*Proc. Natl. Acad. Sci. USA*
Vol. 73, No. 10 pp. 3633–3636, October 1976
Genetics

# Trimodality in the proportion of hemoglobin G Philadelphia in heterozygotes: Evidence for heterogeneity in the number of human alpha chain loci

(α-thalassemia/globin synthesis/α-chain variant/hemoglobin H disease)

ROSALIE M. BAINE*, DONALD L. RUCKNAGEL*, PABLO A. DUBLIN, JR.†, AND JUNIUS G. ADAMS III‡

*Department of Human Genetics, University of Michigan Medical School, Ann Arbor, Mich. 48109; †Center for Sickle Cell Disease, Howard University, Washington, D.C., 20059; and ‡Veterans Administration West Side Hospital, P.O. Box 8195, Chicago, Illinois 60612

*Communicated by James V. Neel, July 8, 1976*

**ABSTRACT** **The extent of variability in the number of human hemoglobin (Hb) alpha chain loci has not yet been conclusively determined. There is evidence that in some populations individuals may possess two α-chain loci, while in other populations only one locus is present. Electrophoresis of peripheral blood from 53 heterozygotes for Hb G Philadelphia ($\alpha^{68\ Asn \rightarrow Lys}$) revealed that the proportion of Hb G is trimodally distributed, with modes at approximately 20, 30, and 40% Hb G. Familial, hematologic, and statistical studies suggest that the proportion of Hb G is not random but is genetically controlled and inversely correlated with mean cell volume. Two alternative genetic models are proposed to explain these findings: one assumes α-thalassemia, while the other postulates variability in the number of α-chain loci in the American Black population. Biosynthetic studies of blood from 15 subjects revealed balanced synthesis of α and β globin chains in heterozygotes from all three classes, strongly supporting variable gene dosage rather than α-thalassemia as the mechanism underlying the observed trimodality in the proportion of Hb G.**

**Incompatibilities between our results and current concepts of α-thalassemia are discussed in the context of differences between Black compared with Oriental and Italian forms of Hb H disease.**

The genetics of the human hemoglobin (Hb) α-chain has generated considerable discussion. In 1968, Lehmann and Carrell observed that most stable α-chain variants comprise only 15–25% of the total hemoglobin in heterozygotes, although most stable β-chain variants comprise 30–50%. On these grounds, they proposed that humans have two α-chain loci but only one β-chain locus (1). Hollan *et al.* provided direct support for this hypothesis with the report of three different α chains, $\alpha^{J\text{-}Buda}$, $\alpha^{G\text{-}Pest}$, and $\alpha^{A}$, in each of three Hungarians (2). Similarly, De Jong *et al.* reported two Indians possessing $\alpha^{Rampa}$, $\alpha^{Koya\ Dora}$, and $\alpha^{A}$ (3). Clearly, these persons must have more than one α-chain locus. In 1975, the DNA·cDNA hybridization studies of Kan *et al.* suggested that only 25% of the normal number of α-chain structural genes are present in persons with Hb H disease, implying that four active α-chain genes are usually present (4).

However, not all α-chain variants support this model. In 1970, Abramson *et al.* described a Melanesian family in which two $\alpha^{J\text{-}Tongariki}$ homozygotes were found. Heterozygotes for Hb J in this family possessed 50% of the abnormal component. This suggests the presence of only one α-chain locus in this family (5). Beavens *et al.* confirmed these observations (6). Recently, Hb J Mexico was found to comprise 30 or 40% of the total hemoglobin in heterozygotes for this α-chain variant (7). Furthermore, a number of other α-chain variants, including hemoglobins G Philadelphia, Anantharaj, Nyanza, G Chinese, J Rovigo, and J Cape Town, have been reported in high proportions (40–50%), as recently reviewed (8).

