Am. J. Hum. Genet. 42:345–352, 1988

Chromosomal Localization and Racial Distribution of the Polymorphic Human Dihydrofolate Reductase Pseudogene (DHFRP1)

Nicholas P. Anagnou,*[,1] Stylianos E. Antonarakis,† Stephen J. O'Brien,‡ William S. Modi,‡ and Arthur W. Nienhuis*

*Clinical Hematology Branch, National Heart, Lung and Blood Institute, National Institutes of Health, Bethesda, MD; †Department of Pediatrics, Genetics Unit, The Johns Hopkins University School of Medicine, Baltimore; and ‡Laboratory of Viral Carcinogenesis, National Cancer Institute, National Institutes of Health, Frederick, MD

Summary

The human dihydrofolate reductase (DHFR) gene family comprises one functional gene and at least four intronless processed pseudogenes. The functional DHFR gene is on chromosome 5, and DHFRP4 is on chromosome 3. Using in situ hybridization, we have now localized the functional DHFR gene to the region q11.1-q13.3 on chromosome 5. By genomic DNA analysis of a panel of human × rodent somatic-cell hybrids, we determined the chromosomal assignment of the DHFRP1 pseudogene to chromosome 18 and that of the DHFRP2 pseudogene to chromosome 6. The DHFRP1 pseudogene exhibits a novel form of polymorphism in humans in that it is present in the DNA of some individuals and absent in that of others. We investigated the racial distribution of this pseudogene in five racial groups. The allelic frequency as defined by analysis of 180 chromosomes was found to be 94% in Mediterraneans, 77% in Asian Indians, 67% in Chinese, 57% in Southeast Asians, and 32% in American blacks. These data suggest that the transposition of this "perfect" pseudogene occurred prior to the inception of the human racial groups.

Introduction

Dihydrofolate reductase (DHFR; 5,6,7,8-tetrahydrofolate [$NADP^+$ oxidoreductase, E.C.1.5.1.3]) plays an essential role in cellular metabolism and growth by converting dihydrofolate into tetrahydrofolate, an important methyl-group donor for the biosynthesis of purines, thymidine, and certain amino acids. We have isolated and characterized the functional human DHFR gene (DHFR) and three other pseudogenes (DHFRP1, DHFRP2, and DHFRP3) (Chen et al. 1982; Shimada et al. 1984) derived from processed RNA molecules. A fourth processed pseudogene (DHFRP4) has been described recently elsewhere (Masters et al. 1982). Thus, the DHFR gene family is comprised of one functional gene and at least four pseudogenes. We and others (Anagnou et al. 1984; Maurer et al. 1984) have shown that all five genes are dispersed in the human genome, each one residing on a different chromosome. The functional gene is on chromosome 5 (Anagnou et al. 1984; Funanage et al. 1984; Maurer et al. 1984), and DHFRP4 is on chromosome 3 (Anagnou et al. 1984; Maurer et al. 1985).

Contrary to the other pseudogenes, DHFRP1 exhibits an open reading frame and its sequences are identical to the coding sequences of the functional gene. The homology extends 2.9 kb beyond the end of the coding sequences (Chen et al. 1982). This perfect sequence homology implies that DHFRP1 has a recent evolutionary origin. In the course of our previous studies, we have found that this pseudogene is present in the DNA of some individuals but is absent in others, thus exhibiting in humans a novel form of polymorphism involving the insertion of the entire gene into an "empty" site of the human genome

Received February 19, 1987; revision received August 18, 1987.

Address for correspondence and reprints: Dr. Arthur Nienhuis, Clinical Hematology Branch, National Heart, Lung and Blood Institute, National Institutes of Health, Bethesda, MD 20205.

1. Present address: Department of Medicine, University of Washington, Seattle, WA 98195.

0002-9297/88/4202-0017$02.00

(Anagnou et al. 1984). The chromosomal localization of DHFRP1 and its racial distribution have now been determined.

