LESSONS FROM THE MOLECULAR BIOLOGY OF G6PD DEFICIENCY

Ernest Beutler

Glucose-6-phosphate dehydrogenase (G6PD) deficiency was the first prevalent enzymatic deficiency to have been described. It is of special interest to the hematologist because this inherited enzymatic defect is more severe in erythrocytes than in most other tissues; as a consequence, the most common clinical effect is anemia. Also of importance is the severe neonatal icterus that may occur in deficient newborns, probably largely because of the effect of the enzyme deficiency on the immature liver. Our understanding of the population genetics of G6PD deficiency has been revolutionized by the application of the techniques of molecular biology. Moreover, mutation detection has become clinically feasible and clinically useful in the evaluation of patients with hemolytic anemia and in prenatal diagnosis. The text of this tutorial is based, in part, upon previous reviews of G6PD by the author,[1-3] and these may be consulted for more detailed information.

Biochemical Characterization of G6PD Variants

By 1988 the existence of nearly 400 variants of G6PD thought to be biochemically distinct had been documented.[4] These variants were distinguished from each other by comparing a variety of characteristics, including residual enzyme activity, electrophoretic mobility, affinity for substrates and substrate analogues, stability to heating, and pH optimum. Although standardization of methodology had been achieved by international agreement,[5,6] the comparison of variants characterized in different laboratories on the basis of qualitative differences in difficult-to-measure parameters was fraught with difficulty. Was a variant with 98% of normal electrophoretic mobility measured in Italy truly different from one that had a mobility measured at 96% of normal three years later in a laboratory in Germany? The fact that variants that were the same might be considered different and vice versa was widely appreciated. Indeed, it was shown that a variant that had been identified as G6PD Chicago in one laboratory and later as G6PD Cornell in another came from different members of the same family.[7]

The difficulty in determining whether variants are the same or different can be overcome with the qualitative data that are provided when G6PD DNA is actually sequenced, and the collection of such sequence data had to await the cloning of the G6PD gene.[8] It has now become clear that some variants that were thought to be distinct are, indeed, the same and that others that were thought to be identical are, in fact, different. The variants that have been characterized at the DNA level have recently been summarized.[1,9,10]

Polymorphic G6PD Variants

Polymorphic variants are those that have achieved a high frequency in some populations. They represent balanced polymorphisms in which the benefit of inheriting

the mutation (probably resistance to malaria) counterbalances the disadvantage (susceptibility to hemolysis and neonatal icterus). Generally, each population has its own characteristics mutations, although, as noted below, there are occasional exceptions to this rule.

African Variants

G6PD deficiency among Africans is relatively mild; red cells contain 10-15% residual enzyme that is electrophoretically rapid. Accordingly it is designated G6PD A- to distinguish its mobility from the normal enzyme, which is designated B. Among Afro-Americans the gene frequency of G6PD A- is about 11%.[11] It is considerably higher in some parts of Africa.[12] Some Africans have an enzyme with the same rapid mobility encountered among deficient individuals, but the activity of the enzyme is normal. This African enzyme is designated G6PD A+, and it is known to be due to an A→G transition at nt 376 that predicts substitution of a negatively charged aspartic acid for asparagine. Its gene frequency is in the range of 20-30%

We now realize that G6PD A- is also quite common in Mexico and in Southern European countries. Presumably because at one time the mind-set was that this was exclusively an African variant, the enzyme in these individuals was regarded as unique and given different designations such as G6PD Distrito Federal, Matera, Castilla, Betica, Tepic, and Ferrara. However, molecular analysis has shown these variants simply to be examples of G6PD A-. It is not surprising that G6PD A- would be found in Southern Europe and in countries to which Southern Europeans migrated when one considers the proximity of these countries to the African continent and their history.

