Invited Review

GATA Transcription Factors: Basic Principles and Related Human Disorders

Tohru Fujiwara

1Department of Hematology and Rheumatology, Tohoku University Graduate School of Medicine, Sendai, Miyagi, Japan

The development of mature blood cell from hematopoietic stem cells is regulated by transcription factors that coordinate the expression of lineage-specific genes. GATA transcription factors are zinc finger DNA-binding proteins that play crucial roles in various biological processes, including hematopoiesis. Among GATA family proteins, GATA-1, GATA-2, and GATA-3 are essential for hematopoiesis. GATA-1 functions to promote development of erythrocytes, megakaryocytes, eosinophils, and mast cells. Mutations in GATA-1 are associated with acute megakaryoblastic leukemia (AMKL), congenital erythroid hypoplasia (Diamond-Blackfan anemia; DBA), and X-linked anemia and/or thrombocytopenia. Conversely, GATA-2 functions early in hematopoiesis and is required for maintenance and expansion of hematopoietic stem cells (HSCs) and/or multipotent progenitors. GATA-2 mutations are associated with immunodeficiency, lymphedema, myelodysplastic syndrome (MDS), and leukemia. Furthermore, decreased GATA-2 expression may contribute to the pathophysiology of aplastic anemia. GATA-3 has an important role in T cell development, and has been suggested to be involved in the pathophysiology of acute lymphoblastic leukemias. This review summarizes current knowledge on hematological disorders associated with GATA-1 and GATA-2 mutations.

Keywords: GATA-1; GATA-2; hematological disease; hematopoiesis; transcription factor

Introduction

Differentiation of hematopoietic stem cells (HSCs) into specific progenitor cells and ultimately into diverse blood cell types is controlled by various transcription factors, represented by GATA family of zinc finger DNA-binding proteins (Bresnick et al. 2010). GATA-1, GATA-2, and GATA-3 are designated as hematopoietic GATA factors based on their important roles in controlling distinct as well as overlapping aspects of hematopoiesis (Orkin 1992; Bresnick et al. 2010). GATA-1 functions to promote erythrocyte, megakaryocyte, eosinophil, and mast cell development. Conversely, GATA-2 functions early in hematopoiesis and is required for maintenance and expansion of HSCs and/or multipotent progenitors, whereas GATA-3 plays a role in T cell development (Yamamoto et al. 1990; Tsai et al. 1994; Lim et al. 2000; Bresnick et al. 2010). GATA-1 and GATA-2 expression patterns overlap partially, such as that observed in primitive erythroblasts, megakaryocytes, and eosinophils (Zon et al. 1993; Tsai and Orkin 1997; Fujiwara et al. 2004), whereas their expression levels are mutually exclusive in other contexts. GATA-2 has its own regulatory locus in HSCs and multipotent progenitors and activates its own expression as well as other genes essential for HSC/progenitor function (Bresnick et al. 2010). Upon induction of GATA-1 expression during commitment of the erythroid lineage, GATA-1 displaces GATA-2 at the GATA-2 locus, leading to transcriptional repression (Bresnick et al. 2010). This so-called GATA switch appears to have an important role as a driver of erythroid maturation and might be contributing to pathophysiology of various human diseases such as anemia, immunodeficiency, myelodysplastic syndrome (MDS), and leukemia.

This review focuses on diseases caused by dysfunctions in GATA-1 and GATA-2, which are termed as “GATA-related diseases” in the present review.

