Fetal Hemoglobin Levels in Sickle Cell Disease and Normal Individuals Are Partially Controlled by an X-Linked Gene Located at Xp22.2

By G.J. Dover, K.D. Smith, Y.C. Chang, S. Purvis, A. Mays, D.A. Meyers, C. Sheils, and G. Serjeant

Fetal hemoglobin (Hb F) production in sickle cell (SS) disease and in normal individuals varies over a 20-fold range and is under genetic control. Previous studies suggested that variant Hb F levels might be controlled by genetic loci separate from the β-globin complex on chromosome 11. Using microscopic radial immunodiffusion and flow cytometric immunofluorescent assays to determine the percentage of F reticulocytes and F cells in SS and nonanemic individuals, we observed that F-cell levels were significantly higher in nonanemic females than males (mean ± SD, 3.8% ± 3.2% vs 2.7% ± 2.3%). F-cell production as determined by F reticulocyte levels in SS females was also higher than in SS males (17% ± 10% vs 13% ± 8%). We tested the hypothesis that F-cell production in both normal and anemic SS individuals was controlled by an X-linked locus with two alleles, high (H) and low (L). Using an algorithm to determine the 99.8% confidence interval of a normal distribution in nonanemic individuals, we estimated that males and females with at least one H allele had greater than 3.3% F cells. Comparisons of male-male or female-female SS sib pairs with discordant F reticulocyte levels distinguished two phenotypes in SS males (L, <12%; H, >12%) and three phenotypes in SS females (LL, <12%; HL, 12% to 24%; HH >24%). Linkage analysis using polymorphic restriction sites along the X chromosome in eight SS and one AA family localized the F-cell production (FCP) locus to Xp22.2, with a maximum lod score (logarithm of odds of linkage vs independent assortment) of 4.8 at a recombination fraction of 0.04.

© 1992 by The American Society of Hematology.

POSTNATALLY, fetal hemoglobin (Hb F) production persists and is confined to a subset of erythrocytes, termed F cells.1 Chernoff2 and Marti and Butler3 demonstrated that Hb F levels in normal adults vary over a 20- to 35-fold range. Hb F levels in individuals with sickle cell anemia (SS), albeit higher than in normal adults, also vary over a 20-fold range.4 Marti and Butler,3 using isotype levels of Hb F, and Zago et al,5 using immunologic techniques for counting F cells, showed that high Hb F (1.1% to 3.4%) and high F-cell levels (>8%) in normal adults were inherited. The terms applied to these examples of hereditary persistence of Hb F (HPFH) were “Swiss type HPFH” or heterocellular HPFH.5,5 Dover et al6 and later Mason et al7 and Milner et al8 showed that either F-cell or Hb F levels in nonanemic sickle trait (AS) parents predicted Hb F production in their SS children. The genetic mechanisms underlying this variation in Hb F production in normal and anemic individuals are unclear.

Because Hb F is unevenly distributed within the red blood cells (RBCs) of normal and SS individuals, differences in peripheral blood Hb F levels may be attributed to at least three distinct variables: differences in the production of F cells, variation in the amount of Hb F per F cell, and, in the case of SS disease, the preferential survival in the peripheral circulation of F cells compared with RBCs containing no detectable Hb F.9 In SS individuals, the variation in F-cell production as measured by the percent F reticulocytes is the major variable contributing to differences in Hb F levels.

The broad distribution of Hb F levels in normal adults and in anemic SS individuals suggests that more than one genetic factor may control Hb F production. Several point mutations in the promoter regions of the two gamma genes have been associated with increased Hb F production in otherwise normal adults, thereby indicating that factors linked to the β-globin gene complex on chromosome 11 influence Hb F production. Differences in mean Hb F levels in SS individuals homozygous for each of the three most common β-globin gene restriction fragment length polymorphism (RFLP) haplotypes also suggest that genetic factors linked to chromosome 11 influence Hb F levels.10,11

Several laboratories have suggested that a locus controlling F-cell production (FCP locus) in normal individuals and in SS disease may be separate from the β-globin gene complex located on chromosome 11. Five families have been described in which heterocellular HPFH was not linked to the β-globin gene complex.12-16 Comparison of F-cell production between SS siblings indicated that one third of the sib pairs had different levels of F-cell production, thus indicating that at least one FCP locus in SS disease is not linked to the β-globin gene complex.17 Recently, Miyoshi et al reported evidence suggesting that F-cell production in heterocellular HPFH within the Japanese population is controlled by an X-linked locus.18

Thus, Hb F production is controlled by multiple genetic factors both linked and unlinked to the β-globin gene complex on chromosome 11. In this report, we have attempted to localize one of the unlinked factors: a gene that controls F-cell production in SS and normal individuals. Our initial analysis tests the simplest genetic model consistent with our analysis of F-cell levels. This model is based on the assumption that a single locus responsible for a major portion of the variation in a quantitative trait can be genetically mapped by linkage to polymorphic DNA loci.19

From the Department of Pediatrics, Medicine, The Johns Hopkins University School of Medicine, and the Kennedy Institute, Baltimore, MD; and the MRC Sickle Cell Center, University of the West Indies, Kingston, Jamaica.