However, G6PD A- has proven to be heterogeneous when studied at the DNA level. Common to all of the types of G6PD A- is the A→G transition at nt 376, the mutation that characterizes G6PD A+. In addition, the various types of G6PD A- have a second mutation; this mutation, although most commonly a G→A at nt 202, may also be a 680 G→T or a 968 T→C. There is, indeed, a Puerto Rican variant, very possibly of African origin, G6PD Santamaria, that shares the nt 376 mutation with a 542 A→T mutation, giving rise to an enzyme with normal electrophoretic mobility. It is apparent from these observations that the G6PD A+ mutation at nt 376 is relatively ancient, and that the other mutations have been superimposed upon it. Other polymorphic sites exist in G6PD, and it has been deduced that these have probably arisen in this order: 1) 376 A→ G; 2) BspHI lost; 3) PvuII gained; 4) 202 G→A.(13-15)

Based on assumptions of cross-over frequencies, we have estimated that the G6PD A- mutation probably arose less than 80,000 years ago.[13]

Mediterranean Variants

G6PD deficiency is very prevalent in some Mediterranean countries. A gene frequency of about 0.7 has been documented among Kurdish Jews.[16] This is the highest incidence known in any group. Among Greeks, Turks, Sardinians, Sephardic Jews and Italians, G6PD deficiency is also quite prevalent, but more commonly with gene frequencies that range from 0.02 to 0.20. The situation regarding Mediterranean variants is in one respect the reverse of the situation of the African variants. Here several different variants (e.g., G6PD “Sassari” and “Cagliari”) were believed to be different on the basis

of biochemical characteristics, but all seem to share the same mutation at nt 563. Variants from other parts the world, thought to be unique - G6PD Dallas, Birmingham, and Panama - proved to be G6PD Mediterranean563T. Other mutations are found in the Mediterranean region as well. G6PD Seattle844T (also described as G6PD Lodi and as G6PD "Modena") and, as noted above, G6PD A- are relatively common.

Oriental Variants

Considerable heterogeneity of G6PD has been documented in various Asian populations. Gaozhou95G has been documented in China, Ube241T and Konan241T in Japan, "Chinese-4"392T in China, "Chinese-3"493G in the Philippines, Mahidol487A in Southeast Asia, China, and Taiwan, G6PDViangchan871A in populations from India, Laos, and the Philippines, G6PD Kalyan949A in India, G6PD Chatham1003A in the Philippines, "Chinese-5"1024T among the Chinese, Canton1376T among in China, and Kaiping1388G in China and Laos.

The Origin of G6PD Mutations

Different polymorphic variants predominate in different regions of the world. True, they spread to adjacent areas. Thus G6PD A- has spread to Southern Europe and G6PD Viangchan871A has been found in a variety of Oriental populations, including India, Laos, and the Philippines. Generally, however, variants are found in contiguous populations. Not so with G6PD Union, a class 2 variant, first documented in Hawaii among Filipinos.(17) Subsequently, the variant has been found in the Vanuatu archipelago in the Southwestern Pacific,(18) in Laos, China, Japan, and also in Spain(19) and in Italy.(20) Another variant that has a very widespread distribution is G6PD Mediterranean, which is found both in Europe and in Asia.

When variants are found in regions quite distant from one another, two explanations may be entertained. First, it is possible the gene was brought from one area to another, presumably by intrepid explorers; the other is that the mutation has arisen more than once. The context with respect to other polymorphic mutations, i.e., the haplotype, with which it is found may provide useful information in this respect. Thus, G6PD A- 202A/376G is found in the same haplotype wherever it is encountered, whether in sub-Saharan Africa, Southern Europe or Mexico. This is consistent with and, indeed, suggestive of a single origin for this mutation. G6PD Mediterranean563T is usually found in the context of the 1311T haplotype, Mediterranean countries, while in the Indian subcontinent it is usually found in the context of the 1311C haplotype. This might argue for a separate origin, but this interpretation is complicated by the fact that even in Europe this mutation is sometimes found in the context of the nt 1311C haplotype. Were there two separate mutational events that gave rise to G6PD Mediterranean? Or did a crossover occur between the nt 563 and the nearby nt 1311 site thousands of years ago? Similarly, G6PD Jammu/Viangchan, both due to a G→A mutation at nucleotide 871, are found in different haplotypes.(21) Again, the same question arises, and there is no way to know whether a crossover or two separate mutations are responsible for these findings.