In the absence of thalassemia, most hemoglobin variants occur in a proportion characteristic of the variant. However, in 1974, Rucknagel and Winter reported a possible bimodality in the proportion of Hb G Philadelphia found in heterozygotes (9). We have now quantitated the hemoglobins in 53 American Black Hb G heterozygotes and found a range of 19–43% Hb G ($\alpha_2^G\beta^2 + \alpha_2^G\delta_2$). This paper is concerned with the distribution of the proportion of Hb G, its significance in genetic and hematologic terms, its relationship to globin synthesis in Hb G heterozygotes, and the implications of these findings for the genetics of the α-chain.

Abbreviation: Hb, hemoglobin.

## MATERIALS AND METHODS

Venous blood was drawn in tubes containing EDTA, and red cell indices were measured using the Coulter counter, model S. In all cases, the presence of Hb G Philadelphia was confirmed by peptide mapping (10). Different hemoglobins were quantitated by electrophoresis on cellulose acetate strips, elution of the hemoglobin bands, and determination of the absorbance of each eluate at 415 nm. All quantitations were performed in triplicate, and the mean values were used to plot the distribution of the proportion of Hb G. Globin was synthesized *in vitro* by incubating peripheral blood with [$^3$H]leucine for 90 min (11). In some cases, reticulocyte enrichment was performed by the method of DeSimone *et al.* (12). After lysis and removal of stroma by centrifugation, globin was prepared by acid acetone precipitation. The component chains were then separated as described by Clegg *et al.* (13), except that 0.045 M $Na_2HPO_4$ was used as the limit buffer. Radioactivity in each fraction was assayed by liquid scintillation counting of an aliquot from each tube. The total counts incorporated into each chain were determined by integrating the counts of all fractions in the corresponding peak after subtracting background counts.

## RESULTS AND DISCUSSION

### Distribution of proportion of Hb G

The histogram in Fig. 1 presents the observed distribution of the proportion of Hb G. At first glance, the observations appear to have been drawn from a trimodal distribution since the histogram has three well-separated peaks. Nonetheless, it is possible that random fluctuation coupled with laboratory error has produced this distribution by chance. To interpret the observed distribution, we performed family studies, hematologic studies, and distribution analyses.

**Family Studies.** If the proportion of Hb G is not genetically determined, within-family means and variances in the pro-

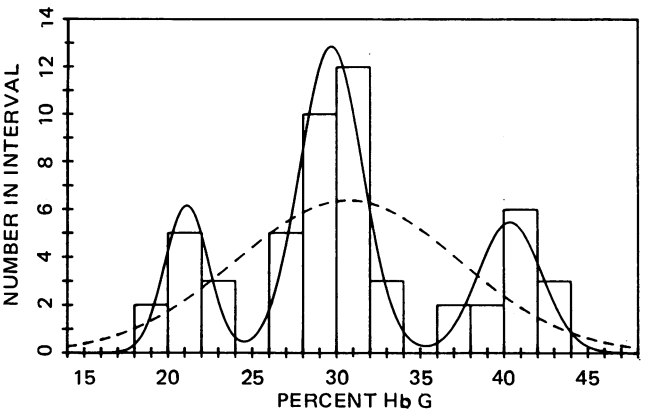


FIG. 1. Observed and hypothetical distributions of Hb G Philadelphia. The hypothetical unimodal (broken line) and trimodal (solid line) distributions are superimposed on the observed distribution of the proportion of Hb G (histogram). The figure was generated by the Calcomp plotter.

portion of Hb G should be approximately equal. The observed data are listed in Table 1. Fisher's $F$-test revealed significant differences in within-family variances [$F(11,200) = 3.1$, $P < 0.001$]. An analysis of variance yielded a highly significant $F$-value [$F(9,25) = 8.8$, $P < 0.001$]. This strongly suggests inequality of within-family mean proportion of Hb G. Since equality of within-family variances is an assumption for this analysis of variance, the validity of this test is questionable and the corresponding, nonparametric Kruskal-Wallace test was performed. Although this test is much less powerful, relying on differences in average ranks rather than actual mean proportion of Hb G, a value of 27.1, significant at the 1% level, was obtained. This confirmed the presence of significant differences in mean proportion of Hb G between families. In two families (*1* and *2*), all Hb G heterozygotes belong to the 20% class. Heterozygotes in six families (*3–8*) all possessed 30% Hb G, while in family *12* all Hb G heterozygotes possessed 40% Hb G. Within each of these families, the variance in the proportion of Hb G was quite low. Thus the proportion of Hb G appears to be genetically determined.