Submitted September 27, 1990; accepted April 14, 1992.

Supported by Grant No. HL28028 from the National Heart Lung and Blood Institute, National Institutes of Health (NIH), Bethesda, MD; and the MRC Laboratories, Kingston, Jamaica. Y.C. Chang is supported by NIH Training Grant No. ST32GM10714-09.

Address reprint requests to George J. Dover, MD, Ross Research Building, Room 1125, Johns Hopkins University Medical School, 720 Rutland Ave, Baltimore, MD 21205.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. section 1734 solely to indicate this fact.

© 1992 by The American Society of Hematology.

0006-4971/92/8003-0027$3.00/0

Blood. Vol 80, No 3 (August 1), 1992: pp 816-824
In this report, we provide evidence for a FCP locus that is linked to the short arm of the X-chromosome within the Xp22.2 region.

**MATERIALS AND METHODS**

**Subjects.** Venous blood was collected in either heparinized or EDTA-containing tubes after verbal or written informed consent as approved by the Johns Hopkins University Medical School Joint Committee on Clinical Investigation. Normal adult samples were obtained from blood donors at Johns Hopkins and from normal controls used in a previously published survey.20 All donors had normal hemoglobin levels and normal RBC indices. Pregnant females or recently pregnant females (<6 months since delivery) were excluded. SS subjects and their families were identified from the clinic population at Johns Hopkins or from the Medical Research Council Sickle Cell clinic at the University of West Indies, Kingston, Jamaica. Two groups of SS patients from Jamaica were studied. A sib pair group was derived from 81 families with more than one child with SS disease. Part of this group was the subject of a previous publication,21 which outlined criteria for the diagnosis of SS disease, reproductibility of F-cell determinations, and validation of paternity. A second group, termed the cohort group, included 103 subjects chosen randomly from 275 SS subjects identified in a prospective neonatal screening program in Jamaica.21 The age range for the sib pair group was 5 to 50 years (mean ± SD, 19.6 ± 12.7 years) and 5 to 15 years for the cohort group (9.8 ± 2.3 years). All subjects in Jamaica have been monitored for more than 5 years and their steady-state hematologic tests have been repeatedly assayed. A white family with heterocellular hereditary persistence of Hb F and X-linked adrenoleukodystropy was also studied.22 Iron-deficiency anemia was excluded in the nonanemic subjects on the basis of a normal hemoglobin for age and normal RBC indices. In SS subjects, elevated serum ferritin excluded iron deficiency.

**F-cell assays.** Assays for percent F cells were performed on peripheral blood to determine F-cell production in normal nonanemic adults. The percent F cells in nonanemic adults were studied. A sib pair group was derived from 81 families with more than one child with SS disease. Part of this group was the subject of a previous publication,21 which outlined criteria for the diagnosis of SS disease, reproductibility of F-cell determinations, and validation of paternity. A second group, termed the cohort group, included 103 subjects chosen randomly from 275 SS subjects identified in a prospective neonatal screening program in Jamaica.21 The age range for the sib pair group was 5 to 50 years (mean ± SD, 19.6 ± 12.7 years) and 5 to 15 years for the cohort group (9.8 ± 2.3 years). All subjects in Jamaica have been monitored for more than 5 years and their steady-state hematologic tests have been repeatedly assayed. A white family with heterocellular hereditary persistence of Hb F and X-linked adrenoleukodystrophy was also studied.22 Iron-deficiency anemia was excluded in the nonanemic subjects on the basis of a normal hemoglobin for age and normal RBC indices. In SS subjects, elevated serum ferritin excluded iron deficiency.

