GATA1 gets personal

Anna Rita Migliaccio

Dipartimento di Scienze Biomediche e NeuroMotorie, Alma Mater Studiorum - Università di Bologna, Bologna and Tisch Cancer Institute, Ichan School of Medicine at Mount Sinai, New York, NY, USA

E-mail: ANNA RITA MIGLIACCIO - annarita.migliaccio@unibo.it or annarita.migliaccio@mssm.edu

doi:10.3324/haematol.2019.246355

Until the beginning of the 20th century, physicians had very few drugs at their disposal to assist them in their fight against disease. The blooming of chemistry into the new field of pharmacology completely revolutionized the concept of "treatment" in medicine. Pharmacology made a large array of chemicals available to physicians and this greatly improved their ability to treat or at least to halt the progression of numerous diseases. The synthesis of new effective chemical entities drove an optimistic belief that it would soon be possible to produce a "compound" to specifically cure every "disease". This belief has been tamed by the fact that, in recent years, the rate of discovery of new effective drugs has greatly decreased. The current consensus in the field is that pharmacology has exhausted its potential, and that any further progress in our ability to cure will be driven by being able to match the "treatments" with the "driver genetic lesions" and/or by personalizing the "treatment modality", including optimization of the supportive treatments, to the "genetic and life-style" profile of the patient. This awareness has given rise to the development of the novel fields of precision[1] and/or personalized[2-4] medicine.

One of the most common consequences of the chemotherapy used to treat cancer is a severe and potentially lethal anemia which is treatable by blood transfusion, which represents an essential part of modern patient care. Transfusion has become a safe and widely available therapy thanks to the discovery of the major blood types by Karl Landsteiner, who received the Nobel price in 1930, and by the establishment of national and international blood banks in the 1940-1950s.[5] Blood, however, is a limited human resource, and although in western countries the blood supply is currently sufficient, in developing countries, the supply rarely meets existing needs.[6] Furthermore, changes in human demographics predict a progressive increase in the population >60 years of age, restricting the pool of eligible donors, and increasing the blood requirements needed to support advanced surgical procedures and medical treatments in older individuals. This means that, even in industrialized countries, by as early as 2050, the blood supply may no longer be adequate.[7] Additional challenges are presented by the outbreaks of novel infective diseases that increase the numbers of tests a transfusion product must undergo before it can be considered safe and by the awareness that the shelf-life of blood products may be much shorter than previously foreseen, which increases the number that go to waste. These challenges are posing an exponential burden to the economic and social costs of producing blood. The transfusion medicine community is aware of these challenges and the need to ensure an adequate blood supply that is safe and affordable in order to meet the clinical needs of the 21st century. More stringent and personalized criteria are, therefore, being implemented to define the patient populations in need of transfusion. These criteria are essential for evidence-based calculations of the blood products required at any given time, to reduce waste, and, therefore, the availability and costs of this therapy.

Human populations vary enormously in the number and biophysical properties of the red blood cells (RBC) present in the circulation and in the ability to recover from blood loss. The genetic basis of this variability has been the subject of several Genome Wide Association Studies (GWAS). These have identified several single nucleotide polymorphisms (SNP). These are regions of the DNA which do not affect the coding sequence of a gene, but rather the efficiency of its mRNA transcription and/or translation, and ultimately regulate the final protein content which predicts, among other factors, the genetic basis of RBC variability. Most of the SNP associated with inherited erythroid traits identified so far are located in the putative regulatory regions of *cKIT*, the gene which encodes the receptor for stem cell factor, the growth factor which, in combination with erythropoietin, regulates RBC production. These SNP lie close to a prominent DNase hypersensitive region approximately 115 kb upstream of *cKIT* and are associated with the variability in RBC counts, mean RBC volume, and mean RBC hemoglobin content observed in the normal population.[8,9] However, so far these studies have failed to drive personalized medicine strategies, for example, by identifying biomarkers for risk stratification of patients subjected to erythroid-related therapies. Surprisingly, these studies did not recognize SNP associated with GATA1.