Sporadic Variants

From the point of view of the clinical hematologist, sporadic variants may be of greater importance that the polymorphic variants. The latter have achieved high gene frequencies because they have relatively few morbid consequences. Sporadic variants, in contrast, often cause hereditary nonspherocytic hemolytic anemia (HNSHA). One might have anticipated that the number of these variants would be virtually limitless, but this proves not to be the case. There are many different mutations, but the same ones are encountered repeatedly in the population. I believe that this is due to the constraints that allow some mutations to cause anemia yet be compatible with life. No G6PD mutations associated with a total lack of G6PD have been found. Thus the null state is probably lethal, and this places one restriction on the number of abnormalities that we can expect to find in patients. It may well be that another constraint is that the mutation causes a more severe deficiency in red cells than in other tissues. Finally, to be brought to our attention, a mutation needs to have an effect on the patient; almost always the effect is anemia. It is interesting that of the 37 different single-point mutations that cause HNSHA to have been documented, 21 are in the narrow range of cDNA nt 1089 to 1361, a portion of the molecule that seems markedly to influence NADP binding. An additional six mutations causing HNSHA are within 100 base pairs of nt 613-615, encoding the lysine that binds glucose-6-phosphate in the G6PD reaction (Figure 1).

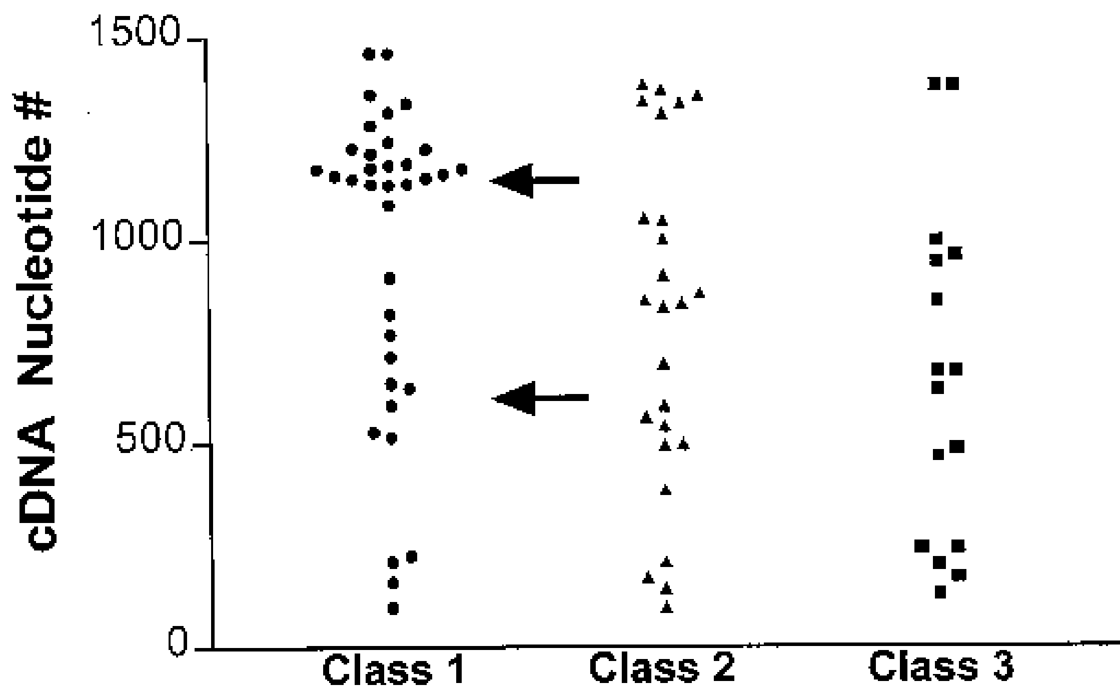

Figure 1. The distribution of single point mutations along the G6PD cDNA in Class 1 (hereditary non-spherocytic hemolytic anemia), Class 2 (severe deficiency) and Class 3 (mild deficiency) mutants. Although mutations are found along the entire length of the coding sequence, Class 1 mutations, in particular, are clustered. Notably, all mutations between nt 1057 and nt 1316 are associated with chronic hemolytic anemia. The arrows denote the glucose-6-P binding site (nt 613-615) and a portion of the molecule that markedly affects NADP binding.