GATA-1

GATA-1 is the founding member of the GATA family. Initial studies with targeted disruption of Gata1 in mice demonstrated defective erythropoiesis and embryonic lethality (Pevny et al. 1991; Fujiwara et al. 1996; Fujiwara...
Subsequent studies based on multiple mouse models revealed that GATA-1 was required for the differentiation of megakaryocytes, mast cells, eosinophils, and basophils (Vyas et al. 1999; Yu et al. 2002a; Nei et al. 2013). GATA-1 is an X-linked gene that encodes a DNA-binding protein containing two zinc finger domains, and the C-terminal zinc finger of GATA-1 mediates DNA binding to (A/T)GATA(A/G) consensus sequences (Evans et al. 1988; Wall et al. 1988; Martin et al. 1989). Analysis of GATA-1 chromatin occupancy in erythroid cells using chromatin immunoprecipitation coupled with massive next-generation sequencing (ChIP-seq) revealed a more complex consensus sequence \([(C/G)(A/T)GATA(A/G)(A/C)(G/A/C)]\) that was enriched at chromatin occupancy sites (Fujiwara et al. 2009). Conversely, the N-terminal zinc finger of GATA-1 enhances endogenous target gene activation repression in a context-dependent manner by interacting with a key cofactor, Friend of GATA-1 (FOG1) (Johnson et al. 2006; Fujiwara et al. 2017), and facilitates GATA-1 binding to a subset of binding sites containing a palindromic motif (Trainor et al. 1996; Hasegawa et al. 2016). Impaired GATA-1–FOG1 interaction appeared to be associated with the onset of hematological disorders, which will be discussed later.

GATA-1 forms a complex with another master regulator of hematopoiesis, the basic-helix-loop-helix transcription factor stem cell leukemia/T-cell acute lymphoblastic leukemia 1 (SCL/TAL1) (Wadman et al. 1997; Lahlil et al. 2004; Bresnick et al. 2010). SCL/TAL1 interacts with the N-terminal zinc finger of GATA-1 and regulates erythropoiesis as well as the functions of HSCs (Lahlil et al. 2004). In addition, in vitro DNA binding studies revealed that GATA-1-SCL/TAL1 assembles a complex containing E2A, LMO2, and LDB1, all of which are required for erythropoiesis (Shivdasani et al. 1995; Wadman et al. 1997; Göttgens et al. 2002; Lahlil et al. 2004; Fujiwara et al. 2009; Inoue et al. 2013). While the subset of genes that are directly suppressed by GATA-1 are known, specific partners of GATA-1 contributing to the repression of these genes are less clear (Kerenyi and Orkin 2010). The hematopoietic corepressor ETO2 associates with the GATA-1–SCL/TAL1 complex and confers transcriptional repression mediated through interaction with histone deacetylases (HDACs) (Amann et al. 2001; Tripic et al. 2009; Fujiwara et al. 2010; Soler et al. 2010; Fujiwara et al. 2013). GATA-1 can recruit polycomb repressive complex 2 (PRC2), a multi-unit complex composed of the EZH2 histone methyltransferase as well as the structural components SUZ12 and EED, leading to transcriptional repression through trimethylation of histone H3 at lysine 27 (H3K27me3) (Müller et al. 2002).

**GATA-1 mutations in inherited and acquired hematological diseases**

Several hematological disorders in humans were found to be associated with GATA-1 mutations. Acquired GATA-1 mutations are strongly associated with acute megakaryoblastic leukemia (AMKL) and transient abnormal myelopoiesis (TAM) in children with Down syndrome (DS). Congenital GATA-1 mutations are associated with congenital erythroid hypoplasia (Diamond-Blackfan anemia; DBA) and X-linked dyserythropoietic anemia and/or thrombocytopenia. Furthermore, decreased GATA-1 expression was also implicated to be associated with primary myelofibrosis (Gilles et al. 2017) and MDS (Höpfer et al. 2012).