GATA1 is a transcription factor which regulates the maturation of erythroid cells in a concentration-specific fashion. It is encoded by a gene, *GATA1*, located both in mice and men on the non-autosomal region of the X chromosome.[10] Therefore, males are hemizygote for the maternal GATA1 allele while, due to the Lyonization effect, the hematopoietic stem cells of females are a mosaic of cells carrying either the maternal or the paternal allele in an active configuration. *GATA1* is so important for erythropoiesis that mice lacking this gene die *in utero* from severe anemia and thrombocytopenia. Genetic alterations leading to a dysfunctional GATA1 protein have been identified in several diseases:[10] (i) point mutations in the coding sequences of the gene impairing, directly or indirectly, the DNA binding ability of the protein were detected in X-linked disorders associated with erythroid and/or megakaryocyte/neutrophil phenotype; (ii) deletions of the regulatory regions of the gene that eliminate a transcription start site required for the transcription of the mRNA encoding the full length form of the protein, making the cells hypomorphic for GATA1 protein, are associated with

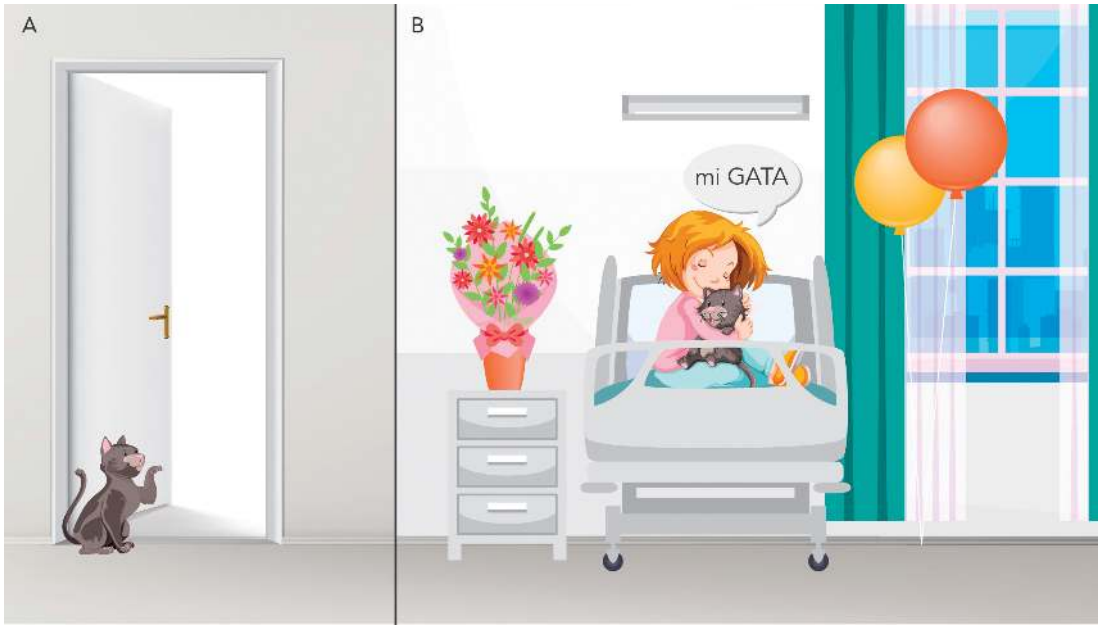

Figure 1. (A) A single nucleotide polymorphism hampering a binding site for the transcription factor GATA1 may predict whether children with acute lymphoblastic leukemia will need transfusion after chemotherapy. (B) This concept is reflected by the image of a little girl in a hospital thanking her GATA site (mi GATA in Spanish) for not having to undergo transfusion.

megakaryocyte leukemia in children with Down Syndrome and in sporadic forms of acute megakaryocyte leukemia in adults. These mutations are also detected in some patients with the inherited Diamond Blackfan anemia. Furthermore, in this disease, hypomorphic levels of GATA1 may result also from mutations in the gene encoding the ribosome protein RSP14 that impairs the ribosome ability to bind and translate *GATA1* mRNA. In addition, progression of benevolent forms of myeloproliferative neoplasms to their fatal myelofibrosis state is associated with a RSP14 signature that makes megakaryocytes of the patients hypomorphic for GATA1.[11] This knowledge has prompted studies that have targeted GATA1 for clinical purposes as pharmacological rescue of GATA1 has been shown to be an effective treatment for myelofibrosis.[12] However, surprisingly, GWAS studies have not yet identified clinically relevant correlations between SNP located either in the regulatory regions of *GATA1* or in GATA1-DNA binding sites, and variability of erythroid traits. One of the few exceptions is an old study which identified a SNP in the GATA1-binding site of the gene encoding the truncated form of the stem cell kinase receptor that determines the susceptibility to develop polycythemia in response to infection with the Friend virus in mice.[13]