Diagnostic Value

It is ordinarily not difficult to diagnose G6PD deficiency by measuring the activity of the enzyme in red blood cells. However, under some circumstances difficulties may arise, and DNA-based diagnosis is then very helpful. When a patient has been transfused, red cell activity will be that of the transfused cells, but the leukocyte DNA, which is used for diagnosis, will be unaffected. Similarly, when a patient with one of the milder variants of G6PD undergoes hemolysis, the remaining or newly produced young red cells may mask the deficiency, but DNA-based diagnosis will still be entirely reliable. Diagnosis in females, too, may pose a problem when enzyme assays are used to detect heterozygotes because of the X-inactivation phenomenon.

Prenatal diagnosis is difficult using biochemical methods and is best achieved by DNA examination. Another situation in which we have found DNA-based diagnosis particularly valuable is the one that arises when a patient from an ethnic group in which G6PD deficiency is common has a chronic hemolytic anemia. In such a circumstance the physician is confronted with a dilemma. Since a high percentage of patients will be G6PD deficient in any case, is there any cause-and-effect relationship between this patient's anemia and G6PD deficiency? Performing DNA analysis for the mutation common in that population will provide the answer. If the African patient has G6PD A-, then the hemolysis must have some other cause. Conversely, finding a different mutation, particularly one located in exon 10, the region where most of the mutations causing NSHA are found, makes a cause-and-effect relationship highly likely.

When the mutation is known, DNA-based diagnosis is very practical under these circumstances; amplifying the appropriate portion of the DNA by PCR and then examining the fragment produced with a restriction enzyme is a relatively minor undertaking in many laboratories nowadays.

Other Clinical and Research Uses of G6PD Mutations

The existence of a common polymorphism in the coding region at nt 1311 has allowed us to develop a new method for determining whether a cell population is clonal, utilizing the X-inactivation phenomenon in female heterozygotes, and determining whether one or both mRNAs are made. The existence of many mutations affecting the G6PD molecule provides an opportunity to deduce structure/function relationships. We showed that a small, highly conserved region, especially amino acids 385, 386 and 387, was associated with defective NADP binding and suggested that this region might be a NADP-binding site. Crystallographic studies on Leuconostoc enzyme by Mason et al,[22] however, suggested that the NADP-binding site might be closer to the amino end of the molecule. Further structural studies will, no doubt, expand our knowledge of how G6PD functions.

Acknowledgements: This is manuscript 10024-MEM from The Scripps Research Institute. Supported by National Institutes of Health grants HL25552 and RR00833 and the Stein Endowment Fund.