**DNA analysis.** The haplotypes resulting from the pattern of RFLPs within the β-globin gene complex were determined for all SS subjects to help exclude nonpaternity in SS sib pairs.23 RFLP analysis in all subjects used in linkage studies was performed on DNA derived from whole blood or Epstein-Barr virus (EBV)-transformed lymphocytes using X-chromosome probes obtained from the ATCC (Rockville, MD), Collaborative Research (Bedford, MA), or individual investigators. Probes used for linkage analysis included DXS 52 (Xq28), HPRT (Xq26), DXS 11 (Xq25-24), DXS 17 (Xq22), PGK (Xq13), DXS 14 (Xp11.21), DXS 7 (Xp11.4-11.3), DXS 84 (Xp21.1), DXS 67 (Xp21.3), DXS 41 (Xp22.1), DXS 274 (Xp22.2-22.1), DXS 16 (Xp22.2), DXS 43 (Xp22.2), DXS 85 (Xp22.2-22.3), DXS 143 (Xp22.3), and DXS 452 (Xp22.3).

**Linkage analysis.** Linkage analysis was performed using the computer program LIPED (courtesy of J. Ott, Columbia University School of Medicine, New York, 1987). This program uses the method of maximum likelihood to calculate lod scores (logarithm of odds of linkage vs independent assortment) at selected recombination fractions for each pedigree, and permits summing lod scores from individual pedigrees. Because no single probe was informative in all families, multiple probes from the same region were tested. For determination of lod score, only one informative probe per region was used with each pedigree. Linkage analysis was performed by treating the variation in F-cell production as either a discrete or a continuous trait. The criteria used for defining FCP phenotypes are described in the Results. For the continuous trait analysis, the mean ± SD of FCP phenotypes were defined as follows: high (HH) 31.8% ± 6.9% F reticulocytes, intermediate (HL) 16.5% ± 3.5% F reticulocytes, and low (LL) 6.2% ± 3.0% F reticulocytes. Complete penetrance was assumed for each allele. Allele frequencies were derived from analysis of the cohort group. Where indicated, we analyzed our maximum lod score as described by Ott.24,26,27 for situations where the mode of inheritance is unknown.

**RESULTS**

**F-cell production in normal and AS adults.** Using our flow cytometry technique for measuring F-cell levels, we examined 292 adults (Fig 1). We found that females have a significantly higher percentage of F cells than males (3.8% ± 3.0%, female; 2.7% ± 2.3%, males; P = .003). Mason et al.7 studying children less than 6 years old, and Rutland et al.28 studying adults, have previously described higher Hb F levels in nonanemic females than males. The

---

Fig 1. F-cell production in nonanemic adults. The percent F cells was determined by flow cytometric assays of 10,000 cells per assay using monoclonal mouse anti-human Hb F antibody.23 The skewed distribution is identical to that seen by Miyoshi et al.18 Results for 292 healthy adults (A, 121 males; B, 171 females) are shown. The broken vertical line (3.3% F cells) is the 99.8% confidence interval for a best-fit normal distribution (μ = 3.3; mean ± SD for the calculated normal distribution, 1.7 ± 0.7).
skewed distribution and the range of F cells in our study (0.2% to 20%) are similar to that observed in the Japanese adults studied by Miyoshi et al.\(^8\) (<1% to 16%) and Zago et al.\(^9\) (<1% to 18%). Black and white adults had similar average percent F cells (3.3% ± 2.9% and 3.0% ± 2.7%, respectively). The range for 25 AS individuals (0.2% to 15%) is identical to that seen in AS subjects studied by Wood et al.\(^8\)

**F-cell production in SS subjects.** Because of preferential survival of F cells in SS subjects, the most accurate estimate of F-cell production in individuals with SS is the F-reticulocyte count.\(^9\) In both the sib pair and the cohort groups of SS patients, the percent F reticulocytes in females was higher than males (sib group: females, 18% ± 11% \([n = 98]\) \(v\) males, 14% ± 9% \([n = 80]\), \(P = .02\); cohort group: females, 13% ± 7% \([n = 45]\) \(v\) males, 10% ± 7% \([n = 58]\), \(P = .006\)). In the case of SS sib pairs, both individuals in a given sib pair must have identical loci at the \(\beta\)-globin gene complex. Thus, the difference in F-cell production must be due to genetic factors elsewhere. This difference was evident at all age levels, suggesting that iron deficiency, which is more prevalent in menstruating females, is not a cause for the difference in F-cell production. The percent F reticulocytes was not correlated with either the hemoglobin level (g/dL) or the percent total reticulocytes. A striking difference in the distribution of percent F reticulocyte is seen when SS males are compared with SS females (Fig 2). While SS males have a skewed unimodal distribution, SS females have a distribution that appears to be bimodal. The limited number of females analyzed did not allow statistical confirmation of a bimodal distribution; however, the probability of bimodality was 80% by the bimodality test of Giacomelli et al.\(^{16}\)