One of the characteristic mutations in GATA-1 was observed in patients with acquired hematological diseases, DS-associated TAM and AMKL (Wechsler et al. 2002; Groet et al. 2003; Hitzler et al. 2003). TAM is abnormal myeloid proliferation that occurs in about 10% of children with DS (characterized by trisomy 21) but tends to resolve spontaneously in most cases (Massey et al. 2006). However, about 20%-30% of TAM cases will develop AMKL (Massey et al. 2006). Mutations of GATA-1 were found in almost all of cases with DS-AMKL. All mutations exist in the N-terminus of GATA-1, resulting in a premature stop codon, and an alternative short form of GATA-1 (GATA-1s) is translated from the methionine residue at amino acid 84 of wild-type GATA-1 (Fig. 1) (Massey et al. 2006). Additionally, a patient with AMKL who did not have DS was reported to possess a mutation within the GATA-1 N-terminal domain that led to the generation of GATA-1s (Harigae et al. 2004). Data from a transgenic mouse model confirmed this phenotype with massive megakaryocytosis resembling TAM (Shimizu et al. 2009), suggesting that GATA-1 mutations were not sufficient to AMKL development. Interestingly, a recent comprehensive whole exome analysis including TAM and DS-AMKL cases demonstrated strong evidence that GATA-1s and trisomy 21 were sufficient to induce TAM (Yoshida et al. 2013). That study further demonstrated that progression from TAM to AMKL was dependent on acquisition of additional mutations, including those encompassing cohesion components, epigenetic regulators (e.g., EZH2, DNMT1, and ASXL1), and common signaling pathways (e.g., MPL, JAK2, and NRAS) (Yoshida et al. 2013). However, the mechanism by which trisomy 21 predisposes children to GATA-1 mutations remains to be elucidated in future studies.

DBA is a bone marrow failure syndrome characterized by macrocytic anemia arising from reduced erythroid precursors in the bone marrow. Whereas more than half of patients with DBA harbor heterozygous loss-of-function mutations involving ribosomal protein genes, the molecular pathogenesis was unclear in a subset of cases (Lipton and Ellis 2009). Subsequently, several groups identified novel mutations involving the exon 2 donor splice site of GATA-1 (Hollanda et al. 2006; Sankaran et al. 2012; Klar et al. 2014) as well as the initiation codon (c. 2T>C) (Parrella et al. 2014) in patients with DBA. These mutations result in the production of GATA-1s in the absence of full-length GATA-1 (Fig. 1). Furthermore, decreased GATA-1 mRNA translation was observed in hematopoietic cells from patients with ribosomal haploinsufficiency (Ludwig et al.
suggesting that GATA-1 dysregulation might be an important event in DBA pathophysiology.

X-linked dyserythropoietic anemia and/or thrombocytopenia was reported in several families (Freson et al. 2001; Mehaffey et al. 2001; Nichols et al. 2002; Yu et al. 2002b; Balduini et al. 2004; Phillips et al. 2007; Tubman et al. 2007; Ludwig et al. 2014; Gao et al. 2015; Crispino and Horwitz 2017). Whereas all related mutations concentrate in the N-terminal zinc finger of GATA-1, clinical phenotypes vary among patients (Fig. 1); V205M leads to anemia and thrombocytopenia (Nichols et al. 2002), whereas D218G results in dyserythropoiesis and thrombocytopenia (Freson et al. 2001). While G208S is associated with thrombocytopenia (Massey et al. 2006) and R216Q is linked to thrombocytopenia and beta thalassemia (Yu et al. 2002b, Balduini et al. 2004; Tubman et al. 2007), R216W leads to congenital erythropoietic porphyria (Di Pierro et al. 2015). All mutations in the N-terminal zinc finger of GATA-1, except for R216 which interferes with DNA binding (Yu et al. 2002b), disrupt FOG1 interaction (Crispino and Horwitz 2017). While these mutations could affect the interaction of GATA-1 with TAL1, the contribution of TAL1 and its interacting partners to the various clinical phenotypes of X-linked dyserythropoietic anemia and/or thrombocytopenia remains to be elucidated.

Mutations in cis elements recognized by GATA-1 were also shown to be associated with human diseases. X-linked sideroblastic anemia (XLSA) is caused by mutations in the X-linked gene ALAS2 that encodes the first enzyme of heme biosynthetic pathway in erythroid cells (Fujiwara and Harigae 2013). Whereas XLSA-related mutations in ALAS2 are usually missense mutations that affect conserved amino acids and lead to a loss of function, mutations in ALAS2 regulatory regions involving intron 1 were demonstrated to lead to decreased ALAS2 expression (Campagna et al. 2014; Kaneko et al. 2014).