In this issue of Haematologica, the paper by Xie *et al.* identifies by whole genome sequencing a single nucleotide polymorphism (SNP) in the intron of a gene (*ARHGEF12*) encoding a RhoA guanine nucleotide exchange factor which predicts whether children with acute lymphoblastic leukemia (ALL) will require multiple transfusions after chemotherapy.[14] On the basis of the assumption that chemotherapy hampers RBC production, and that the requirement for transfusion after this treatment depends on the efficiency with which the patient activates the stress erythropoiesis pathway necessary to overcome anemia, the authors hypothesized that *ARHGEF12* encodes a protein that plays an important role in the control of terminal erythroid maturation and that this function is affected by the SNP. To test this hypothesis, Xie *et al.* carefully determined that the SNP is located in a regulatory site disrupting a functional binding site for GATA1 and that, at the homozygous state, it reduces the gene transcription rates by 60%. Loss-of-function studies of *ARHGEF12* in animal models indicated that the gene regulates erythropoiesis at the proerythroblast level and studies in cell lines indicated that ARHGEF12 is upstream of a rhoA-p38 signaling pathway, the forced activation of which rescued impaired erythropoiesis induced by ARHGEF12 deficiency in animal models.

This study is important because it provides information on many levels. (1) It identifies a new signaling pathway, ARHGEF12/rhoA/p-38, which regulates the progression of proerythroblasts toward terminal maturation. Although the ligand which activates this pathway has not yet been identified, it is conceivable that it may represent a novel erythroid-stimulating agent other than erythropoietin which may have clinical value in the treatment of anemia. (2) It identified an SNP which is associated with the requirement for blood transfusion in children with ALL undergoing chemotherapy. If this association is validated by additional prospective studies on children with ALL, and possibly in other cancers, genetic testing for this SNP may represent a biomarker to stratify cancer patients at risk of anemia, allowing the blood supply necessary to support these patients during therapy to be calculated from the data base. (3) Last but not least, it suggests that GWAS studies designed on the basis of well-characterized patient cohorts may be more informative for devising precision/medicine

strategies than studies based on large numbers of healthy donors which may be flawed by the great level of noise in the signals generated with this strategy.

Funding

Supported by grants from the National Cancer (P01-CA108671), Heart, Lung and Blood (1R01-HL134684) Institute, Associazione Italiana Ricerca Cancro (AIRC IG23525)