The discrepancy between male and female F-reticulocyte levels suggested the possibility that gender differences might have falsely elevated the proportion of SS sib pairs with discordant F-reticulocyte levels in the previous SS sib study.\(^7\) We had previously defined F-reticulocyte levels between SS sibs as being discordant if the probability that F-reticulocyte levels for SS sibs were alike was 10\(^{-4}\) as determined by Student's \(t\)-test. Therefore, we reexamined our now expanded sib pair study group to determine if differences in F-cell production among SS sibs were related to gender. Among 97 sib pairs, 39 (40.2%) had significantly different percent F reticulocyte levels. Among sibs of the same sex, 15 of 45 (33%) had discordant percent F reticulocytes. Thus, gender differences between sib pairs cannot explain the differences in F-cell production between all SS sibs. Since these SS sibs must have identical \(\beta\)-globin loci, this reanalysis confirms our previous finding\(^17\) that one control of F-cell production in SS disease is separate from the \(\beta\)-globin locus.

**X-linked FCP locus.** It is clear that there are marked differences in the number, range, and distribution of F-cell production in males and females. While sex-limited expression could explain these differences, the simplest hypothesis is that the responsible genetic locus is on the X-chromosome. Therefore, we tested the hypothesis that a single FCP locus with two alleles, high (H) and low (L), exists on the X-chromosome.

**Phenotypes and genotypes of nonanemic adults.** Because nonanemic adults have a skewed unimodal distribution of percent F cells, it has been difficult to distinguish high or low FCP phenotypes. However, Miyoshi et al were able to distinguish high and low F-cell groups using a least-squares method.\(^{18}\) An algorithm that sequentially eliminates the highest values one by one was used to determine the best fit for a normal distribution of the total data. A 99% confidence interval for the normal distribution was used to distinguish the high and low phenotypes. In Miyoshi's analysis, using microscopic examination of blood smears treated with a rabbit anti-human Hb F antibody, high F-cell production was set at greater than 4.4%. Applying the same algorithm to our flow cytometry–determined percent F cells, we identified a high group beginning at 3.3% (Fig 1). This analysis allows assignment of high (H) and low (L) phenotypes for males and females. Based on the hypothesis of a X-linked gene, we can deduce that males have either an H allele (>3.3% F cells) or an L allele (<3.3% F cells). In females, we can infer the LL genotype (<3.3% F cells), but
cannot distinguish females who were heterozygous for H and L alleles from females who were homozygous for the H allele. We can only deduce from this analysis that nonane- and L alleles from females who were homozygous for the H allele cannot distinguish females who were heterozygous for H allele. We reasoned that if differences between discordant SS males is attributable to H and L alleles at a locus on the X chromosome, a single boundary should separate the percent F reticulocytes among discordant male sibs. Similarly, one or two demarcation boundaries should exist for females (Bimodal LL v HL or HH; Trimodal LL v HL v HH). Figure 3 illustrates that among the 15 same sex sib pairs discordant for F-reticulocyte levels, discordant SS male sibs are separated by a single boundary between 10% and 12% F reticulocytes, and females are separated by two demarcation boundaries at 10% to 12% and 24% to 26% F reticulocytes. These observations are consistent with two genotypes among SS males (L < 12% and H > 12% F reticulocytes) and three genotypes in females (L < 12%, HL 12% to 24%, and HH > 24% F reticulocytes). The apparent three genotypes in females suggests that the alleles are codominant and that the putative X-linked FCP gene may not be subject to X-inactivation. This is in accord with the elevated level of F reticulocytes in females compared with males.

**Hardy-Weinberg analysis.** Comparison of F-reticulocyte levels in SS individuals shows that a smaller proportion of females have a low percentage of F reticulocytes than do males (see Fig 2). We calculated the fraction of individuals in both groups with the low phenotype (<12% F reticulocytes). Among SS females, 49% (22/45) had less than 12% F reticulocytes as compared with 76% (44/58) among males. A significant difference between males and females in the distribution of F reticulocytes is not compatible with classic autosomal inheritance. However, if a major determinant of F-cell production is X-linked, the proportion of males with an L phenotype (q) should be greater than the proportion of females with the LL phenotype (q^2). Similar differences in the distribution of G6PD activity among males and females was one of the initial indications that it was X-linked.39

Using the phenotype designations described above and assuming X-linkage, the frequency of the L allele in our male SS population is 0.76. Using this value for the L allele, we calculated the expected frequencies of the three female phenotypes using the Hardy-Weinberg equation. Observed and expected numbers were not significantly different (Observed: LL = 22, HL = 19, HH = 4; Expected: LL = 26, HL = 16, HH = 3), χ^2 = 1.51, P = .47. Given our data set (N = 103), the 95% confidence limits of this chi-square analysis allow the frequency of the L allele in males to be between 0.57 and 0.79, if X-linkage is assumed.