GATA-2 is highly expressed in HSCs, multipotent hematopoietic progenitors, erythroid precursors, megakaryocytes, eosinophils, and mast cells (Zon et al. 1993; Tsai and Orkin 1997; Minegishi et al. 1999; Fujiwara et al. 2004; Bresnick et al. 2010). GATA-2 expression is essential for HSCs maintenance but is dispensable for erythroid and myeloid terminal differentiation (Tsai and Orkin 1997). Targeted disruption of Gata2 in mice was embryonically lethal due to failure of definitive hematopoiesis (Tsai et al. 1994), whereas mice with heterozygous deletion of Gata2, albeit viable, exhibited compromised HSC longevity (Rodrigues et al. 2005). Conversely, GATA-2 overexpression in murine bone marrow was shown to inhibit hematopoiesis (Persons et al. 1999), whereas mice with heterozygous deletion of Gata2, albeit viable, exhibited compromised HSC longevity (Rodrigues et al. 2005). Conversely, GATA-2 overexpression in murine bone marrow was shown to inhibit hematopoiesis (Persons et al. 1999). Beyond its role in hematopoietic cells, GATA2 is specifically required for mesenchymal stem cell differentiation, which will be discussed later (Tong et al. 2000; Okitsu et al. 2007; Kamata et al. 2014).

Human diseases caused by GATA-2 dysregulation

GATA-2 deficiency syndrome: Recently, heterozygous GATA-2 germline mutations, both inherited and de novo, were reported to cause three overlapping clinical entities characterized by a predisposition to MDS and acute myeloid leukemia (AML): (i) familial MDS/AML, (ii)
Emberger syndrome, and iii) an immunodeficiency due to monocytopenia characterized by *Mycobacterium avium* complex (MonoMAC)/dendritic cell (DC), monocyte, B- and natural killer (NK)-lymphoid deficiency (DCML) (Dickinson et al. 2011; Hsu et al. 2011; Ostergaard et al. 2011). These conditions are collectively termed GATA-2 deficiency syndrome. Nearly 100 different GATA-2 mutations that have been identified thus far appear to share certain characteristics such as reduced or abrogated GATA-2 transcriptional activity (Collin et al. 2015). Approximately two-thirds of all cases have mutations in the N- and C-terminal zinc finger domains (Fig. 2). Whereas familial MDS/AML was specifically caused by GATA-2 missense mutation within C-terminal zinc finger domain, mutations observed in MonoMAC/Emberger’s syndromes include whole gene deletion, frameshift mutation, missense mutation, and regulatory mutation involving GATA-2 intronic enhancer region (Fig. 2) (Johnson et al. 2012; Collin et al. 2015). In patients with GATA-2 deficiency syndrome, monocyte, B cell, NK cell, and DC populations are profoundly diminished or undetectable (Dickinson et al. 2011; Hsu et al. 2011), whereas neutrophil, macrophage, and T cell populations remain unaltered (Dickinson et al. 2011; Hsu et al. 2011). In addition, these patients sometimes develop pulmonary alveolar proteinosis resulting from dysregulated phagocytic activity and cytokine production in alveolar macrophages (Vinh et al. 2010). On the other hand, Emberger syndrome is characterized by primary lymphedema and predisposition to AML (Ostergaard et al. 2011).