In contrast, three families (*9–11*) contained heterozygotes from both the 30 and 40% classes, yielding larger within-family

Table 1. Family data

| Family | $N$ | Mean % Hb G | Variance | Average rank |
|---|---|---|---|---|
| 1 | 6 | 20.9 | 1.70 | 4.2 |
| 2 | 2 | 21.5 | 0.20 | 5.5 |
| 3 | 2 | 27.1 | 2.35 | 11.8 |
| 4 | 2 | 28.9 | 0.01 | 16.5 |
| 5 | 2 | 29.5 | 6.16 | 17.0 |
| 6 | 2 | 29.2 | 1.94 | 17.3 |
| 7 | 2 | 29.4 | 0.20 | 18.5 |
| 8 | 4 | 32.1 | 1.90 | 24.5 |
| 9 | 2 | 32.5 | 38.46 | 20.0 |
| 10 | 5 | 34.7 | 38.57 | 24.4 |
| 11 | 4 | 38.8 | 50.65 | 29.8 |
| 12 | 4 | 39.9 | 4.64 | 31.5 |

Significant differences were found in within-family variances, as measured by Fisher's $F$-test, and in within-family means, as measured by analysis of variance. The nonparametric Kruskal-Wallace test revealed significant differences in average within-family ranks. These results strongly support genetic determination of the proportion of Hb G.

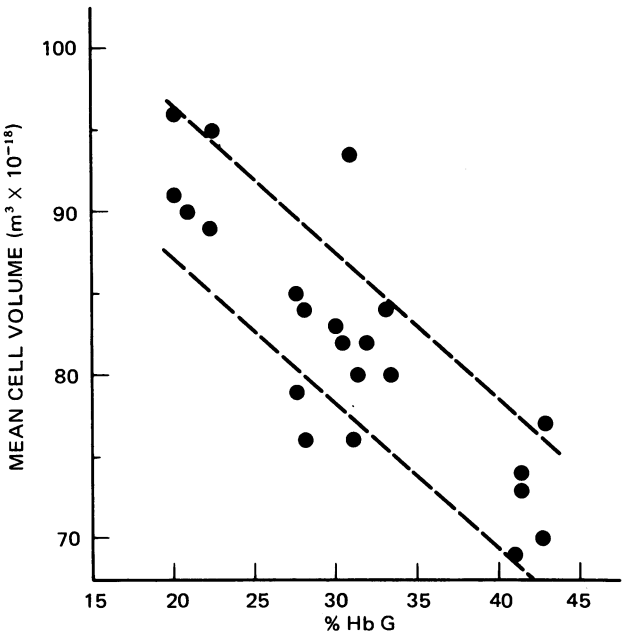


FIG. 2. Mean cell volume plotted against proportion of Hb G. The broken lines are removed 1 SD from the regression line.

variances. Some genetic determinant affecting the proportion of Hb G seems to be segregating within these families.

**Hematologic Studies.** If the variation in the proportion of Hb G is random, no correlation between proportion of Hb G and any hematologic parameter should be expected. However, mean cell volume is negatively correlated with the proportion of Hb G (Fig. 2); the correlation coefficient is −0.82 (degrees of freedom = 20, $P < 0.001$). Similarly, the proportion of Hb G and mean cell hemoglobin content are also negatively correlated; the correlation coefficient is −0.81 (degrees of freedom = 20, $P < 0.001$).

**Distribution Analysis.** Nonsystematic variation should give the proportion of Hb G a unimodal normal distribution whose mean and variance we estimate by the sample mean and variance. Since we have discovered families in which all Hb G heterozygotes possessed approximately 20, 30, or 40% Hb G, a trimodal distribution seemed the most reasonable alternative to a unimodal distribution. Under the trimodal hypothesis, we represent the distribution function as the weighted sum of the distribution functions of the three classes. Each class is normally distributed with a mean and variance set equal to the class mean and variance; each class is weighted in proportion to the number of persons in that class. Fig. 1 shows the hypothetical unimodal and trimodal distributions superimposed on the observed distribution.

