Sickle Cell Disease: A Multigenic Perspective of a Single-Gene Disorder

Abdullah Kutlar
Medical College of Georgia, Augusta, Ga., USA

Key Words
Sickle cell disease • Genetic polymorphisms

Abstract
Although significant progress has been made in our understanding of sickle cell disease (SCD) and in the development of new therapies, many questions are still unanswered, and a cure remains elusive. This is particularly evident in the clinical heterogeneity of the disease. Studies have shown the importance of high hemoglobin F determinants and α-thalassemia as modifiers of disease severity, but these alone do not explain the diversity that is seen. This paper focuses on recent advances on the effect of nonglobin genetic modifiers on the SCD phenotype. The roles of polymorphic variants of (1) methylene-tetrahydrofolate reductase gene in the pathogenesis of avascular necrosis, (2) factor V R485K and risk of venous thrombosis, and (3) UDP glucuronosyltransferase-1 polymorphism on serum bilirubin levels in SCD are discussed. Mention is made of genetic polymorphisms that might predispose to stroke. The application of gene expression profiling to the study of SCD is very promising and some preliminary data are provided.
the disease has provided further impetus to search for other genetic modifiers of SCD.

The well-known consequences of the sickle mutation and its downstream effects are (1) a chronic hemolytic anemia, (2) episodic vaso-occlusion with resultant characteristic painful episodes, and (3) chronic organ damage. The clinical phenotype of SCD is highly variable; complications such as acute chest syndrome, stroke, leg ulcers, and avascular necrosis (AVN) do not occur in all of the patients, suggesting that factors beyond the β6 Glu→Val mutation play a role in the pathogenesis of these syndromes. Recent work, particularly by Hebbel and Verrellotti [2], has established that SCD is an ‘inflammatory state’ and that the endothelium is activated. Thus, the emerging view of the pathogenesis of many complications of SCD involves complex interactions between sickle reticulocytes, neutrophils, monocytes, and the endothelium. It therefore follows that many factors that are important in the pathways leading to inflammation, cell adhesion, and coagulation will be important in the pathophysiology of the disease and that variations in the expression of a variety of molecules in these pathways will contribute to the heterogeneity of SCD.

The approach to the study of this complexity in the postgenome era should involve several different methodologies including: (1) analysis of polymorphisms in ‘candidate genes’ by high throughput methods and the association of these polymorphisms with distinct phenotypic features of the disease, (2) the study of differential gene expression in various tissues using cDNA microarrays, and (3) the application of proteomics. The tissues and cells that are particularly important in this regard include blood cells (neutrophils, monocytes, reticulocytes), bone marrow, endothelial cells, and liver tissues. A candidate gene approach should include a large number of patients with and without a certain complication (e.g. stroke) to determine the role of these polymorphisms as risk factors. Pooling studies to discover the role of genome-wide SNPs as risk factors for certain complications of the disease is also emerging as a viable approach. In all of these studies, careful characterization of the phenotype remains of critical importance. A study of sickle cell patients from different populations is likely to yield important information because of the differences of the genetic backgrounds in these populations and their potential implications on the disease phenotype. The important SCD-related ‘phenotypes’ should include complications such as stroke and CNS disease, frequent vs. infrequent vaso-occlusive episodes (VOEs), frequency of the acute chest syndrome, pulmonary hypertension, AVN of the hips and shoulders, gallbladder disease and cholecystectomy, leg ulcers, renal involvement, priapism, and retinopathy.

For the past several years, our center and many other centers have focused on the study of nonglobin genetic modifiers of sickle cell disease. Some of the findings of these early studies will be summarized below.