References

1. Beutler E: G6PD deficiency. Blood 84:3613, 1994
2. Beutler E: Déficit en G6PD: Pathologie molèculaire. Hèmatologie 1:385, 1995
3. Beutler E: G6PD: Population genetics and clinical manifestations. Blood Rev 10:45, 1996
4. Beutler E, Yoshida A: Genetic variation of glucose-6-phosphate dehydrogenase: A catalog and future prospects. Medicine (Baltimore) 67:311, 1988
5. Betke K, Beutler E, Brewer GJ, Kirkman HN, Luzzatto L, Motulsky AG, Ramot B, Siniscalco M: Standardization of procedures for the study of glucose-6-phosphate dehydrogenase. Report of a WHO scientific group. WHO Tech Rep Ser No. 366:1967
6. Glucose-6-phosphate dehydrogenase deficiency. Report of a WHO working group. WHO Geneva, 3-4 September 1985
7. Fairbanks VF, Nepo AG, Beutler E, Dickson ER, Honig G: Glucose-6-phosphate dehydrogenase variants: Reexamination of G6PD Chicago and Cornell and a new variant (G6PD Pea Ridge) resembling G6PD Chicago. Blood 55:216, 1980
8. Toniolo D, Persico MG, Battistuzzi G, Luzzatto L: Partial purification and characterization of the messenger RNA for human glucose-6-phosphate dehydrogenase. Mol Biol Med 2:89, 1984
9. Vulliamy T, Beutler E, Luzzatto L: Variants of glucose-6-phosphate dehydrogenase are due to missense mutations spread throughout the coding region of the gene. Hum Mutat 2:159, 1993
10. Beutler E, Westwood B, Melemed A, Dal Borgo P, Margolis D: Three new exon 10 glucose-6-phosphate dehydrogenase mutations. Blood Cells Mol Dis 21:64, 1995
11. Heller P, Best WR, Nelson RB, Becktel J: Clinical implications of sickle-cell trait and glucose-6-phosphate dehydrogenase deficiency in hospitalized black male patients. N Engl J Med 300:1001, 1979
12. Luzzatto L: Studies of polymorphic traits for the characterization of populations. African populations South of the Sahara. Isr J Med Sci 9:1181, 1973
13. Kay AC, Kuhl W, Prchal JT, Beutler E: The origin of G6PD polymorphisms in Afro-Americans. Am J Hum Genet 50:394, 1992
14. Beutler E, Kuhl W: Linkage between a PvuII restriction fragment length polymorphism and G6PD A-202A/376G: Evidence for a single origin of the common G6PD A- mutation. Hum Genet 85:9, 1990
15. Vulliamy TJ, Othman A, Town M, Nathwani A, Falusi AG, Mason PJ, Luzzatto L: Polymorphic sites in the African population detected by sequence analysis of the glucose-6-phosphate dehydrogenase gene outline the evolution of the variants A and A-. Proc Natl Acad Sci USA 88:8568, 1991
16. Oppenheim A, Jury CL, Rund D, Vulliamy TJ, Luzzatto L: G6PD Mediterranean accounts for the high prevalence of G6PD deficiency in Kurdish Jews. Hum Genet 91:293, 1993

17. Beutler E, Westwood B, Kuhl W, Hsia YE: Glucose-6-phosphate dehydrogenase variants in Hawaii. Hum Hered 42:327, 1992
18. Ganczakowski M, Town M, Bowden DK, Vulliamy TJ, Kaneko A, Clegg JB, Weatherall DJ, Luzzatto L: Multiple glucose-6-phosphate dehydrogenase-deficient variants correlate with malaria endemicity in the Vanuatu archipelago (southwestern Pacific). Am J Hum Genet 56:294, 1995
19. Vives Corrons JL, Rovira A, Pujades MA, Estrada M, Vulliamy TJ: Diverse point mutations of glucose-6-phosphate dehydrogenase (G6PD) gene in Spanish and Cuban patients with hemolytic anaemia. La Revista de Investigacion Clinica 46(Suppl):234a, 1994
20. Calabrò V, Mason PJ, Filosa S, Civitelli D, Cittadella R, Tagarelli A, Martini G, Brancati C, Luzzatto L: Genetic heterogeneity of glucose-6-phosphate dehydrogenase deficiency revealed by single-strand conformation and sequence analysis. Am J Hum Genet 52:527, 1993
21. Beutler E, Prchal JT, Westwood B, Kuhl W: Definition of the mutations of G6PD Wayne, G6PD Viangchan, G6PD Jammu and G6PD “LeJeune”. Acta Haematol (Basel) 86:179, 1991
22. Naylor CE, Rowland P, Basak AK, Gover S, Mason PJ, Bautista JM, Vulliamy TJ, Luzzatto L, Adams MJ. Glucose 6-phosphate dehydrogenase mutations causing enzyme deficiency in a model of the tertiary structure of the human enzyme. Blood 87:2974, 1996