We were unable to reject the unimodal alternative on standard statistical grounds alone; we found no significant skew or kurtosis in the observations, and both the Kolmogorov-Smirnov and the Cramer-von Mises tests yielded values that were not statistically significant. However, when likelihood ratio methods were used to determine the relative probability that the observations were drawn from the trimodal rather than the unimodal distribution, a highly significant result was obtained. The value of −2 ln (likelihood ratio), which is asymptotically distributed as chi-square with four degrees of freedom, is approximately 36, significant at less than the 0.1% level. Thus, it is at least a thousand times more likely that the observations were drawn from the trimodal rather than the unimodal distribution. We conclude that the variation in the proportion of

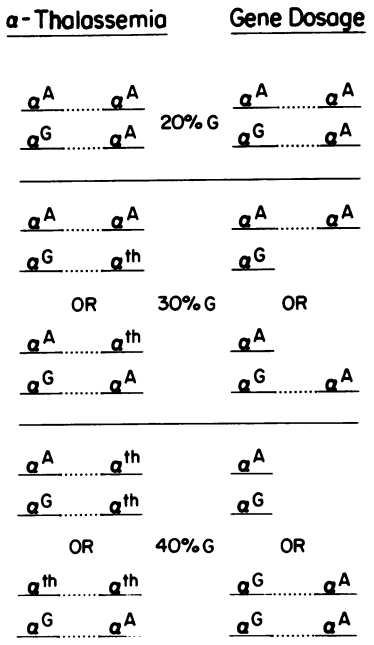


FIG. 3. Possible configurations for two alternative genetic models that can explain the trimodal distribution of the proportion of Hb G. The two loci may or may not be linked.

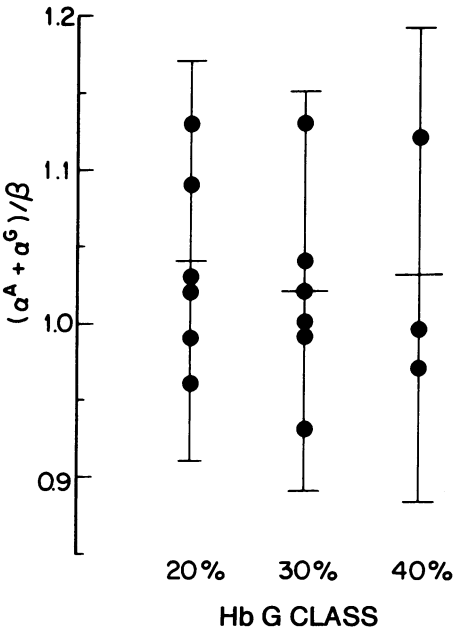


FIG. 4. Ratios of incorporation of [$^3$H]leucine into $\alpha$ and $\beta$ chains of 15 Hb G heterozygotes in the three phenotypic classes. For each class, data points are superimposed on the mean ±2 SD.

Hb G is meaningful, genetically determined, and trimodally distributed.

## Alternative genetic models

Fig. 3 presents two genetic models that can explain the observed trimodality in the proportion of Hb G. These are referred to as the $\alpha$-thalassemia model and the variable gene dosage model.

The $\alpha$-thalassemia model assumes that the phenotype reflects the number of $\alpha$-thalassemia genes. One $\alpha$-thalassemia allele at one of the two $\alpha$-chain loci results in a barely detectable phenotype referred to as the silent carrier (or $\alpha$-thal-2 trait). Two such genes result in more overt thalassemia (or $\alpha$-thal-1 trait), three in Hb H disease, and four in a lethal condition characterized by hydrops fetalis. The $\alpha$-thalassemia model also rests upon the well-known observation that the presence of $\alpha$-thalassemia increases the proportion of an $\alpha$-chain structural variant. Thus, an individual with one $\alpha^G$ gene, two $\alpha^A$ genes, and one $\alpha$-thalassemia gene would have 30% Hb G. One $\alpha^G$, one $\alpha^A$, and two $\alpha$-thalassemia genes would result in 40% Hb G.

