Sickle cell disease in the era of precision medicine: looking to the future

Martin H Steinberg, Sara Kumar, George J. Murphy & Kim Vanuytsel

To cite this article: Martin H Steinberg, Sara Kumar, George J. Murphy & Kim Vanuytsel (2019): Sickle cell disease in the era of precision medicine: looking to the future, Expert Review of Precision Medicine and Drug Development, DOI: 10.1080/23808993.2019.1688658

To link to this article: https://doi.org/10.1080/23808993.2019.1688658
Sickle cell disease in the era of precision medicine: looking to the future

Martin H Steinberg, Sara Kumar, George J. Murphy and Kim Vanuytsel

Department of Medicine, Division of Hematology/Oncology, Center of Excellence for Sickle Cell Disease and Center for Regenerative Medicine, Boston University School of Medicine and Boston Medical Center, Boston, MA, USA

ABSTRACT

Introduction: Sickle cell anemia is a Mendelian disease that is noted for the heterogeneity of its clinical expression. Because of this, providing an accurate prognosis has been a longtime quest.

Areas covered: Reviewed are the benefits and shortcomings of testing for the major modulators of the severity of the disease, like fetal hemoglobin and a thalassemia, along with studies that have attempted to link genetic variation with sub-phenotypes of disease in a predictive fashion. Induced pluripotent stem cells driven to differentiate into erythroid precursor cells provide another area for potential patient-specific drug testing.

Expert opinion: Fetal hemoglobin is the strongest modulator of sickle cell anemia but simply measuring its blood levels is an insufficient means of forecasting an individual’s prognosis. A more precise method would be to know the distribution of fetal hemoglobin levels across the population of red cells, an assay not yet available. Prognostic measures have been developed using genetic and other signatures, but their predictive value is suboptimal. Widely applicable assays must be developed to allow a tailored approach to using several new treatments that are likely to be available in the near future.

1. Introduction

Adult hemoglobins are tetramers formed by 2 α- and 2 non-α-globin polypeptides. Hemoglobin A (HbA; α2β2) comprises more than 96% of the total hemoglobin while HbA2 (α2δ2) and fetal hemoglobin (HbF; α2γ2) contribute, respectively, about 2.5% and <1%. Hemoglobin remains in solution during transitions between oxy (relaxed or R) and deoxy (tense or T) states as it loads oxygen in the lungs and delivers it to tissues. Sickle hemoglobin (HbS; α2β6G; glutamic acid (E) 7 valine (V) GAG-G7G), one of about 1200 mutant hemoglobins, has the nearly unique property of undergoing reversible polymerization within seconds of deoxygenation. This deforms the sickle erythrocyte – a process called sickling – and is accompanied by a cascade of pathologic events that damage the cell causing hemolytic anemia and vasoocclusion (Figure 1). HbS has reached polymorphic frequencies in some populations because of the increased genetic fitness of heterozygotes under selective pressure from Plasmodium. Sickle cell anemia, defined as homozygosity for the HbS mutation, is a prototypical mendelian disease. Each patient has the same HbS-coding mutation (11p15.4; HBBS;141900; rs334; OMIM 603903). Nevertheless, the disease is noted for its clinical heterogeneity.

Compound heterozygosity for HbS and HbC, HbD and other less common β-globin variants, along with compound heterozygotes with HbS and β thalassemia, also cause sickle cell disease. Generally, most of these compound heterozygous conditions are less common than HbS homozygotes; patients. They also tend to have less severe phenotypes. Most information on phenotypic heterogeneity comes from studies of sickle cell anemia, therefor the following discussion focuses on this genotype.

Biomarkers that reflect the clinical diversity of sickle cell anemia would be useful for disease management. Herein, we discuss genetic markers associated with phenotypic variability; complex approaches to defining disease severity; the possible use of individualized disease modeling with induced pluripotent stem cells or iPSCs. The advent of new disease-modifying therapy and the more efficacious application of the current treatment, both of which are likely to alter the trajectory of the disease, make the quest for prognostic biomarkers and a personalized approach to treatment more important than ever.

2. Genetic markers of phenotypic variability

2.1. HbF

HbF is the predominant modulator of the phenotype of sickle cell anemia, an effect mediated by the exclusion of HbF and its mixed hybrid tetramer, α2β6γ, from the HbS polymer. This has prompted efforts toward understanding the genetic regulation of HbF gene expression and therapeutically inducing increased expression of the HbF genes (HBG2, HBG1).

HbF is restricted to a subset of erythrocytes called F-cells that comprise 2% to 80% of sickle erythrocytes. These cells contain at least 6 pgs. of HbF, which is the lower limit of HbF...
detectable by the common FACS assay. While the numbers of F-cells and the levels of HbF in each F-cell are likely to be controlled genetically this regulation is complex, multigenic and poorly understood. Our knowledge of hemoglobin switching, the process whereby HbF genes are silenced and the adult hemoglobin genes maximally expressed, is better understood, providing targets for reactivation of these genes or genetically increasing their expression in erythroid progenitors [1]. Total HbF level in sickle cell anemia is determined by the absolute number of F-cells, the amount of HbF/F-cell and the improved survival of F-cells compared with non-F-cells [2]. The primacy of HbF as an inhibitor of HbS polymerization is vividly demonstrated in compound heterozygotes for HbS and gene deletion hereditary persistence of HbF (HPFH). These individuals have nearly 30% HbF that is distributed almost equally amongst their erythrocytes so that each contains about 10 pg. This pancellular distribution endows each cell with sufficient HbF to protect it from HbS polymer-induced injury resulting in a condition where patients are asymptomatic, at least when young, and have only a slight reduction in hemoglobin level [3]. Based on these observations, the achievement of 10 pg of HbF in nearly all sickle erythrocytes becomes the ‘holy grail’ of HbF-based therapeutics as it would result in the clinical ‘cure’ of disease.