References

1. Dzau VJ, Ginsburg GS. Realizing the Full Potential of Precision Medicine in Health and Health Care. JAMA. 2016;316(16):1659-1660.
2. Vogenberg FR, Isaacson Barash C, Pursel M. Personalized medicine: part 1: evolution and development into theranostics. P T. 2010;35(10):560-576.
3. Vogenberg FR, Barash CI, Pursel M. Personalized medicine: part 2: ethical, legal, and regulatory issues. P T. 2010;35(11):624-642.
4. Vogenberg FR, Barash CI, Pursel M. Personalized medicine: part 3: challenges facing health care plans in implementing coverage policies for pharmacogenomic and genetic testing. P T. 2010;35(12):670-675.
5. Alter HJ, Klein HG. The hazards of blood transfusion in historical perspective. Blood. 2008;112(7):2617-2626.
6. Report of the US Department of Health and Human Services. The 2009 National Blood Collection and Utilization Survey Report. Washington, DC: Services, T.U.D.o.H.a.H; 2010.
7. 3. Ali A, Auvinen MK, Rautonen J. The aging population poses a global challenge for blood services. Transfusion. 2010;50(3):584-588.
8. Kamatani Y, Matsuda K, Okada Y, et al. Genome-wide association study of hematological and biochemical traits in a Japanese population. Nat Genet. 2010;42(3):210-215.
9. van der Harst P, Zhang W, Mateo Leach I, et al. Seventy-five genetic loci influencing the human red blood cell. Nature. 2012;492:369-375.
10. Crispino JD, Horwitz MS. GATA factor mutations in hematologic disease. Blood. 2017;129(15):2103-2110.
11. Ling T, Crispino JD, Zingariello M, Martelli F, Migliaccio AR. GATA1 insufficiencies in primary myelofibrosis and other hematopoietic disorders: consequences for therapy. Expert Rev Hematol. 2018;11(3):169-184.
12. Gangat N, Marinaccio C, Swords R, et al. Aurora Kinase A Inhibition Provides Clinical Benefit, Normalizes Megakaryocytes, and Reduces Bone Marrow Fibrosis in Patients with Myelofibrosis: A Phase I Trial. Clin Cancer Res. 2019;25(16):4898-4906.
13. Persons DA, Paulson RF, Loyd MR, et al. Fv2 encodes a truncated form of the Stk receptor tyrosine kinase. Nat Genet. 1999;23(2):159-165.
14. Xie Y, Gao L, Xu C, et al. ARHGEF12 regulates erythropoiesis and is involved in erythroid regeneration after chemotherapy in acute lymphoblastic leukemia patients. Haematologica. 2020;105(4):925-936.

Genomic profiling of histiocytic sarcoma: new insights into pathogenesis and subclassification

Jonathan Said

Department of Pathology and Laboratory Medicine, UCLA Medical Center and David Geffen School of Medicine, Los Angeles, CA, USA

E-mail: JONATHAN SAID - jsaid@mednet.ucla.edu

doi:10.3324/haematol.2019.246314

A revised classification of histiocytoses and their neoplasms was long overdue when members of the Histiocyte Society suggested dividing them into five groups, designated L (Langerhans cell), C (Cutaneous), M (Malignant), R (Rosai-Dorfman), and H (Hemophagocytic) (Figure 1).[1] As an example, the L group, which includes Langerhans cell histiocytosis and Erdheim Chester disease, is characterized by mutations in the MAPK pathway and BRAF V600E. In contrast histiocytic sarcomas in the M group remain an elusive category. These rare and highly malignant neoplasms occur at all ages, and frequently involve extranodal sites including skin, soft tissues, and the gastrointestinal tract. There are few objective criteria for diagnosis other than expression of histiocyte markers (CD68, CD163, CD4, lysozyme, CD21, CD35, S100) and exclusion of other tumors by a panel of antibodies including, but not limited to, S100, keratins, EMA, Melan-A, HMB45, and B- and T-lymphoid markers.[2] Further complications arise in the apparent plasticity between histiocytic sarcomas and other malignancies, such as follicular lymphomas, demonstrated by translocations, immunoglobulin gene rearrangements, or mutational analysis.[3] Because of these pitfalls, histiocytic sarcoma has been vastly over diagnosed, with mimics including T-cell and other lymphomas with histiocyte-rich backgrounds, and, in particular, CD30 positive anaplastic large cell lymphomas. Clearly, any progress in diagnosing and treating these aggressive neoplasms will depend on identifying specific molecular and other markers for their accurate diagnosis.

Until recently, molecular analysis of histiocytic sarcomas has given confusing results, and there have been no consistent cytogenetic abnormalities. Mutations involving the RAS-MAPK signaling pathway, BRAF V600E mutations, as activation of PI3K and the tumor suppressor gene CDKN2A[4] have been most frequently reported, and there have been no reports of ALK translocations. Some cases, particularly in patients with associated B-cell lymphomas, have demonstrated immunoglobulin heavy or light chain gene rearrangements.

In this issue of Haematologica, Egan *et al.* performed genomic profiling of 21 primary histiocytic sarcomas, and in addition to confirming frequent alterations in the RAS/MAPK pathway, identified a novel intra-chromosomal transcript between exon 12 of TTYH3 and exon 8 of BRAF on chromosome 7. Moreover, differential expression analysis identified two distinct molecular subgroups with distinct molecular profiles and associated clinical characteristics based on the presence or absence of NF1/PTP11 mutations.[5] Cases that had NF1/PTPN11