If autosomal inheritance is assumed, the frequency of the L allele in our SS male population is the square root of 0.76 or 0.87. Using this value to predict the expected frequencies of phenotypes in SS females gave values that were significantly different from those observed (Observed: LL = 22, HL = 19, HH = 4; Expected: LL = 34, HL = 10, HH = 1). Chi-square analysis shows that our data are incompatible with autosomal inheritance: χ^2 square = 21.34, P = ~<.0001.

**Pedigree analysis.** If F-cell production in SS disease is regulated by an X-linked locus, the FCP phenotype in SS or AS males should not be inherited from their fathers. Alternatively, if F-cell production is controlled by an autosomal locus, one would expect that one half of the males should inherit the FCP phenotype from their fathers. We examined the parents of 20 SS males with F-reticulocyte levels greater than 12% from 18 pedigrees, and one AS male with greater than 3.3% F cells (Fig 4). In nine pedigrees where one parent was L or LL (<3.3% F cells) and the other parent had at least one H allele (>3.3% F cells), only one of 12 possible father-to-son transmissions of the H phenotype was observed (top row, Fig 4). Two pedigrees (bottom row, left, Fig 4) were not informative, since both parents had at least one H allele. Where only the mother was available for analysis, five of six pedigrees indicated that the mother had at least one H allele (bottom row, right, Fig 4). Although not unequivocal, these data strongly support X-linked inheritance of F-cell production. We have no explanation for the apparent father-son transmissions (see arrow), but they may indicate that F-cell production is also influenced by non-X-linked factors.

**Linkage analysis.** Formal linkage analysis using RFLPs associated with known positions on the X-chromosome was performed because the data described above were most...
Fig 4. Pedigrees of males with the H-FCP phenotype (>12% F reticulocytes for SS and >3.3% F cells for AS). The numbers associated with AS individuals are percent F cells, while the numbers associated with SS individuals are percent F reticulocytes. Note on the top row, where both mother and father were assayed, 11 out of 12 times the mother was either HL or HH (>3.3% F cells). On the second row, in two pedigrees, both parents had an H phenotype, or, in six pedigrees where only the mother was available for analysis, the mother was either HL or HH five out of six times. Transmission of the H allele from mother to son was observed 11 out of 12 times (top row) and possibly 19 out of 21 times (total pedigrees). The arrows indicate two pedigrees where mother-to-son transmission of the H allele is tentatively excluded. Half-filled symbols, AS; solid symbols, SS.

consistent with an X-linked gene having a significant effect on F-cell production. We analyzed our data first as discreet phenotypes, and then as a continuous trait. Note that a lod score of greater than 2 is considered sufficient for establishing linkage to the X-chromosome.31

Linkage analysis treating FCP as discreet phenotypes. Using the discreet phenotypes discussed above (see phenotypes of nonanemic adult and phenotypes of SS subjects), recombinants can be easily identified. This treatment eliminates minor variations within each phenotype that are attributable to sampling error or other factors unlinked to the X-chromosome. Erroneous assignment of phenotype can easily lead to false recombinants and thus nonlinkage (low lod score), even when two loci are linked. Because the probability of obtaining false evidence for linkage is small,26 a significant lod score would provide strong evidence for an X-linked gene. Our strategy was to use a family with nine sibs (six SS females, two AS females, and one AS male) to screen for a candidate linkage site. This family has been studied repeatedly over several years and full paternity has been established.17 As seen in Fig 5, at least two recombinants were detected between the FCP phenotype and each of the six markers on the long arm of the Xp-chromosome, leading to lod scores of less than −2 (data not shown). This excludes Xq as being a candidate region for the FCP locus. There was no recombination with the Xp loci DXS 43 and DXS 41, suggesting that the putative FCP locus is located on the short arm of the X-chromosome.