Complications of sickle cell anemia can be divided into those associated with sickle vasoocclusion and those that are a consequence of the intravascular hemolysis of varying numbers of sickle erythrocytes (Figure 1) [4,5]. Naturally occurring or therapeutically induced high levels of HbF are strongly associated with fewer of the complications that have been associated with sickle vasoocclusion and blood viscosity, like acute painful episodes, acute chest syndrome, and osteonecrosis. With the exception of leg ulceration, there is little association of HbF levels with complications closely associated with the severity of intravascular hemolysis, like cerebral, pulmonary and systemic vasculopathy and nephropathy. Lysis of

Figure 1. Pathophysiology of sickle cell anemia. The sickle hemoglobin mutation (top left), codes for a β-globin gene variant, which as with normal HbA, is in solution when oxygenated. In distinction to HbA, HbS polymerizes when deoxygenated (center left). Polymerization of deoxyHbS initiates the pathophysiology of disease by injuring the sickle erythrocyte. This leads to a heterogeneously red cell population that includes adhesive cells, dense cells and irreversibly sickled cells (bottom left). These abnormal red cells damage blood vessels and cause the vasoocclusive and hemolytic components of the disease. Sickle vasoocclusion (top right) provokes reperfusion injury and inflammation. It is associated with disease sub-phenotypes like acute pain episodes, acute chest syndrome, and osteonecrosis. A variable portion of hemolysis occurs intravascularly (bottom right). Intravascular hemolysis depletes haptoglobin and hemopexin and liberates free heme and arginase into the plasma consuming nitric oxide leading to platelet activation, endothelial damage, and inflammation. The severity of intravascular hemolysis is associated with the sub-phenotypes of pulmonary and systemic hypertension, stroke and nephropathy.
cells that contain insufficient HbF to protect them from HbS polymer-provoked injury liberates hemoglobin and arginase into the circulation depleting bioavailable nitric oxide (NO) and triggering proliferative vasculopathy [6].

HbF levels and the highly correlated number of F-cells are heritable [7–9]. One to 2 months following birth the normal switch from HbF to adult hemoglobin is nearly complete with <1% HbF remaining. Most, but not all of the normal variance in HbF gene expression is regulated by 3 quantitative trait loci (QTL). One QTL is linked to HBB on chromosome 11; the other 2 QTL are trans-acting and are encoded on chromosomes 2 and 6.

2.1.1. Cis-acting regulation

Five common haplotypes of the HbS gene, defined by polymorphisms or SNPs cis to HBB, are associated with different HbF levels and disease phenotypes [10–12]. Patients with the Arab-Indian (AI) and Senegal haplotype have the mildest disease; Bantu or Central African Republic (CAR) haplotype patients have the most severe disease; Benin and Cameroon haplotype patients are intermediate. The association of a haplotype with disease severity correlates with the HbF level typical of each haplotype. In adult haplotype homozygotes, levels of HbF average about 20% in the AI, 10% in the Senegal and 7% to 4% in the Cameroon, Benin and Bantu haplotypes. The mechanism underlying differential HbF expression among haplotypes is unclear and is most likely to be a cis-acting effect. In children with the AI haplotype HbF levels average about 30%. These children can have a mild disease but as they age and HbF levels fall to 15% to 20% the disease becomes more severe [13,14]. The AI haplotype has unique cis and trans-acting variants but whether or not they are responsible for the high HbF is unknown [15–19]. A C-T polymorphism (rs7482144) 158 bp 5’ to HBG2 (~158) in the proximal promoter of this gene creates a cleavage site for the restriction endonuclease Xmn1. This SNP is present only in the AI and Senegal haplotypes. It is associated with increased expression of only HBG2 with a corresponding increase only in $\gamma$-globin. There is little evidence that rs7482144 is functional. Recent studies of BCL11A binding in the HbF gene promoters do not support a mechanistic role for rs7482144 suggesting that it is in linkage disequilibrium with the functional cis-acting element of these haplotypes [20].

Ascertaining HbS-associated haplotypes has been useful epidemiologically and anthropologically but their prognostic relevance in individuals is minimal.

BCL11A binds TGACCA motifs present at 35 sites within the HBB gene cluster, 2 of which are in the $\gamma$-globin gene promoters. The distal of these 2 sites at positions –118 to –113 is the locus of 2 point mutations and a 13 bp deletion associated with the phenotype of HPFH [21]. BCL11A binds preferentially to this site in adult erythroid progenitors. Its occupancy in the –118 to –113 motif represses this promoter and favors locus control region (LCR) interactions with $\beta$-globin gene promoters. A BCL11A binding motif is not present in the region surrounding the –158 Xmn1 cleavage site [22,23].

Other cis-acting elements with putative roles in HbF gene expression were located within the HBD-HBG1 intergenic region, in the HBB LCR hypersensitive site-2 core, ~530 bp 5’ to HBB and in the olfactory gene cluster upstream of the LCR. An additional candidate region was a 3.5 kb element near the 5’ portion of HBD but this site is devoid of BCL11A binding sites. Regions remote from the HbF genes are less likely to have major roles in switching from fetal to embryonic to adult hemoglobins.