Material and Methods

Cell Lines

A panel of 45 somatic-cell hybrid cell lines (human × mouse RAG and human × Chinese hamster E36) deficient for hypoxanthine phosphoribosyltransferase was used as previously described elsewhere (O'Brien et al. 1982, 1983; Anagnou et al. 1984). Extraction of high-molecular-weight DNA (O'Brien et al. 1983), isozyme analysis, and cytogenetic analysis were performed on the same passage of each cell line according to a method described elsewhere (O'Brien et al. 1982; Anagnou et al. 1984).

Blood Samples

For purposes of polymorphic screening, peripheral blood samples were obtained from a total of 90 unrelated individuals and β-thalassemia patients from five racial groups: American blacks, Mediterraneans (Greeks and Italians), Asian Indians, Chinese, and Southeast Asians. High-molecular-weight DNA was extracted according to a method described elsewhere (Anagnou et al. 1985).

Southern Blot Analysis

High-molecular-weight DNAs from the panel of somatic-cell hybrids and from the 90 individuals were digested with various restriction endonucleases, and the fragments were separated in 1.0% agarose gels, transferred to nitrocellulose filters, hybridized with nick-translated ^{32}P-labeled probes, and washed according to a method described elsewhere (Anagnou et al. 1984, 1985).

In Situ Hybridization

The procedure followed was essentially that of Harper and Saunders (1981). One-half microgram of DHFR 1.8-kb *Eco*RI DNA fragment was labeled, using three ^{3}H-labeled triphosphates, to a specific activity of 1.7×10^7 cpm/μg. Hybridization to metaphase chromosomes from fresh primary lymphocytes was performed at 37 C for 18 h in 50% formamide, 10% dextran sulfate, 2 × SSCP (300 mM NaCl, 30 mM trisodium citrate, 40 mM NaH_2PO_4/Na_2HPO_4, pH 7.2), the DNA probe at a concentration of 0.01–0.15 μg/ml, and sheared salmon-sperm DNA at a 500-fold higher concentration than that of the probe DNA. Following hybridization, the slides were washed five times at 39 C in 50% formamide, 2 × SSC and then in 2 × SSC before being dehydrated in an ethanol series. Slides were then dipped in Kodak NTB-2 and allowed to expose at 0 C for 4–30 days. After development in Dektol, slides were either Q-banded by means of immersion into a quinacrine-mustard solution (0.5 mg/ml) in McIlvaine's buffer (pH 5.5) or G-banded with Wright's stain.

Results

Assignment of the Polymorphic DHFRP1 Pseudogene to Chromosome 18

The chromosomal position of DHFRP1 was determined by means of DNA analysis of a panel of 45 somatic-cell hybrids, containing different combinations of human chromosomes, previously characterized elsewhere (O'Brien et al. 1983; Anagnou et al. 1984). On digestion with *Eco*RI and using a DHFRP1 DNA fragment containing the coding sequences of exons 2–6 as a probe, we detected diagnostic fragments (Anagnou et al. 1984) of each of the five human DHFR genes, as well of those of the mouse and hamster DHFR genes, as shown in figure 1*A*. DHFRP1 is contained in a 4.0-kb *Eco*RI fragment that comigrates with the 4.0-kb fragment containing exon 5 of the functional DHFR gene (Anagnou et al. 1984), as shown in figure 1*A*. Thus, on the basis of this approach, the presence or absence of DHFRP1 could be scored accurately only in cell lines that lack the functional DHFR gene, such as cell lines 80H-1 and 80H-12 (fig. 1*A*), in which DHFRP1 is present but DHFR-specific fragments are absent.