The alternative variable gene dosage model postulates variability in the number of $\alpha$-chain loci in the American Black population, so that an individual may have two, three, or four $\alpha$-chain genes (9, 14, 15). (For simplicity of presentation, the loci are assumed to be linked.) This model would explain the observed trimodality by differences in the number of $\alpha^A$ genes in each class. Thus, the genotype of the 40— hb G class would be composed of one $\alpha^G$ gene and one $\alpha^A$ gene; that of the 30% class would comprise one $\alpha^G$ and two $\alpha^A$ genes. Theoretically, a person with 40% Hb G could also have inherited one $\alpha^A$ plus one $\alpha^G$ gene from each parent. The presence of $\alpha^G$ on both one- and two-locus chromosomes must be the result of either two independent mutations or crossingover.

We have attempted to distinguish between the $\alpha$-thalassemia and variable gene dosage models by performing biosynthetic studies *in vitro* using peripheral blood from persons drawn from all three classes. If $\alpha$-thalassemia is the cause of the observed variation in the proportion of Hb G, then in all persons with 40% Hb G we would expect that the sum of the counts incorporated into the $\alpha^A + \alpha^G$ chains would equal only about 70–80% of the counts in the $\beta$-chain (16, 17). A lesser imbalance would be expected in the 30% class. Balanced synthesis would be expected in the 20% class, which serves as a control.

The observed $\alpha/\beta$ synthetic ratios are presented in Fig. 4. The means and standard deviations were 1.04 ± 0.07, 1.02 ± 0.07, and 1.03 ± 0.08 for the 20, 30, and 40% classes, respectively. These are virtually identical to each other and to the normal control values reported by others, e.g., Kan *et al.*, 1.02 ± 0.07 (16), and Chalevelakis *et al.*, 1.03 ± 0.14 (18). Thus no evidence for unbalanced globin synthesis was found in either the 30% or the 40% class, arguing against $\alpha$-thalassemia as the cause of the trimodality in the proportion of Hb G. Similar results have been reported by Tatsis (19), who found balanced globin synthesis in eight Hb G heterozygotes, including one with 45% Hb G. However, McCurdy *et al.* found $\alpha/\beta$ ratios ranging from 0.42 to 0.85 in seven double heterozygotes for $\alpha^G$ and either $\beta^C$ or $\beta^S$ (20). These data are difficult to interpret: while the $\alpha/\beta$ ratios seem to segregate into two groups with means of 0.50 and 0.76, these ratios are lower than those usually seen in $\alpha$-thal-1 trait and $\alpha$-thal-2 trait, respectively. In addition, the Hb G quantitation is obscured by the $\beta$-chain variant, so that one cannot ascertain if the lower $\alpha/\beta$ ratios correspond to the higher proportions of Hb G. We can only conclude that while synthesis may appear unbalanced in some Hb G heterozygotes, either because of $\alpha$-thalassemia or some technical problem in analysis of double heterozygotes, synthesis is clearly balanced in others.

There is other evidence against the $\alpha$-thalassemia model. In 1971, French and Lehmann (21) attempted to reconcile the high proportion of Hb G found in some heterozygotes with Lehmann and Carrell's two-locus model. Because the Hb G heterozygotes appeared microcytic, the authors attributed the 40% Hb G phenotype to linkage between an $\alpha^G$ and an $\alpha$-thalassemia gene, the homologous chromosome having two $\alpha^A$ genes. This now seems an unlikely explanation: since the 20%

class is most likely accounted for by one $\alpha^G$ gene plus three $\alpha^A$ genes, the 30% phenotype cannot be explained. Nor does the $\alpha$-thalassemia model proposed in Fig. 3 explain French and Lehmann's family (21) and one subsequently described by Martinez *et al.* (22), because children of 40% Hb G heterozygotes also had 40% Hb G. Therefore, under the $\alpha$-thalassemia model, the parents not heterozygous for $\alpha^G$ must be obligate $\alpha$-thalassemia silent carriers. However, the authors emphasized that they appeared entirely normal; no abnormality in hematologic parameters, red cell morphology, or osmotic fragility was found. (Biosynthetic studies were not performed.)

