore, detection of OTEs with screening strategies, such as GUIDE-seq or CIRCLE-seq, should be used before any clinical translation of this gene therapy approach.

CONCLUSION

GATA2 deficiency is a monogenic disorder with complex clinical manifestations and a very high propensity for MDS/AML development. Carriers of GATA2 mutations show nearly complete life-time penetrance towards the development of myeloid neoplasia; however, it is not clear what factors (genetic, epigenetic, or inflammatory) modify the phenotype leading to variable penetrance observed for identical mutations. Due to the phenotypic diversity, clinical overlap and evolving phenotype of this disease, a timely diagnosis is often difficult to achieve. Therefore, a crucial unmet need for novel, genotype-specific, more efficacious therapies for GATA2 deficient patients remain. Recently, Catto and colleagues suggested that early recovery of haematopoiesis in GATA2 deficiency either by transplant or potentially by gene therapy may be beneficial for haematopoiesis and to prevent other clinical complications.

In this context, with the breakthroughs in CRISPR/Cas9 technology, it is an imaginable scenario where GATA2 carriers at early stage of the disease can be actively assisted by an innovative tailored and precise gene-based treatment. The reinfusion of corrected autologous HSCs would represent a promising and less toxic alternative treatment for GATA2-related MDS. However, a multidisciplinary effort is still needed to establish the clonal origin of leukemogenesis in GATA2 carriers in order to select the appropriate time frame to treat these patients by gene editing.

AUTHOR CONTRIBUTIONS
LK, DRM, CB and AG wrote the manuscript, OM-B prepared figures and reviewed the manuscript, EK, AC, AB, MWW reviewed and edited the manuscript.

ACKNOWLEDGEMENTS
This work was supported by ERA PerMed GATA2-HuMo Funding Mechanism (Spain: Acció instrumental de SLT011/18/00006 of the Department of Health of the Government of Catalonia, to AG and AB; Hungary: ED-18-1-2019-001 grant from the National Research, Development and Innovation Office to CB; and Germany: German Federal Ministry of Education and Research (BMBF) 2018-123/01KU1904 to MWW), ÚNKP-21-2-1-SE-21 and Hungarian National Academy of Scientist Education grant to KL, TKP2021-NVA-15, TKP2021-EGA-24 and EU’s Horizon 2020 research and innovation programme under grant agreement no. 739593 and Elixir Hungary to CB, MSCA No 101029927-scGATA2track (H2020-MSCA-IF-2020) to OM-B, the Spanish Ministry of Economy, Industry, and Competitiveness (MINECO PID2020-15591RB-100), La Marató de TV3 (202001-32), FPS Grant 2018 by Fondazione Pisana per la Scienza ONLUS and CERCA Programme/Generalitat de Catalunya for institutional support to AG, Vera and Joseph Dresner Foundation, Evans MDS Foundation, American Lebanese Syrian Associated Charities, and BMBF MyPred 01GM1911A to MWW. Open access funding enabled and organized by ProjektDEAL.

COMPETING INTEREST
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this study.

DATA AVAILABILITY STATEMENT
All data used in this manuscript was collected using the PubMed database (accessed: 20 April 2022).

ORCID
Lili Kotmayer https://orcid.org/0000-0002-5089-7659
Damia Romero-Moya https://orcid.org/0000-0001-9947-3577
Oskar Marin-Bejar https://orcid.org/0000-0002-6889-6566
GATA2 DEFICIENCY AND MDS/AML: EXPERIMENTAL STRATEGIES FOR DISEASE MODELLING AND FUTURE THERAPEUTIC PROSPECTS

Emilia Kozyra  https://orcid.org/0000-0003-1107-8818
Albert Catalá  https://orcid.org/0000-0003-0133-1752
Anna Bigas  https://orcid.org/0000-0003-4801-6899
Marcin W. Włodarski  https://orcid.org/0000-0001-6638-9643
Csaba Bódor  https://orcid.org/0000-0002-0729-692X
Alessandra Giorgetti  https://orcid.org/0000-0001-5803-3567

REFERENCES