DHFRP1 was also studied in all hybrids, by using a 0.42-kb *Xba*I-*Eco*RI fragment from the cloned 3′-flanking region of DHFRP1 as a probe (Anagnou et al. 1984) (fig. 1*B*). This probe detects a 4.0-kb *Eco*RI fragment specific for DHFRP1 (fig. 1*B*). By combining data of the cytogenetics, isozyme analysis, and the DNA analysis of the panel of the hybrids, we assigned DHFRP1 to human chromosome 18. As shown in figure 2*A*, chromosome 18 and its included isozyme locus of peptidase A (PEPA) showed 98% and 100% concordance, respectively, with the 4.0-kb *Eco*RI fragment of DHFRP1 whereas the remaining chromosomes showed high discordancy (25%–60%). This high degree of concordance implies that the human donor of the panel is homozygous for the presence of DHFRP1. Thus, these data document the

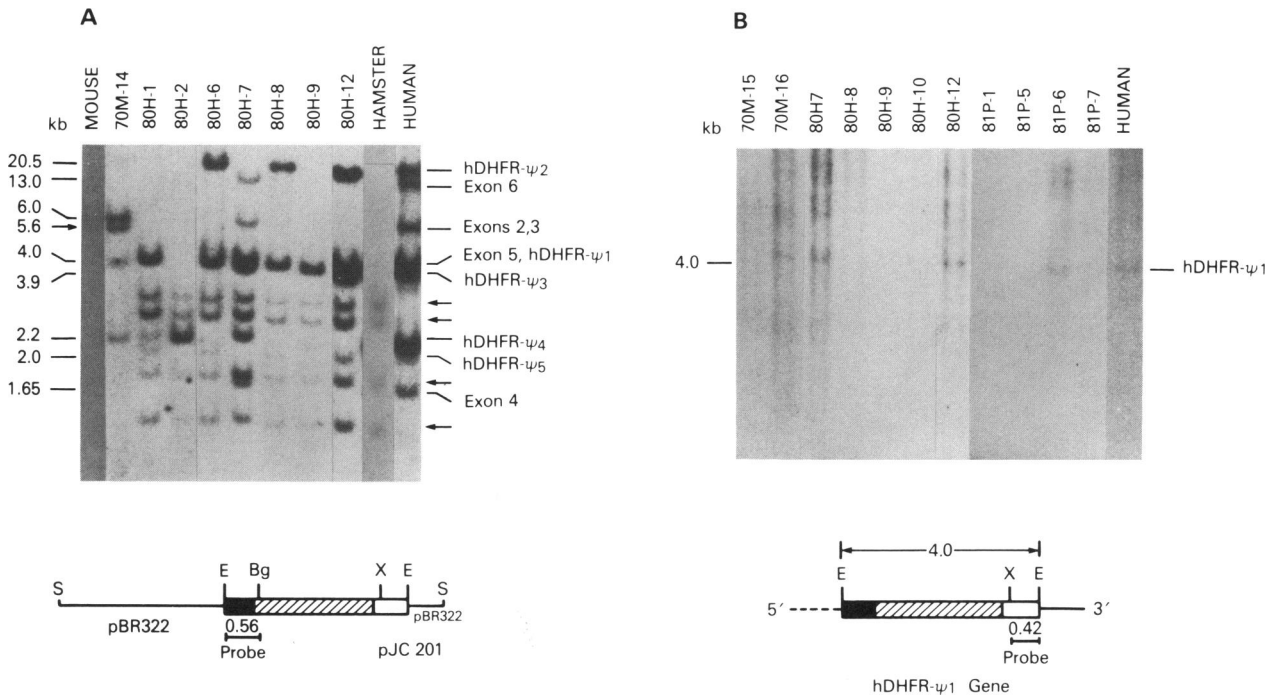

Figure 1 Southern blot analysis of DNAs from a somatic-cell hybrid panel used for the detection of the human DHFR genes. *A*, DNA from each cell line, either human × mouse (70 series) or human × hamster (80 and 81 series), was digested with *Eco*RI and probed with a 0.56-kb *Eco*RI-*Bgl*II fragment containing the coding sequences of exons 2–6, as shown in the diagram under the autoradiogram. This probe, as described elsewhere (Anagnou et al. 1984), generates diagnostic fragments specific for each of the five genes, as indicated on the right side of panel *A*. Hamster-specific bands are indicated with arrows on the right side of panel *A*. Mouse DNA generates two fragments (6.0 and 5.6 kb), as indicated on the left side of panel *A*. *B*, After being washed, the filters were rehybridized with a 0.42-kb *Xba*I-*Eco*RI probe from the 3′-flanking region of the cloned DHFRP1 (Chen et al. 1982; Anagnou et al. 1984). This probe generates a 4.0-kb diagnostic fragment for DHFRP1. E = *Eco*RI; B = *Bgl*II; X = *Xba*I; and S = *Sal*I. The designations used in figs. 1, 2, and 4 for the DHFR genes correspond to the Human Gene Mapping workshop nomenclature as follows: hDHFR = DHFR; hDHFR-ψ_1 = DHFRP1; hDHFR-ψ_2 = DHFRP2; hDHFR-ψ_3 = DHFRP3; and hDHFR-ψ_4 = DHFRP4.