2.1.2. Trans-acting regulation

The breadth of HbF levels within individual haplotype groups implies that trans-acting factors interacting directly or indirectly with promoters and enhancers have essential roles in $\gamma$-globin gene expression. MYB (6q23.3) and BCL11A (2p16.1) are the 2 trans-acting QTL whose polymorphisms are associated with HbF levels and whose mechanisms of action vis-à-vis $\gamma$-globin gene expression are understood in some detail [24,25]. Both of these QTL affect HBG2 and HBG1 expression. MYB and BCL11A encode silencers of HbF gene expression. Their expression levels are regulated by polymorphisms in their enhancers. BCL11A works directly to repress HbF gene expression; MYB works indirectly through KLF1 and by its effects on hematopoiesis.

MYB regulates the proliferation and maturation of erythroid cells and gene expression within the HBB gene cluster [24,26]. A 3 bp deletion polymorphism (rs66650371) is the probable functional variant of this QTL that affects $\gamma$-globin gene expression [27]. This SNP is highly associated with HbF in multiple populations. In its proximity are binding sites for multiple erythropoiesis-related transcription factors and it is in a locus with enhancer-like activity [28,29]. Downregulation of a long noncoding RNA, transcribed from this enhancer increased $\gamma$-globin gene mRNA 200-fold [30]. The frequency of rs66650371 in African and Saudi populations is low compared with its frequency in normal Europeans or Chinese with $\beta$ thalassemia reducing the effect of this variant on HbF in sickle cell anemia.

BCL11A, chromosome 2 QTL, encodes a zinc-finger protein that represses the $\gamma$-globin genes by binding their promoters, as discussed above. The genetic association of BCL11A with HbF levels was proven in studies of multiple cohorts of normal individuals and patients with sickle cell anemia and $\beta$ thalassemia [25,31]. BCL11A favorably modified the features of both diseases because of its effects on HbF concentration [32,33]. Sentinel SNPs marking the effects of BCL11A on HbF were located in the second intron of this gene in an erythroid-specific gene enhancer. The high HbF-associated SNP is common and also has a large effect size [34,35]. High-resolution studies of BCL11A binding sites in the HBB gene cluster suggested that promoter repression was the major mechanism of action of this protein that controls most of the hemoglobin switching.

2.1.3. HbF distribution in F-cells

It is likely that the distribution of HbF concentrations amongst the F-cell population, which differs amongst individuals with similar HbF levels, has a greater impact on the disease than
the total amount or percent of HbF or the number of F-cells measured by conventional FACS. For near-absolute protection of the sickle erythrocyte from polymer-induced damage, 9 to 10 pg. of HbF/F-cell is required. Lesser amounts, while likely to be helpful, can leave the cell prone to polymer-induced injury. Such heterogeneity of HbF concentrations among F-cells might explain why individuals with similar total HbF levels, even as high as 20% to 30%, can have very different clinical and hematologic findings. Therefore, any of the available HbF measures cannot foretell the likelihood of sickle cell-related complications in an individual [36]. Although not yet tested experimentally because of the absence of a widely available assay, knowing the distribution of HbF concentrations amongst F-cells is likely to be the best way of gauging the protective effects of HbF, developing an accurate prognosis and be the most useful way of determining the efficacy of HbF-induction therapeutics.

2.1.4. Genetic risk scores for HbF

Even though imperfect, knowing at birth the HbF level that will be present after the γ- to β-globin switch is complete might allow a better patient-specific prognosis and also guide an individualized approach to HbF-induction therapy. An ensemble of genetic risk scores used to predict HbF in patients with sickle cell anemia explained more HbF variability and had a higher predictive accuracy compared with single SNP analyses. The risk scores used as few as 14 SNPs, mostly in the 3 HbF QTL with a few on several other chromosomes, predicting HbF in 3 cohorts independent of the training cohort. Although 23.4% of the variability in HbF was explained, the heritability of HbF levels is far greater. The extension of this work to larger cohorts with the inclusion of SNPs detected by whole genome sequencing might increase the predictive value of this approach [37]. Another genetic prediction model used 4 SNPs (rs6545816, rs1427407, rs66650371, rs7482144) in the 3 HbF QTL to account for 21.8% of HbF variability [38].

Collectively, the 3 HbF QTL explain about one-third to half of HbF variation within populations but the genotypes of these QTL cannot predict HbF in an individual. HbF QTL also cannot predict the HbF response to hydroxyurea whose major mechanism of action is to induce increased HbF levels. Hemolysis, erythroid marrow expansion, and stress erythropoiesis also affect HbF levels by perturbing the kinetics of erythropoiesis.

3. α Thalassemia

α Thalassemia (MIM141850/141800) due to the deletion of 1 or 2 α-globin genes is present in about one-third of most sickle cell anemia populations of African descent and in half of the Middle Eastern and Indian patients. The most common α-thalassemia variant is the &alpha;3.7 α-globin gene deletion. By reducing the intracellular concentration of HbS, α-thalassemia decreases HbS polymer-induced cellular damage thereby attenuating hemolysis [39]. The hematologic and laboratory changes in sickle cell anemia-α thalassemia include; increased erythrocyte lifespan, higher hemoglobin concentration, lower MCV, lower reticulocyte count, lower bilirubin level, lower LDH level and fewer dense and irreversibly sickled cells (ISCs). The magnitude of these changes is related to the number of deleted α-globin genes. Coincident α thalassemia is usually associated with fewer complications that have been associated with intravascular hemolysis [40]. The likelihood of developing complications associated with sickle vasoocclusion is increased. This has been attributed to increased blood viscosity resulting from decreased hemolysis and higher hemoglobin level.