**AVN and Methylenetetrahydrofolate Reductase C677T Mutation**

Studies of large patient populations such as the Cooperative Study of Sickle Cell Disease have shown that co-inherited α-thalassemia, age, high hematocrit, and frequent VOEs are risk factors for the development of AVN of the femoral head in SCD [3]. We determined the frequency of the thermolabile methylenetetrahydrofolate reductase (MTHFR) C677T mutation in our sickle cell population, because of its association with elevated serum homocysteine levels and resultant vascular complications. Overall, the MTHFR mutation was present in 16% of the sickle cell patients in our center (1.8% homozygote). There was a strong association of the presence of MTHFR with AVN, with 35.6% of the AVN patients having the MTHFR mutation as opposed to only 12.9% of sickle cell patients without AVN (p = 0.006). Furthermore, the presence of concomitant α-thalassemia and MTHFR was found to be additive in terms of the risk of developing AVN [4]. This association, however, was not confirmed in the high Hb F population of Kuwaiti sickle cell patients, suggesting that different genetic factors may be operative in different populations [5].

**Factor V R485K Polymorphism**

First described as a neutral polymorphism in the factor V gene in nucleotide 1628 in exon 10 by Gandrille et al. [6], the Arg→Val substitution at residue 485 was found at high frequency (32.4%) in sub-Saharan Africans. This polymorphism was later associated with a risk of thrombosis and coronary artery disease in Asian populations and reported to lead to a mild activated protein C resistance [7]. A study of the frequency of this polymorphism in our sickle cell population showed a similar allele frequency of 0.29 (45.1% heterozygous, 6.6% homozygous). Although no association of this polymorphism has been established with certain complications of SCD (frequent VOEs, acute chest syndrome, leg ulcers, AVN, and priapism), interestingly, the presence of this mutation has
been shown to be associated with an increased risk of central venous catheter-associated thrombosis ($p = 0.006$, odds ratio = 4.9; table 1) [8].

**UGT1A1 Polymorphism and Bilirubin Levels**

The gene UGT1A1 encodes the enzyme UDP-glucuronosyltransferase-1 that mediates the glucuronidation of bilirubin. A common polymorphism in the TATA box of this gene was found to be associated with a decreased enzyme activity and indirect hyperbilirubinemia (Gilbert’s syndrome). The sequence A(TA)6TAA is presumed to be the wild type while A(TA)7TAA is associated with Gilbert’s syndrome. The coinheritance of this (TA)7 polymorphism with hereditary hemolytic anemias, such as hereditary spherocytosis, α-thalassemia, and G6PD deficiency, results in marked hyperbilirubinemia and increased incidence of gallstones [9]. We, and others, have studied the frequency of this polymorphism in the sickle cell population and its impact on bilirubin levels and incidence of gallstones and cholecystectomy [10, 11]. In all of these studies, an association was found between the (TA)7 genotype and bilirubin levels; those with 7/7 had the highest mean bilirubin levels compared to the genotypes with 6/6 and 6/7. Our results are summarized in table 2. Furthermore, the impact of the 7/7 genotype on treatment with hydroxyurea has recently been reported by Henney et al. [12]. These authors reported a normalization of the bilirubin levels in sickle cell patients with the 6/6 genotype upon treatment with hydroxyurea in contrast to those with 7/7 and 6/7 genotypes whose bilirubins failed to fall to normal levels.