assignment of this pseudogene to chromosome 18. Previous studies (Maurer et al. 1984) had provisionally assigned DHFRP1 to chromosome 3 or X. Our data clearly establish its assignment to chromosome 18, since both chromosomes 3 and X showed a very high discordancy vis-à-vis the presence of the 4.0-kb fragment.

Assignment of the DHFRP2 Pseudogene to Chromosome 6

Digestion of the DNA from the panel of the hybrids with *Eco*RI and hybridization with the 0.56-kb *Eco*RI-*Bgl*II probe containing exons 2–6 of the functional DHFR (fig. 1*A*) detects a 20.5-kb fragment that is specific for DHFRP2 (Chen et al. 1982; Anagnou et al. 1984), as shown in hybrids 80H-6, 80H-8, and 80H-12 (fig. 1*A*). DHFRP2 was 94%–98% concordant with human chromosome 6, as defined both cytogenetically and by the presence of the human isozymes of malic enzyme-1 (ME1) and glyoxylase (GLO) (fig. 2*B*). All of the remaining 22 chromosomes exhibited a high degree of discordancy (17%–55%). A previous study (Maurer et al. 1984) had implied that DHFRP2 is on chromosome 2 or chromosome 6. The present data provide definitive evidence for the assignment of DHFRP2 to human chromosome 6.

Regional Assignment of the Functional DHFR Gene to the q11.1-q13.3 Region of Chromosome 5

The functional DHFR gene has been assigned to chromosome 5 by this and other laboratories previously (Anagnou et al. 1984; Funanage et al. 1984; Maurer et al. 1984). The regional localization is unresolved, since different positions have been reported by two groups that utilized somatic-cell hybrids containing chromosomes with interstitial deletions (Funanage et al. 1984; Maurer et al. 1985). To resolve this conflict, in the present paper we report the