We then applied linkage analysis of discreet phenotypes to the rest of our collection of pedigrees, using the linkage program LIPED. In six additional SS pedigrees and one AA pedigree, the mothers were heterozygous for one or more probes (DXS 43 and DXS 41) on the short arm of the X-chromosome. The numbers below AS individuals refer to percent F cells and the numbers below SS individuals refer to percent F reticulocytes. The assignments listed on the top row (H, HH, HL) refer to the FCP phenotypes given to male and female AS and SS subjects as illustrated in Figs 1 to 3 and discussed in the text. The designation for the X-linked probes (left column) includes the abbreviation for the restriction enzyme used to test for linkage. The location for each probe on the X chromosome is listed in the right column. With each probe, the mother's two X-chromosomes can be distinguished by the presence (+) or absence (−) of the polymorphic restriction enzyme site. For simplicity, only the segregating maternal chromosome is shown for each offspring. Note that no recombinants are seen with either Xp probe. None of the long arm sites appear to be linked. Half-filled symbols, AS; solid symbols, SS.

Fig 5. Summary of linkage analysis in one SS pedigree (Ba) for six RFLP probes on the long arm of the X-chromosome and two anonymous RFLP probes (DXS 43 and DXS 41) on the short arm of the X-chromosome. The numbers below AS individuals refer to percent F cells and the numbers below SS individuals refer to percent F reticulocytes. The assignments listed on the top row (H, HH, HL) refer to the FCP phenotypes given to male and female AS and SS subjects as illustrated in Figs 1 to 3 and discussed in the text. The designation for the X-linked probes (left column) includes the abbreviation for the restriction enzyme used to test for linkage. The location for each probe on the X chromosome is listed in the right column. With each probe, the mother's two X-chromosomes can be distinguished by the presence (+) or absence (−) of the polymorphic restriction enzyme site. For simplicity, only the segregating maternal chromosome is shown for each offspring. Note that no recombinants are seen with either Xp probe. None of the long arm sites appear to be linked. Half-filled symbols, AS; solid symbols, SS.
ADULT Hb F LEVELS ARE INFLUENCED BY AN Xp GENE

RFLP on the short arm of the X-chromosome and had at least one H allele (Fig 6). Of 27 informative meiosis (families BA, M, P, SW, H), only one recombinant (M, II-3) was identified between FCP phenotypes and markers previously mapped to Xp22.2. Using allele frequencies derived from analysis of the male cohort data (see Hardy-Weinberg analysis, L = 0.76; H = 0.24), we observed a maximum lod score of 4.6 at a recombination fraction of 0.036 (Table 1). Using the allele frequencies determined to be acceptable by the Hardy-Weinberg analysis (L = 0.57 to 0.79, the 95% confidence range), we obtained maximum lod scores between 4.6 and 4.5, at a recombination fraction of 0.036. The regions proximal and distal to Xp22.2 were either unlinked or weakly linked to the FCP locus (Table 1). This analysis provides evidence for close linkage between the FCP locus and Xp22.2. Note the maximum lod score for SS families (BA, M, P, SW) alone is 3.4 at recombination fraction of 0.04, and the lod score for a single informative family (BA) was 2.408 at a recombination fraction of 0.

The number of available pedigrees was insufficient to allow formal segregation analysis to statistically test our pedigrees for modes of inheritance. Without independent tests of modes of inheritance, linkage analysis alone can lead to false assignment of an otherwise unconfirmed putative gene.\(^3\)\(^2\)\(^3\) Methods for correcting lod scores in analyses where the mode of inheritance is unknown have been suggested.\(^2\)\(^7\) Essentially, the maximum lod score is not significant unless it is greater than the accepted lod score plus the log of the number of polymorphic sites tested while searching for linkage. Ott\(^2\) also points out that such

![Pedigree Diagrams](image)

**Fig 6.** Pedigrees of six SS families and one AA family.\(^2\)\(^3\) The segregation of Xp probes (DXS 84, 41, 43, 16, 85, 143, and 452) in families where these probes were informative are shown. Note that the probes at Xp22.2 (DXS 43, 16, and 85) show only one recombination (pedigree M; II-3) out of 27 informative meioses (pedigrees M, II-3; pedigree II-1,2; pedigree II-2,3,4,5, 1, 2, 3, 4, 5, 6, 7, 8); Empty symbols, AA; half-filled symbols, AS; solid symbols, SS.
validity of the test. Applying this conservative standard, our markers, Ott’s correction is not required to ensure the correction is only necessary if the number of markers tested is greater than 100. Because we only used 16 polymorphic markers, Ott’s correction is not required to ensure the validity of the test. Applying this conservative standard, our analysis requires a maximum lod score of 3.2 (2 + log 16) to provide evidence for X-linkage, and thus the lod of 4.6 remains significant.