**Genetic Polymorphisms as Risk Factors for Stroke**

Clinical stroke occurs in ~11% of US sickle cell patients by 20 years of age [13]. Increased velocities of blood flow (>200 cm/s) as measured by transcranial Doppler (TCD) ultrasound in major intracranial arteries has been shown to be a predictor of stroke risk in children with SCD [14]. The value of TCD as a surrogate marker of cerebrovascular disease and a predictor of stroke risk has been validated in the randomized STOP trial, and a therapeutic intervention based upon this risk stratification (prophylactic transfusions in children with high TCD velocities) has been proven to reduce the stroke risk [14]. Despite the validation of the utility of TCD as a clinical tool for predicting stroke risk, biologic factors underlying the development of cerebrovascular disease and stroke in the SCD population are poorly understood. The presence of α-thalassemia has been shown to be protective of cerebrovascular disease and stroke [13]. Several recent studies have sought an association between certain genetic polymorphisms and stroke risk in SCD. Driscoll et al. [15] found a higher rate of stroke in siblings with SCD ($p = 0.0012$), suggesting the role of genetic factors. A recent study has found an association with HLA genotypes and stroke; of particular interest is the association of certain genotypes with distinct subtypes (large vessel vs. small vessel stroke) in a retrospective study of the Cooperative Study of Sickle Cell Disease population [16]. Another recent study [17] reported a protective role for VCAM G1238C polymorphism against symptomatic stroke (odds ratio = 0.35, $p = 0.04$). An ongoing study at our center is looking at the association of high TCD velocity with 28 genetic polymorphisms among 25 candidate genes in 630 patients (230 high TCD, 400 normal TCD) enrolled in the STOP and STOP II trials. The candidate genes include those associated with coagulation and thrombophilia (factor V, factor VII, factor XIII, prothrombin, thrombomodulin, fibrinogen, PAI-1, MTHFR), endothelial cell function and inflammation
Gene expression profiling utilizing cDNA microarrays is a relatively novel method, particularly in terms of its applications to the study of SCD. Few published reports exist in this area. These include the study of gene expression profiling from the kidney tissue in transgenic sickle cell mouse [18], profiling of bronchial endothelial cells in patients with SCD and acute chest syndrome [19], and a study of differential gene expression in peripheral blood cells from SCD patients with differing vaso-occlusive crisis frequency [20].

The quality of the data obtained from microarray studies depends upon several factors: (1) a meticulous study design, particularly a clear and unambiguous definition of the 'phenotype(s)' to be studied, (2) elimination of technical/methodological pitfalls (e.g. quantity/quality of the RNA, methods of specimen collection and processing, conditions and timing of hybridization, and elimination of intrapatient variation by repeated sampling), and (3) expertise in bioinformatics in the interpretation of vast amounts of data generated.

In patients with SCD, the tissues that may be amenable to gene expression profiling studies include peripheral blood cells, endothelial cells (circulating endothelial cells, as well as endothelial cells from various tissues), bone marrow, and liver tissue. We have conducted preliminary studies of gene expression with commercially available cDNA microarrays (Affymetrix U95, Affymetrix, Santa Clara, Calif., USA) from peripheral blood neutrophils and reticulocytes in SCD patients with discordant rates of painful episodes in an effort to elucidate the factors contributing to different crisis frequency. Two patients who were quite similar in terms of demographic (age, gender, ethnic/racial background) and hematologic (hemoglobin, reticulocyte count, neutrophil count, Hb F level, α-globin genes and β5 haplotypes) characteristics, but very different in terms of annual crisis frequency (0.75/year vs. 6/year) were chosen for the preliminary study. An analysis of gene expression patterns from purified peripheral blood neutrophils revealed that of the 7,070 genes represented on the microarray, 536 had a >2-fold expression difference between the 2 patients. Sixteen genes had a higher expression and 27 genes had a lower expression in the patient with a high crisis frequency as compared to the patient with a lower vaso-occlusive crisis frequency. The genes with a higher expression in the patient with frequent VOEs included CD11b/CD18, phospholipase A2, C and D, and IL-8, suggesting a higher level of neutrophil activation in this patient.

The application of the microarray technology to the study of SCD is in its infancy. Although the preliminary data look interesting, these results need to be confirmed and validated through a large number of experiments in a large number of patients. The available data appear promising and show the feasibility of the application of this technology to SCD.

Conclusions

The development of high throughput genotyping methods, the microarray technology, and the completion of the Human Genome Project have already begun to revolutionize medicine and biology. The application of these techniques to the study of SCD is in its very early stages and is likely to yield novel information.

A candidate gene approach and study of genome-wide SNPs and their association with certain complications of SCD is expected to further clarify the basis of clinical/phenotypic heterogeneity of this single-gene disorder. A study of gene expression profiling coupled with proteomics will shed further light on the functional aspects and differences between patients. This information will not only lead to a better understanding of the pathogenesis and pathophysiology of many complications of the disease, but will also likely result in the identification of novel therapeutic targets and discovery of new genes with prognostic and therapeutic implications.
References


2. Hebbel RP, Vercellotti GM: The endothelial biology of sickle cell disease. J Lab Clin Med 1997;129:288–293.