Given that DCs play crucial roles in the immune system and as their numbers are profoundly decreased in GATA2 deficiency syndrome, we recently examined the roles of GATA2 in DC differentiation (Onodera et al. 2016). We found greatly reduced numbers of splenic DCs in conditional *Gata2* knockout mice. In *Gata2* knockout mice, we further determined that DC generation was impaired in Lin−/Sca-1+/Kit+ (LSK) cells, CMPs, and CDPs but not in granulocyte-macrophage progenitors (GMPs) or common lymphoid-restricted progenitors (CLPs). Furthermore, gene expression analysis data revealed downregulation of myeloid-related genes and upregulation of T cell-related genes, including *Gata3*, in *Gata2* knockout DC progenitors. This finding suggested that GATA2 might play an important role in cell fate specification toward myeloid versus T lymphocyte lineage by regulating lineage-specific transcription factors in DC progenitors, thereby contributing to DC differentiation. Our findings, together with those of previous reports (Vinh et al. 2010; Dickinson et al. 2011; Hsu et al. 2011; Ostergaard et al. 2011), provide new insights into the role of GATA-2 in immune function and regulation of growth, maturation, and apoptosis in the myeloid compartment.

Nearly half of individuals harboring GATA-2 mutations will eventually develop MDS/AML in association with fibrosis and megakaryocyte dysplasia (Dickinson et al. 2011; Hsu et al. 2011; Ostergaard et al. 2011; Collin et al. 2015). Disease progression from MDS to AML in patients with GATA-2 deficiency appear to be more rapid compared to MDS patients with wild-type GATA-2, despite comparable prognostic scores (Grossman et al. 2014). Considering the important role of GATA-2 in HSC proliferation (Tsai et al. 1994; Tsai and Orkin 1997; Ezoe et al. 2002; Rodrigues et al. 2005), additional genetic events might explain the progression of patients with GATA-2 deficiency to MDS/AML. In this regard, the most commonly associated cytogenetic finding is monosomy 7 as well as additional acquired mutations such as those in *ASXL1* (Hahn et al. 2011; Bödör et al. 2012; West et al. 2014; Churpek et al. 2015). However, the molecular basis for the evolution of GATA-2 deficiency into MDS/AML has not been elucidated, which impacts early detection and treatment of the disease. We recently conducted whole genome sequencing of MDS samples, which were compared with matched samples from nails and leukocytes of patients with immunodeficiency and those from bone marrow-derived mesenchymal stem cells (BM-MSCs). We identified three candidate mutations in *EZH2*, *HECW2*, and *GATA-1* that might contribute to the evolution of the disease (Fujiwara et al. 2014). In addition to the germline mutations, acquired
GATA-2 mutations also contribute to the development of MDS and leukemias. A recent large-scale genetic analysis demonstrated that GATA-2 mutations, which were critical for clonal evolution in MDS, had a relatively weaker impact on the progression of AML (Makishima et al. 2017). Chromosomal translocations involving GATA-2 upstream enhancer region (−77 kb), such as inv(3)(q21q26) and t(3;3) (q21;q26), contribute to the onset of AML (Gröschel et al. 2014; Yamazaki et al. 2014). Both inv(3)(q21q26) and t(3;3) (q21;q26) were proposed to translocate oncogenic EVI1 gene to an area close to the GATA-2 −77 kb enhancer region, leading to the aberrant overexpression of GATA-2 (Gröschel et al. 2014; Yamazaki et al. 2014). Additional GATA-2 mutations were identified in acute myeloid transformation of chronic myeloid leukemia, which are caused by activating GATA-2 mutation (L359V) (Zhang et al. 2008), leading to the enhanced inhibitory effect on the myeloid transcription factor PU.1.

Aplastic anemia: Aplastic anemia (AA) is characterized by a decrease in the number of HSCs and fatty marrow replacement (Young 1999). However, AA pathogenesis is complex, and many questions remain unanswered. Immunological injury to HSCs has been proposed to underlie the reduction in the number of HSCs in bone marrow (Solomou et al. 2007; Young et al. 2008, 2010), based on in vitro analysis of samples obtained from AA patients as well as clinical evidence that immunosuppressive therapy with drugs such as anti-thymocyte globulin and cyclosporine A is effective in 75% of AA cases (Young et al. 2008, 2010). Conversely, some patients do not respond to immunosuppressive therapy and may even develop MDS or acute leukemia as a consequence of clonal evolution (Bacigalupo et al. 2000), which suggests that immunological aberration is not the sole disease mechanism. In these latter cases, intrinsic abnormalities such as telomere attrition (Dumitriu et al. 2015) and loss of heterozygosity at short arm of chromosome 6 (6pLOH) (Katagiri et al. 2011) were also suggested.