4. Genetic markers associated with clinical and hematologic subphenotypes

4.1. Pain

Pain in sickle cell anemia has a complex and multifactorial etiology. Acute and chronic pain often coexist and can be difficult to separate. Neuropathic pain and opioid-induced hyperalgesia are also common. Impression defining pain phenotypes and the huge impact of social and environmental factors makes it very hard to study pain in genetic association studies. This is reflected in the underwhelming results of both candidate gene and genome-wide association studies (GWAS) of pain, most of which have not been replicated.

GCH1 encodes a GTP cyclohydrolase that is rate-limiting for tetrahydrobiopterin (BH4) synthesis. BH4 is a cofactor for NO synthases. The most robust genetic association with sickle cell pain is with variants of GCH1. In a discovery cohort of 228 patients and replication cohort of 513 patients, rs8007267 in the GCH1 promoter was a significant risk factor for acute painful episodes. Subjects homozygous for the risk allele produced more BH4. Physiological studies of traits likely to be modulated by GCH1 showed that rs8007267 was associated with altered endothelial-dependent blood flow in females. Association with pain was limited to females and attributable to a distinct African haplotype [41]. In 131 additional patients, rs8007267 in GCH1 was associated with acute care utilization rate and with chronic pain [42].

4.2. Cerebrovascular disease

Stroke is a sub-phenotype that can be more precisely ascertained than pain lending itself to more convincing genetic association studies. In 130 stroke patients and 103 controls with sickle cell anemia, in addition to the known association of gene deletion α thalassemia with reduced stroke risk, SNPs in ANXA2, TEK, ADCY9, and TGFB3 were associated with the risk of stroke. The model predicted the correct outcome for 7 additional individuals with stroke and for 105 of 107 subjects without stroke, an overall predictive accuracy of 98% [43]. Partial replication of this association was confirmed in an independent cohort of 130 patients with thrombotic stroke and 103 patients without stroke where ANXA2, TGFB3, and TEK were associated with increased risk and a thalassemia and ADCY9 were associated with decreased stroke risk [44]. Other studies have associated different SNPs in VCAM1 with both increased and decreased stroke risk [45,46].
4.3. Nephropathy

As sickle cell disease population age, renal failure, culminating in dialysis or renal transplantation, becomes more common. Albuminuria is a harbinger of sickle nephropathy. Early detection of albuminuria is an indication for the use of inhibitors of the renin-angiotensin system and hydroxyurea that can prevent advancing renal disease. Homozygosis or compound heterozygosity for the G1 (rs73885319) or G2 (rs71785313) alleles of APOL1 were associated with albuminuria (P = 0.02) and shortened the time to first developing albuminuria (P = 0.0003) [47]. Screening for variants of APOL1 before albuminuria is detectable might identify patients for early renoprotective treatment.

4.4. Other clinical sub-phenotypes

Multiple sub-phenotypes were associated with several genes of the large TGF-β (transforming growth factor-β)/Smad/BMP (bone morphogenetic protein) pathway that regulates diverse cellular processes important in disease pathophysiology [48]. These pathways include inflammation, fibrosis, cell proliferation and hematopoesis, osteogenesis, angiogenesis, wound healing and the immune response. Genes outside of the globin gene clusters that have been associated with sub-phenotypes of sickle cell disease or with hematologic changes are shown in Table 1.

Many of these studies have limited prognostic value because of small sample sizes, imprecision of phenotype definition, small effects of the variant on the phenotype, gene–environment interactions and the multigenic nature of most clinical sub-phenotypes.

4.5. Erythrocyte volume

Sickle erythrocytes are heterogeneous. Low-density cells include reticulocytes and high-density cells, including the iconic ISC, can have a mean corpuscular hemoglobin concentration (MCHC) of 50g/dL. HbS polymerization, is affected by the MCHC. The Gardos channel (KCNN4) and K-Cl cotransport channels (KCC1, KCC3, KCC4) have major roles maintaining red cell hydration in sickle cell anemia and therefore impact HbS polymerization; however, variants in these genes have not been associated with a red cell phenotype [49].

A gain-of-function variant in the gene PIEZO1 (E756del; rs572934641) that codes for a mechanosensitive and deoxygenation-activated cation transport channel, or P<sub>π</sub>ickle, is present in one-third of people of African descent [50]. In 788 patients with sickle cell anemia or HbSC disease, there was no correlation between E756del or with cation leak measured by P<sub>π</sub>ickle and hemoglobin, reticulocyte count, MCHC, or hospital admissions [51]. Based on whole exome sequencing in 226 sickle cell disease patients, the E756del allele was associated with increased cell density, a result replicated in 375 additional patient samples. However, there was no association with markers of hemolysis and renal function or with priapism and leg ulcers.

4.6. Hemolysis

Intravascular hemolysis can be assessed using a principal component analysis that includes reticulocyte count, LDH, AST and bilirubin levels that computes a hemolytic component. When this estimate was used as a phenotype in a GWAS, hemolysis was associated with rs7203560 in NPR3L3. The HBA2/HBA1 regulatory elements, termed HS-48 [R1], HS-40 [R2] and HS-33 [R3], are located in introns of NPR3L3. They act independently and additively to regulate α-globin gene expression. It was hypothesized that α-globin gene regulatory loci tagged by rs7203560 downregulated the expression of the α-globin genes [52]. Two variants, rs11185131 and rs11248850, upregulate α-globin gene expression [53]. Unlike rs7203560, these variants were not associated with hemolysis. Adjusting for the effects of these SNPs did not change the association of rs7203560 with hemolysis. However, in patients with sickle cell anemia and concurrent gene deletion a thalassemia, rs111248850 and rs11185131 nullified the hemolysis-reducing effect of a thalassemia [54]. Regulatory elements of the α-globin gene cluster can up- or