## Implications for the $\alpha$-thalassemia syndromes

Recent studies (4, 23–25) indicate that $\alpha$-thalassemia may be caused by at least partial deletion of the $\alpha$-chain structural genes. The varying degrees of severity in the $\alpha$-thalassemia syndromes are believed to reflect the number of $\alpha$-chain genes deleted. Viewed simplistically, the number of transcribed structural genes should be equivalent in the two models discussed. The relative microcytosis exhibited in persons with 40% and 30% Hb G could be interpreted as evidence for $\alpha$-thalassemia. However, globin synthesis is balanced in the small number of heterozygotes we have studied. This is in contrast to the $\alpha$-thalassemia syndromes, in which the synthetic balance is impaired. We therefore believe that the single-locus chromosome produces a phenotype hematologically similar to $\alpha$-thalassemia but whose underlying molecular mechanisms must be different.

With our present knowledge, we can only speculate as to what these differences might be. Perhaps lowering the gene dosage does decrease the absolute amount of $\alpha$-chain synthesized, but limitation of $\beta$-chain production by transcriptional and/or translational control mechanisms maintains balanced synthesis. In contrast, the lesion in $\alpha$-thalassemia must both diminish $\alpha$-chain synthesis and disturb the balancing mechanism.

These differences are especially interesting in the context of studies on Black persons with Hb H disease (26, 27). While these individuals showed a degree of microcytosis similar to that observed in Orientals and Italians with Hb H disease, the course of the disease was milder. Only about 5% of their hemoglobin was Hb H, while in other racial groups the proportion of Hb H ranges up to 30% (28). No Hb Bart's was observed. In six of these patients, biosynthetic studies yielded $\alpha/\beta$ ratios of 0.46, 0.56, 0.66, 0.69, 0.74, and 0.79, considerably higher than the $0.41 \pm 0.11$ reported by Kan *et al.* (16) in Oriental and Italian patients with Hb H disease. Thus the comparison of Hb H disease in Blacks with Hb H disease in Orientals and Italians is analogous to our comparison of 40% Hb G with $\alpha$-thal-1 trait: the hematologic results are similar and the synthesis results differ. Possibly the Black form of Hb H disease is caused by interaction between $\alpha$-thal-1 and a one-locus chromosome, resulting in a milder condition than the $\alpha$-thal-1/$\alpha$-thal-2 genotype associated with Hb H disease on other racial groups. Recently, Rieder *et al.* (29) have reported a person with mild Hb H disease, Hb G Philadelphia, and no Hb A. The $\alpha^G/\beta$ biosynthetic ratio was 0.66, consistent with the above explanation of Hb H disease in Blacks and with the existence of $\alpha^G$ on a one-locus chromosome. Perhaps some of the variability seen in the $\alpha$-thalassemia syndromes may be due to variation in the number of $\alpha$-chain loci.

We thank Dr. Rose Schneider and Dr. Marjory Cates for their generous contributions of blood samples, Dr. James V. Neel for useful criticisms and suggestions, Dr. Peter Smouse for statistical advice, Dr. Gerald Cederquist for assistance in computer programming and manuscript preparation, and Dr. Samir Hanash for performing several incubations. We also gratefully acknowledge the excellent technical assistance of Ms. Floretta Reynolds, Ms. Dorothy Sweet, and Ms. Catherine LaDu. This work was supported by Grants NIH-5-TO1-GM-71, 2-PO1-GM-15419, and NIH-HL-16008.