Given the crucial role of GATA-2 in the function and maintenance of HSCs, it is possible that changes in GATA-2 expression levels might lead to aberrant proliferation and differentiation of HSCs and may be responsible for the development of AA. Previously, we demonstrated decreased expression of GATA-2 in CD34-positive cells in AA (Fujimaki et al. 2001). This observation was confirmed in another study utilizing microarray-based transcriptional profiling of CD34-positive cells from patients with AA (Zeng et al. 2004). Furthermore, the merged analysis of GATA-2 ChIP-seq and expression profiling identified HOXB4, which encodes a protein with an important role in HSC expansion (Antonchuk et al. 2002), as a direct downstream target (Fujiwara et al. 2009). We demonstrated that GATA-2 positively regulates HOXB4 expression, through binding to the GATA binding motif located at the promoter (Fujiwara et al. 2012). Overall, these results indicate that some stem cell-specific genes, including HOXB4, are aberrantly expressed as a consequence of GATA-2 downregulation in HSCs in AA, which might be contributing to the development and/or progression of the disease.

BM-MSCs are self-renewing precursor cells that differentiate into bone, fat, cartilage, and stromal cells of the bone marrow, thereby forming a microenvironment that maintains HSCs (Frenette et al. 2013). However, the mechanism of BM-MSC differentiation into adipogenic progenitors and ultimately into mature adipocytes in the bone marrow remains to be elucidated. In this regard, we and others demonstrated that GATA-2 was expressed in BM-MSCs and played a central role in the control of adipogenesis (Tong et al. 2000; Okitsu et al. 2007; Kamata et al. 2014). In addition to the regulation of BM-MSC differentiation into adipocytes, GATA-2 was reported to have an important role in cell cycle regulation of BM-MSCs as well as in hematopoietic support, as demonstrated in in vitro co-culture studies (Kamata et al. 2014). In conjunction with the evidence that GATA2 expression was decreased in BM-MSCs from patients with AA (Xu et al. 2009), these findings suggest that GATA-2 might be important in maintaining the BM microenvironment and contributing to the pathophysiology of AA.

Finally, given the pathophysiological links between GATA-2 and human diseases including AA, GATA-2 deficiency syndrome, lung cancer (Kumar et al. 2012), and prostate cancer (Rodriguez-Bravo et al. 2017), a comprehensive and clear understanding of the details related to the mechanisms regulating GATA-2 transcription is critical. We recently conducted a high-throughput screening analysis based on a small interfering RNA (siRNA) library to gain novel insights into factors involved in the regulation of GATA-2 expression. We screened 995 transcription factor genes and found that CITED2 acted as a GATA-2 activator in human hematopoietic cells (Saito et al. 2015). We believe that targeting or otherwise boosting residual GATA-2 activity, directly or indirectly, may represent a novel therapeutic strategy.

Conclusion

Dysfunctions in GATA-1 and GATA-2 were observed in several hematological disorders, primarily with apparent involvement of erythroid and myeloid lineages. Given the rapid advances in understanding the role of GATA transcriptional factors in human hematological diseases, further preclinical studies, in conjunction with extensive analyses of clinical samples, has the potential to yield pivotal insights into the molecular basis of various human diseases and contribute to the development of targeted therapies.

Acknowledgments

Author thanks the staff of the Department of Hematology and Rheumatology at Tohoku University Graduate School of Medicine in Sendai, Japan for preparing the manuscript.
Conflict of Interest

The author declares no conflict of interest.

References


Gilles, L., Arslan, A.D., Marinaccio, C., Wen, Q.J., Arya, P.,