| Table 1. Genes outside of the globin gene clusters associated with sub-phenotypes of sickle cell anemia. These reports include primarily HbS homozygotes. Most studies are small and their results should be considered exploratory as they often do not correct for co-existent factors. Mechanistic and functional studies to follow-up on these correlation analyses are nearly always lacking. References are in the text and in [48]. |
|---|---|
| Sub-phenotypes | Genes and effects |
| Survival | Multiple genes, including TGFBR3 |
| Stroke, silent infarction | Multiple genes increase or decrease likelihood, VCA1M, ILR4, ADRB2, HLA, LDLR, ENP1, GOLGB1 |
| Acute and chronic pain | GCCH1, KIAA1109, ADRB2, NR3C1, MBL2, CACNA2D3, DRD2, KCNS1, COMT, FAH, OPR1M1, ADRB2, UGT2B7, FAM193A, PLA2G4A, IL1A, GLA5L3 |
| Acute chest syndrome | Genes have been identified, e.g., COMMD7, HMOX1, NOS1, VEGFA, but few studies have been validated. |
| Bacteremia/Infection | MBL2-contradictory evidence in different populations that that low level protective. Other genes include CCL5, various |
| Osteonecrosis | MTHFR (weak evidence); BMP6-results validated in 2 different populations |
| Priapism | KL, TEK, TGFBR3, AQP1 |
| Leg ulcers | TGF-β/Smad/BMP pathway, KL possibly HLA alleles |
| Sickle vasculopathy/TRV | BMP6, TGFBR3, ACVR1, BMP2, THBS1, DRD2 |
| Cholelithiasis | Promoter repeats in UGT1A1 associated with serum bilirubin and gallstones |
| Renal function/albuminuria/glucomerolar hyperfiltration | DARC Fy- associated with proteinuria, TGF-β/Smad/BMP pathway, MYH9, APO1, HMOX1, AGGFI, CYP4B1, TOR2A, PKD1L2, CD163, |
| Multiple subphenotypes | Duffy antigen receptor (DARC) No relationship to leg ulcers, nephropathy, priapism, osteonecrosis, response to opioids |
| Erythrocyte density | ATP2B4, PIEZO1 increase cell density (see text) |
| Alloimmunization | HLA locus, chr5 rs7585387, FCGR2C, CTLA4 |
| Hemolysis | HBA2/1 deletions, NPR3L3, ADcy6 improve red cell survival |
down-regulate gene expression; variants of these elements might account for some of the phenotypic heterogeneity of sickle cell anemia.

### 4.7. Inflammation

Inflammation in sickle cell anemia is a byproduct of the foundational pathophysiology of the disease. Polymerization and depolymerization of HbS induce multiple loci of erythrocyte damage leading ultimately to vasoocclusion and intra and extravascular hemolytic anemia. These intertwined pathophysiologies cause tissue damage and are responsible for features of the disease ‘downstream’ of HbS polymerization. Ischemia-reperfusion cycles along with the reduction of anti-inflammatory effects of NO due to its consumption by cell-free hemoglobin lead to the activation of leukocytes, platelets, and endothelial cells. A recent review thoroughly explores inflammation in sickle cell disease and how this might be countered therapeutically [55]. Studies have attempted to link genetic variants in many genes in inflammatory pathways to complications of sickle cell anemia. None are definitive; most are very small; few have used unbiased genome-wide approaches. Despite pathophysiologic relevance and numerous points for therapeutic intervention the complexity of the biological pathways involved in inflammation coupled with the small individual contribution of the many cellular elements and genes of these pathways has hindered the development of biomarkers useful for prognosis or therapeutics.

### 5. Complex approaches to disease severity modeling

A single clinical complication of sickle cell anemia or an isolated laboratory measurement has limited predictive value for defining disease severity or the likelihood of near-term death. The exception to this is the association of tricuspid regurgitant velocity, or TRV with mortality. A predictive model of disease severity used Bayesian networks to integrate biomarkers with disease complications to define an overall measure of disease severity predicting near-term death [56,57]. Although not designed to predict complications, the model is useful to understand the relationships between clinical and laboratory measures of the disease. Along with previously known risk factors for mortality, like renal insufficiency and leukocytosis, the network identified markers of the severity of hemolytic anemia and its associated clinical events as contributing risk factors. It can compute a personalized disease severity score that might be used to guide therapeutic approaches. Signatures based on a combination of blood mononuclear cell gene expression profiles and medical history have also been used to predict mortality [58].

Between a genetic variant and a physiologic outcome – or genotype and phenotype – are proteins that can be measured directly and whose effects can also be captured by some laboratory measures. A system-type approach was used to discover profiles of multiple, common biomarkers that correlated with morbidity and mortality using cluster analysis. Seventeen signatures of 17 common circulating biomarkers were found in more than 2300 patients and partially replicated in 2 additional cohorts. Using longitudinally collected data, some of the signatures were associated with reduced risk while others were associated with increased risk for complications like stroke, pain, leg ulceration, acute chest syndrome, avascular necrosis, seizure, death, HbF level, hemolysis and biomarkers of pulmonary vasculopathy. Such signatures could become an important and affordable precision medicine tool to aid the treatment and management of the disease [59].

### 6. Disease modeling with induced pluripotent stem cells (iPSCs)

iPSC technology has altered our approach to disease modeling, cell-based therapeutics and precision medicine. iPSCs are capable of self-renewal and differentiation into tissues of any germ layer. They can capture the genetic background of the subject from whom they are created. While functional equivalents of embryonic stem cells they can be used without their corresponding ethical and technical barriers. A virtually unlimited supply of target cells can be generated by first reprogramming a small number of adult somatic cells obtainable by venipuncture toward pluripotency and infinite self-renewal by the forced expression of four key transcription factors, Oct4, Sox2, Klf4, and c-Myc. This is followed by directed differentiation into the desired cell type [60,61].