**Linkage analysis of FCP as a continuous trait.** Linkage analysis was also performed after entering the actual value of F-cell or F-reticulocyte percentage into LIPED, rather than using discrete phenotypic categories. This treatment avoids arbitrary assignment of each individual to a particular phenotype and requires only the means ± SD expected from each phenotype. Using the means ± SD estimated from the female population of our sib pair (see Materials and Methods) study, the maximum lod score was 3.1 at a recombination fraction of 0.05 for linkage between the putative FCP locus and Xp22.2 DNA markers. Since this lod score is greater than the usually accepted value of 2, it also provides significant support for X-linkage. The lod score from regions other than Xp22.2 were again negative or insignificant. When dealing with continuous traits, the maximum lod score is expected to be lower than the one calculated from discrete phenotype, since variations within each peak of the distribution lower the significance of each nonrecombinant identified. Our analysis has the added disadvantage that the means ± SD used were calculated from relatively small sample sizes within the total test population (N = 98). Nevertheless, the result of this analysis is compatible with our previous assumptions and provides evidence for linkage of the FCP locus to Xp22.2.

**DISCUSSION**

Because Hb F interferes with the polymerization of Hb S, SS individuals with elevated levels of Hb F appear to have less-severe disease.\(^{35,36}\) Indeed, a recently completed natural history study of SS subjects in the United States shows an inverse correlation between percent Hb F and vaso-occlusive crises, which is continuous down to levels of approximately 4% Hb F.\(^{37}\) Therefore, it has been of considerable interest to determine what genetic mechanisms control the diverse levels of Hb F seen in SS disease. Deletions of large portions of the b-globin gene complex (deletion HPFH) and point mutations in the promoter region of the g-globin genes (nondeletion HPFH) have been shown to increase Hb F levels. However, these mutations are rare and cannot account for the wide variation seen in the general SS population.\(^{38}\) Several investigators have suggested that Hb F levels may be linked to particular b-globin gene haplotypes as determined by RFLP analysis.\(^{10,11}\) The Senegal haplotype has been associated with elevated Hb F levels in African, Saudi, and Indian populations of SS patients. However, when Miller et al analyzed Hb F production in individual pedigrees of SS subjects from the Eastern Saudi Arabia, they noted that elevated Hb F production could not be explained solely by association with the Senegal-like haplotype.\(^{39}\) In our analysis of haplotypes in the Jamaican SS sib and cohort groups, we find that less than 10% of the variation in Hb F levels or percent F reticulocytes is associated with a particular b-globin haplotype. This undoubtedly is due to the relative paucity of homozygotes for the less frequent haplotypes (Central African Republic and Senegal) in Jamaican and American SS patients. Our earlier data suggest that 70% of the variation in Hb F levels in SS disease is attributable to the variation in percent F reticulocytes, and that, as described above, the percent F reticulocytes is at least partially controlled by the FCP locus on the X-chromosome. Further studies comparing the influence of H and L alleles of the FCP locus in patients with different haplotypes will have to be done to see whether factors associated with different chromosome 11 haplotypes affect the expression of the FCP locus.

The mechanism by which the FCP locus modulates F-cell production is unknown. Trans-acting proteins bind to enhancer and promoter regions in the b-globin complex and may modulate differential globin gene expression. Indeed, Zon et al\(^{40}\) have isolated a DNA-binding protein, GATA-1, which may serve as one of the major regulators of erythroid gene expression. The gene for this protein has been localized to the X-chromosome, at Xp21.1. Another factor that influences erythroid differentiation in vitro, erythroid

<table>
<thead>
<tr>
<th>Location</th>
<th>Locus</th>
<th>Lod Score (2) at Recombination Fraction (%)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>.005</td>
<td>.10</td>
</tr>
<tr>
<td>Discrete phenotypes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xp22.3</td>
<td>DXS 143</td>
<td>-4.12</td>
<td>.43</td>
</tr>
<tr>
<td>Xp22.2</td>
<td>DXS 452</td>
<td>4.07</td>
<td>4.22</td>
</tr>
<tr>
<td>Xp22.1</td>
<td>DXS 85</td>
<td>-3.76</td>
<td>.74</td>
</tr>
<tr>
<td>Xp21.1</td>
<td>DXS 43</td>
<td>-10.49</td>
<td>-2.02</td>
</tr>
<tr>
<td>Continuous distribution</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xp22.3</td>
<td>DXS 143</td>
<td>-2.23</td>
<td>.44</td>
</tr>
<tr>
<td>Xp22.2</td>
<td>DXS 452</td>
<td>2.54</td>
<td>3.02</td>
</tr>
<tr>
<td>Xp22.1</td>
<td>DXS 43</td>
<td>-3.21</td>
<td>.29</td>
</tr>
<tr>
<td>Xp21.1</td>
<td>DXS 84</td>
<td>-11.35</td>
<td>-3.04</td>
</tr>
</tbody>
</table>
ADULT Hb F LEVELS ARE INFLUENCED BY AN Xp GENE