1. Lehmann, H. & Carrell, R. W. (1968) *Br. Med. J.* **4**, 748–750.
2. Hollan, S. R., Szelenyi, J. G., Brimhall, B., Duerst, M. & Jones, R. T. (1972) *Nature* **235**, 47–50.
3. De Jong, W. W., Meera Khan, P. & Bernini, L. F. (1975) *Am. J. Hum. Genet.* **27**, 81–90.
4. Kan, Y. W., Dozy, A. M., Varmus, H. E., Taylor, J. M., Holland, J. P., Lie-Injo, L. E., Ganesan, J. & Todd, D. (1975) *Nature* **255**, 255–256.
5. Abramson, R. K., Rucknagel, D. L., Shreffler, D. C. & Saave, J. J. (1970) *Science* **169**, 194–196.
6. Beaven, G. H., Hornabrook, R. W., Fox, R. H. & Huehns, E. R. (1972) *Nature* **235**, 46–47.
7. Trabuchet, G., Dahmane, M., Pagnier, J., Labie, D. & Benabadji, M. (1976) FEBS *Lett.* **61**, 156–158.
8. *International Hemoglobin Information Center Bulletin* (Comprehensive Sickle Cell Center, Augusta, Ga.) (1976) 1, No. 5.
9. Rucknagel, D. L. & Winter, W. P. (1974) *Ann. N.Y. Acad. Sci.* **241**, 80–92.
10. Winter, W. P. & Rucknagel, D. L. (1974) in *The Detection of Hemoglobinopathies*, eds. Schmidt, R. M., Huisman, T. H. J. & Lehmann, H. (CRC Press, Cleveland, Ohio), pp. 51–70.
11. Adams, J. G. III, Winter, W. P., Rucknagel, D. L. & Spencer, H. H. (1972) *Science* **176**, 1427–1429.
12. DeSimone, J., Kleve, L. & Shaeffer, J. (1974) *J. Lab. Clin. Med.* **84**, 517–524.
13. Clegg, J. B., Naughton, M. A. & Weatherall, D. J. (1965) *Nature* **207**, 945–947.
14. Rucknagel, D. L. & Rising, J. (1975) *Am. J. Med.* **59**, 53–60.
15. Nute, P. E. (1974) *Ann. N.Y. Acad. Sci.* **241**, 39–60.
16. Kan, Y. W., Schwartz, E. & Nathan, D. G. (1968) *J. Clin. Invest.* **47**, 2515–2522.
17. Pootrakul, S., Sapprapa, P., Wasi, P., Na-Nakorn, S. & Suwanik, R. (1975) *Humangenetik* **29**, 121–126.
18. Chalevelakis, G., Clegg, J. B. & Weatherall, D. J. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 3853–3857.
19. Tatsis, B. (1976) *Blood* **46**, 1029.
20. McCurdy, P. R., Sherman, A. S., Kamuzora, H, & Lehmann, H. (1975) *J. Lab. Clin. Med.* **85**, 891–897.
21. French, E. A. & Lehmann, H. (1971) *Acta Haematol.* **46**, 149–155.
22. Martinez, G., Vidal, H. & Colombo, B. (1973) *Hum. Hered.* **23**, 157–163.
23. Ottolenghi, S., Lanyon, W. G., Paul, J., Williamson, R., Weatherall, D. J., Clegg, J. B., Pritchard, J., Pootrakul, S. & Wong, H. B. (1974) *Nature* **251**, 389–391.
24. Taylor, J. M., Dozy, A., Kan, Y. W., Varmus, H. E., Lie-Injo, L. E., Ganesan, J. & Todd, D. (1974) *Nature* **251**, 392–393.
25. Ramirez, F., Natta, C., O'Donnell, J., Canale, V., Bailey, G., Sanguensermsri, T., Maniatis, G., Marks, P. & Bank, A. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 1550–1554.
26. Schwartz, E. & Atwater, J. (1972) *J. Clin. Invest.* **51**, 412–418.
27. Stamatoyannopoulos, G., Heywood, D. & Papayannopoulou, T. (1972) *Birth Defects, Orig. Artic. Ser.* **8**, 29–38.
28. Weatherall, D. J. & Clegg, J. B. (1972) *The Thalassemia Syndromes* (Blackwell Scientific Publications, Oxford), 2nd ed.
29. Rieder, R. F., Woodbury, D. H. & Rucknagel, D. L. (1976) *Br. J. Haematol.* **32**, 159–166.