Critical for sickle cell disease, progress has been made in developing differentiation protocols that mimic the natural sequence of hematopoietic progenitor development for the generation of iPSC-derived erythroid-lineage cells. These ever-improving iPSC-based platforms allow for the functional testing and mechanistic study of genetic variants that may play a role in HbF regulation. An example of this is shown in Figure 2 for a library of sickle iPSCs derived from patients of different genetic backgrounds. iPSCs offer an unlimited supply of cellular material making them ideal for screening novel therapeutics in a patient-specific manner. Usually, in vitro testing that precedes clinical trials is performed using immortalized cell lines, which have been subject to genetic alterations, or patient CD34+ cells whose supply is limited, potentially compromising the fidelity of drug screens. Promising results in immortalized cell-based assays and animal models might not always translate to the clinic [62]. iPSC technology offers the prospect of using primary cells linked to the patient genetic background to obtain toxicity and efficacy data well before significant investment in human clinical trials [63,64]. This is particularly useful when developing new therapies for sickle cell disease with its highly variable clinical presentations and equally variable therapeutic responses.

A diverse collection of sickle cell disease patient-derived iPSCs capturing a range of genetic backgrounds and phenotypes has been assembled, banked and is available for distribution. This iPSC library used patient samples from four distinct geographic locations that encompass the four common HbS gene haplotypes [65]. Their use might help further our understanding of variable responses to existing treatments and could better inform future treatment recommendations (Figure 2).
iPSCs are not perfect systems. A current challenge is differentiating iPSCs into progeny that accurately resemble their naturally-occurring counterparts. iPSC-derived erythroid cells retain certain fetal/embryonic characteristics such as their inability to synthesize β-globin levels that approximate those in post-natal erythrocytes. They have levels of HbF and embryonic globins that are much higher than adult cells. The efficiency of differentiation varies slightly with different protocols, but all seem to result in the formation of developmentally stunted nucleated erythrocytes [66]. The latest advances in differentiation platforms have addressed these limitations by generating adult-type progeny from iPSCs via control of the Notch and the aryl hydrocarbon receptor signaling pathways [64]. Recently, a β-globin reporter iPSC line was established as a tool to map real-time β-globin gene expression during erythropoiesis at a single-cell resolution [67]. In conjunction with single-cell RNA sequencing, the study of this line found that definitively patterned iPSC-derived erythroblasts show similar gene expression patterns as their naturally occurring counterparts, verifying the utility iPSCs bestow upon disease modeling. Such advances will ultimately enable the use of specific iPSC libraries to create more relevant disease models and to study the potential impact of novel therapeutics on globin gene expression in adult-type blood cells.

In the 10 years since the discovery of iPSCs, the field has moved forward rapidly. Although the biologic insights from this work have been impressive, the application to human therapeutics has been minimal. Notably, the use of iPSC-derived erythroid progenitor cells as a potential cell-based therapeutic for sickle cell anemia would take advantage of 2 impactful characteristics of how the cells are currently produced: the ability to make unlimited numbers of isogenic, iPSC-derived erythroblasts and the fetal/embryonic hemoglobin phenotype of these cells. Isogenic cells should not evoke an immune response. Cells producing predominantly HbF will prevent polymerization of any HbS present while adequately transporting oxygen. Thus, without manipulation of the genome or chronic immunosuppression, a potentially curative cellular therapy could be manufactured centrally and be available for reinfusion in any infusion center.

Figure 2. Clinical trial in a test tube. A. Global distribution of sickle cell disease and the 5 common haplotypes of the HbS globin gene. As detailed in the text, HbF levels are lowest in the Bantu haplotype, highest in the Arab-Indian haplotype and intermediate between these in the other haplotypes. B. Reprogramming somatic peripheral blood cells from patients representative of these haplotypes allows their differentiation to erythroid precursors that can reflect the HbF level of native erythrocytes. These cells might be used to test the effects of HbF-inducing therapeutics. The heat map at the lower right depicts possible outcomes of such a screen in terms of HbF levels. This output can be used to guide treatment recommendations.
7. Conclusions

Disease-modifying treatment is now limited to hydroxyurea. Matched sibling hematopoietic stem cell transplantation is applicable to only 10% to 20% of the patients because of constrained donor availability. Starting hydroxyurea in the first year of life at an average dose of 27 mg/kg was associated with HbF levels of 33.3 ± 9.1% about one-third of patients had HbF >40% and a hemoglobin concentration of 10.1 ± 1.3 g/dL—about one-third had hemoglobin concentrations >11 g/dL—without toxicity, with a marked reduction in sickle cell-related events [68]. If these HbF levels are sustained, both the vasoocclusive and hemolytic components of disease should improve. However, we know little about the long-term toxicity of this agent when it is started in infancy and continued for life.

Two other drugs, which like hydroxyurea should be disease-modifying based on their mechanisms of action, could soon be available but their approval will be based on short-term clinical trials. Nothing is known about the long-term effects of these agents, especially when they are started in early life, the optimal time to begin disease-modifying treatment. A third drug, L-glutamine was FDA approved (Endari®) in 2017 to prevent acute vasoocclusive events. Its mechanism of action is ill-defined. the benefits marginal and its potential for modifying the course of disease unknown [69,70].