Potentiating activity, has been localized to Xp11.4-11.1.\textsuperscript{41} Since we have clearly defined families that show nonlinkage (negative lod scores) of the FCP locus to the Xp21.1cen region, our data suggest that the FCP locus represents a gene on the X chromosome which is separate from the two previously discussed genes. Furthermore, recent experiments by Zitnik et al,\textsuperscript{42} using hetero-specific hybrids made from human lymphoid cells fused with MEL cells, are compatible with the possibility that the gamma gene on human chromosome 11 may be reactivated by genetic material on chromosomes other than chromosome 11. Characterization of the FCP locus itself will be necessary to determine whether it encodes for a protein that also binds to the globin region or, alternatively, affects F-cell production without directly interacting with the β-globin gene complex on chromosome 11.

Our data suggest that the X-linked FCP alleles are codominant and not subject to X-inactivation. The mean F-cell and F- reticulocyte levels are higher in females than in males (Figs 1 and 2), while the F-reticulocyte level in HL females is approximately half of that seen in HH females (Fig 2). In addition, sib pair analysis (Fig 3) indicates the presence of three phenotypes in females and just two in males. Although most X-linked genes are subject to dosage compensation in females, an increasing number of X-linked genes are being described that escape X-inactivation.\textsuperscript{43,44} These genes exist on both the long and short arm of the X-chromosome, indicating that escape from X-inactivation is not a local phenomenon. Currently, four areas of Xp are known to escape X-inactivation. Two of these, the distal pseudoautosomal region and ZFX, bracket Xp22.2, the presumed site of the FCP locus. Brown and Willard\textsuperscript{44} have speculated that much of the short arm of the X-chromosome was a recent addition to the eutherian X-chromosome, and suggested that genes throughout its length might escape X-inactivation.

We presume that the X-linked factor controlling F-cell production in nonanemic Japanese adults described by Miyoshi et al\textsuperscript{18} is the same as the FCP factor described in this report. Although Miyoshi et al suggested that their factor was X-linked, their pedigrees could not establish formal linkage to any region on the X-chromosome. In discussing their observations, Miyoshi et al pointed out that most of the pedigrees describing heterocellular HPFH in the literature were compatible with X-linked transmission.\textsuperscript{18} The fact that the FCP locus is mapped to the same region on the short arm of the X-chromosome in both black SS families and in a white family, suggests that in humans there is at least one FCP locus on the X-chromosome. However, it should be emphasized that our analysis does not exclude other loci either on the X-chromosome, chromosome 11, or other autosomes which may contribute to the heterocellular HPFH phenotype.

ACKNOWLEDGMENT

We thank R. Schnur, L. Kunkel, P. Pearson, K. Davis, G. Sack, and B. Vogelstein for DNA probes. We thank Drs H. Franklin Bunn and Samuel H. Boyer for their thoughtful advice and encouragement.

REFERENCES

2. Chernoff AI: Immunologic studies to hemoglobins. II. Quantitative Precipitin test using anti fetal hemoglobin sera. Blood 8:413-421, 1953
8. Milner PF, Leibfarth JD, Ford J, Barton BP, Grenett HE, Garver FA: Increased Hb F in sickle cell anemia is determined by a factor linked to the B\textsuperscript{0} gene from one parent. Blood 63:64, 1984
persistence of fetal hemoglobin regulated dominantly by gene(s) on X chromosome. Blood 72:1854, 1988


23. Dover GJ, Boyer SH: Fetal hemoglobin-containing cells have the same mean corpuscular hemoglobin as cells without fetal hemoglobin: A reciprocal relationship between γ- and β-globin gene expression in normal subjects and in those with high fetal hemoglobin production. Blood 69:1109, 1987


34. Thompson EA: Interpretation of LOD scores with a set of marker loci. Genet Epidemiol 1:357, 1984


Fetal hemoglobin levels in sickle cell disease and normal individuals are partially controlled by an X-linked gene located at Xp22.2

GJ Dover, KD Smith, YC Chang, S Purvis, A Mays, DA Meyers, C Sheils and G Serjeant