7.1. Voxelotor

Voxelotor is an orally available agent that increases hemoglobin-oxygen affinity by binding to the α-globin chain and preventing HbS from assuming the deoxy conformation that is capable of polymerization. Clinical trials have shown that treatment reduced markers of hemolysis and hemoglobin concentration increased about 1 g/dL in more than half of the patients [71]. In this short-term trial, the effects of this agent on sickle vasoocclusion were inconclusive. Based on our current understanding of the pathophysiology of sickle cell anemia, a reduction in intravascular hemolysis should, over time, decrease the risk of stroke, pulmonary and systemic vasculopathy, nephropathy and mortality. However, because its mechanism of action involves an increased affinity of hemoglobin for oxygen it decreases tissue oxygen delivery and the potential exists for untoward effects, especially in a developing central nervous system [72].

7.2. Crizanlizumab

Crizanlizumab is a P-selectin blocking monoclonal antibody that impedes sickle erythrocyte adherence to endothelium; it does not retard HbS polymerization. When given by monthly intravenous infusion it was associated with a 50% reduction in acute vasoocclusive events [73]. This agent has no effect on HbF or chronic hemolysis but should modify the course of the disease by improving blood flow. Although safe in its initial trials, long-term toxicity of selectin blocking is unknown.

7.3. Gene therapies and haploidentical bone marrow transplantation

Also, likely to be available in the future are different forms of gene therapy and the more widespread use of haploidentical hematopoietic stem cell transplantation. Both modalities could be ‘curative’ [31,74,75].

Lentiviruses are being used to insert into patient CD34+ hematopoietic progenitor cells a β-globin gene that contains a HbS polymer-inhibiting amino acid substitution or an RNAi that inhibits the expression of BCL11A. CRISPR/Cas and zinc finger nucleases are being employed to disrupt the BCL11A enhancers or the BCL11A binding sites in the γ-globin gene promoters. It is ‘early days’ for therapeutic genetic manipulation some of which involve random insertion or deletion of DNA. The ultimate risks and long-term benefits of these treatments, which all involve myeloablative bone marrow conditioning are unknown. New methods of preparing the hematopoietic niche to accept transplanted stem cells and alternative approaches to procuring additional stem cell donors are actively being explored [76,77].

The current pace of therapeutic advances prioritizes the need for accurate prognostic tools as the possibility for drug therapy with multiple agents, expanded use of stem cell transplantation and multiple approaches to gene therapy approach clinical reality.

8. Expert opinion

Foretelling the type and severity of the complications of sickle cell anemia should help the clinician choose the most efficacious individualized therapy. Currently, except for hematopoietic stem cell transplantation from an HLA-matched sibling donor, a modality available to the very few, the course of this disease is beneficially affected only with hydroxyurea treatment that is recommended for nearly every patient regardless of symptoms. The next 5 years is likely to see the approval of at least two new drugs that, like hydroxyurea, might be disease-modifying. In addition, cell-based gene-based therapy, either by lentiviral introduction of an anti-sickling globin into patient stem cells or by genomic editing in erythroid precursor cells that reduces the repression of HbF gene expression, might also become available. New drug treatments will be used in a polypharmaceutical approach compounding the possibility of adverse effects; gene therapy, although potentially curative, will not be risk-free. Compared with hydroxyurea, all these therapies will be incredibly expensive. This change in the therapeutic landscape makes more imperative the quest for individualized treatment regimens based on genetic, cellular or clinically based approaches.

Currently, treatment monitoring is inadequate. Better means are needed to tailor therapy to the individual. One hope is that whole genome sequencing will provide more complete genetic information that can be used for genetic risk scores, which could be available at birth, if not before. Other non-genetically based methods of individual prognostication and focused treatment can also be developed. When monitoring treatments designed to increase HbF gene expression, HbF levels and F-cell analysis by flow cytometry provides only an approximation of possible benefit as neither assay measures the distribution of HbF concentrations.
among F-cells. This distribution is likely to be a critical measurement enumerating the proportion of sickle erythrocytes that are protected from polymer-induced damage. To establish the likely benefit of a drug or cell-based HbF-induction therapy a rapid, flow cytometry-based assay that can provide an estimate of the distribution of HbF concentrations among F-cells would be ideal. For agents like crizanlizumab that improve blood flow an ideal would be a validated, replicable and physiologically relevant in vitro assay, perhaps employing a microfluidic chamber where patient cells can be repeatedly tested during treatment. Some agents, like voxelotor, improve hemolysis. A test to foretell the level of hemolysis suppression that is needed to prevent the common vasculopathy complications of disease would be a valuable tool to judge an individual’s likely response to treatment. This might be approached using the hemolytic component that is computed using reticulocyte count, LDH, bilirubin, and AST, as a quantitative measure of vasculopathy and tricuspid regurgitant velocity (TRV) as an outcome measure to derive an age-adjusted estimate of the likelihood of increased TRV according to the hemolytic component. Prospective long-term longitudinal studies involving large precisely phenotyped patient cohorts with multi-omic analyses will be needed to document the replicability and validity of a 21st century precision medicine-based treatment paradigm.

**Funding**

This work was supported in part by the NIH, NHLBI under Grants U01 HL107443; R01 HL 87681; R01 HL 068970; RC2 HL101212.

**Declaration of interest**

The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

**Reviewers Disclosure**

Peer reviewers on this manuscript have no relevant financial relationships or otherwise to disclose.

**ORCID**

George J. Murphy  
http://orcid.org/0000-0003-3464-793X

**References**

Papers of special note have been highlighted as either of interest (†) or of considerable interest (‡) to readers.

This study shows that proper dosing of hydroxyurea in the

Stadtholders R, Aktuna S, Thongjuea S, et al. HBSS1-MYB intergenic
variants modulate fetal hemoglobin via long-range MYB enhancers.

Canver MC, Lessard S, Pinello L, et al. Variant-aware saturating
mutagenesis using multiple Cas9 nucleases identifies regulatory

the HBSS1-MYB intergenic region on chromosome 6q23 is associated

Vijnamur DS, Bauer DE, Orkin SH. Recent progress in understanding
and manipulating haemoglobin switching for the

Lettre G, Sankaran VG, Bezerra MAC, et al. DNA polymorphisms at
the BCL11A, HBSS1-MYB, and beta-globin loci associate with fetal
hemoglobin levels and pain crises in sickle cell disease. Proc Natl

shows BCL11A associated with persistent fetal hemoglobin and
amelioration of the phenotype of beta-thalassemia. Proc Natl

Bauer DE, Kamran SC, Lessard S, et al. An erythroid enhancer of
BCL11A subject to genetic variation determines fetal hemoglobin

Canver MC, Smith EC, Sher F, et al. BCL11A enhancer dissection by
Cas9-mediated in situ saturating mutagenesis. Nature. 2015;527
(7577):192–197.

Steinberg MH, Chui DHK, Dover GJ, et al. Fetal hemoglobin in sickle

In this paper, the arguments for the importance of the distribution
of HbF concentrations amongst F-cells as a determinant
for young children with sickle cell anemia. Am J Hematol.

Taylor JG, Nolan VG, Mendelsohn L, et al. Chronic hyper-hemolysis
in sickle cell anemia: association of vascular complications and

Belfer I, Youngblood V, Darbari DS, et al. A GCH1 haplotype confers
sex-specific susceptibility to pain crises and altered endothelial
function in adults with sickle cell anemia. Am J Hematol. 2014;89

Jhun EH, Sadhu N, Yao Y, et al. Glucocorticoid receptor single
nucleotide polymorphisms are associated with acute crisis pain in

Sebastiani P, Ramoni MF, Nolan V, et al. Genetic dissection and
prognostic modeling of overt stroke in sickle cell anemia. Nat Genet.

stroke in children with sickle cell anemia. Blood. 2011;117

Taylor JG, Tang DC, Savage SA, et al. Variants in the VCA1M1 gene
and risk for symptomatic stroke in sickle cell disease. Blood.
2002;100(13):4303–4309.

Hoppe C. Gene interactions and stroke risk in children with sickle

cell anemia and APOL1 genetic variants develop albuminuria early

Steinberg MH. Clinical variability in sickle cell anemia. 2019.
UpToDate. Walters Kluwer, Waltham, MA.

study of erythropoietin density in sickle cell disease patients. Blood

populations causes RBC dehydration and attenuates Plasmodium

Rooks H, Brewin J, Gardner K, et al. A gain of function variant in
PIEZ1 (E56del) and sickle cell disease. Haematologica. 2019;104
(3):e91–e93.

Milton JN, Rooks H, Drasar E, et al. Genetic determinants of hae-

modify hematologic and other clinical phenotypes in sickle cell

Milton JN, Shaikhho EM, Steinberg MH. Haemolysis in sickle cell
anemia: effects of polymorphisms in α-globin gene regulatory

Conran N, Belcher JD. 2018; Inflammation in sickle cell disease. Clin

Sebastiani P, Nolan VG, Baldwin CT, et al. A network model to
predict the risk of death in sickle cell disease. Blood. 2007;110
(7):2727–2735.

One of the few studies to derive a clinically relevant severity
index for sickle cell disease.

Belini Junior E, Silva DGH, de Torres LS, et al. Severity of Brazilian
sickle cell disease patients: severity scores and feasibility of the
Bayesian network model use. Blood Cells Mol Dis. 2015;54

Desai AA, Leitz, Bahroos N, et al. Association of circulating tran-
scriptomic profiles with mortality in sickle cell disease. Blood.
2017;129(22):3009–3016.


Takahashi K, Yamanaka S. Induction of pluripotent stem cells from
mouse embryonic and adult fibroblast cultures by defined factors.
Cell. 2006;126(4):663–676.

Takahashi K, Yamanaka SA. Developmental framework for induced

Deshmukh RS, Kovacs KA, Dinnyés A. Drug discovery models and
toxicity testing using embryonic and induced pluripotent
cell-derived cardiac and neuronal cells. Stem Cells Int.
2012;2012:379569.

Chun YS, Byun K, Lee B. Induced pluripotent stem cells and perso-
nalized medicine: current progress and future perspectives. Anat

Park S, Gianotti-Sommer A, Molina-Estevez FJ, et al. A comprehensive,
ethnically diverse library of sickle cell disease-specific induced pluripotent stem cells. Stem Cell Rep.

Vanuytsel K, Matte T, Leung A, et al. Induced pluripotent stem cell-
based mapping of β-globin expression throughout human ery-

Leung A, Zulick E, Skvir N, et al. Notch and aryl hydrocarbon
receptor signaling impact definitive hematopoiesis from human

Kola I, Landis J. Can the pharmaceutical industry reduce attrition

McGann PT, Niss O, Dong M, et al. Robust clinical and laboratory
response to hydroxyurea using pharmacokinetically guided dosing
for young children with sickle cell anemia. Am J Hematol. 2019;94

This study shows that proper dosing of hydroxyurea in the
first year of life can have profound clinical benefit.

Quinn CT. l-Glutamine for sickle cell anemia: more questions than

Nihara Y, Miller ST, Kanter J, et al. Investigators of the phase 3 trial
of l-glutamine in sickle cell disease, a phase 3 trial of l-glutamine in

**An excellent review by leaders in the field of the possibilities of gene therapy in hemoglobinopathies**