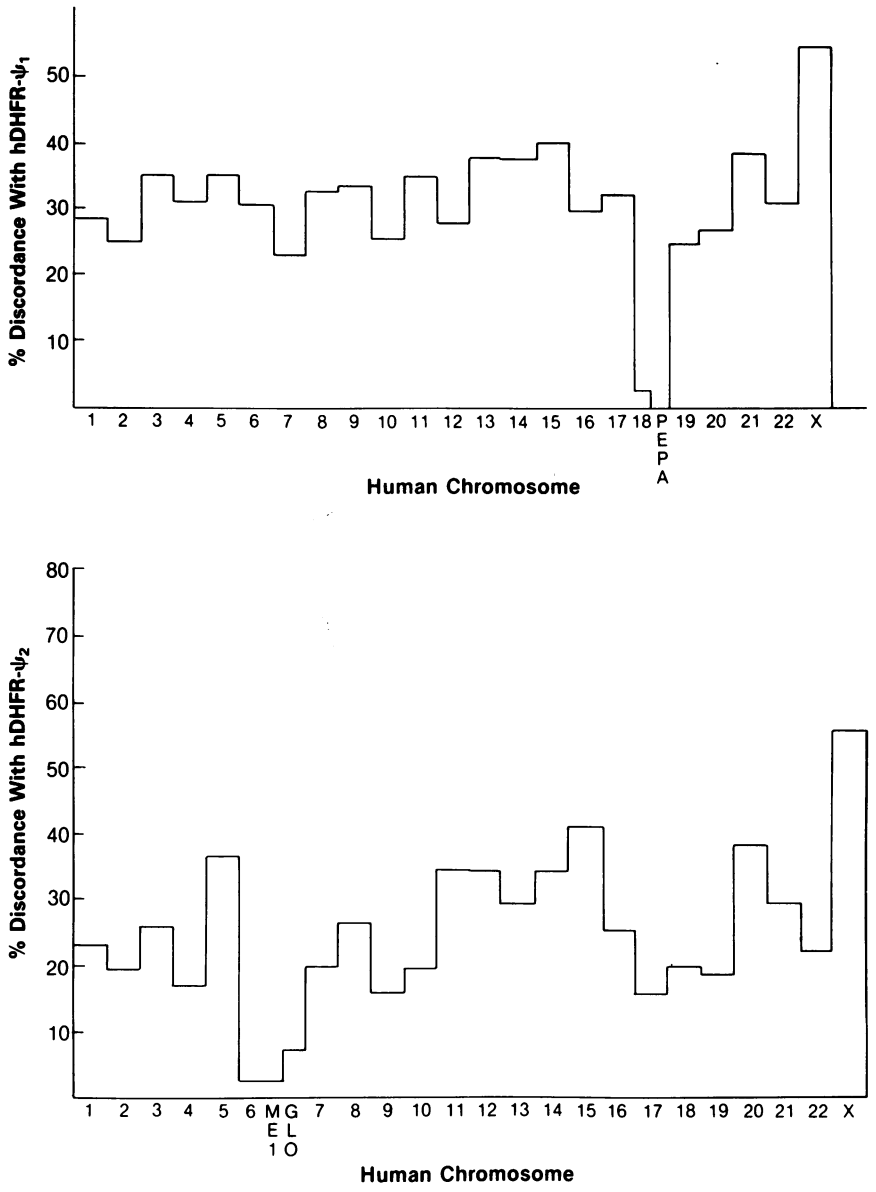

Figure 2 Analysis of the association of DHFRP1 (hDHFR-ψ_1)- and DHFRP2 (hDHFR-ψ_2)-specific fragments with segregant chromosomes in a panel of 45 somatic-cell hybrids. Chromosome scores represent the consensus result of karyotyping (G-banding) and isozyme scores. Thirty-six isozyme systems diagnostic for each human chromosome were tested for each hybrid (O'Brien et al. 1982, 1983; Anagnou et al. 1984). PEPA = Peptidase A; ME1 = malic enzyme-1; and GLO = glyoxylase.

results of in situ hybridization of the functional DHFR gene. As a probe we used a 1.8-kb *Eco*RI fragment from the 5′ end of the functional DHFR gene (Anagnou et al. 1984). A total of 183 silver grains were distributed along the chromosomes of 71 metaphase cells examined (fig. 3*A*). Of this total, a significant number of grains—i.e., 25 (14.2%)—were found along chromosome 5, and 15 of them were localized on the chromosomal region 5q11.1-q13.3, as shown in figure 3*B*. No other sites were labeled above background. We conclude from these data that the functional DHFR gene is located on bands q11.1-q13.3 of chromosome 5. These data are in agreement with a previous study using deletion-mapping analysis (Funanage et al. 1984) but do not agree with the results of another study (Maurer et al.

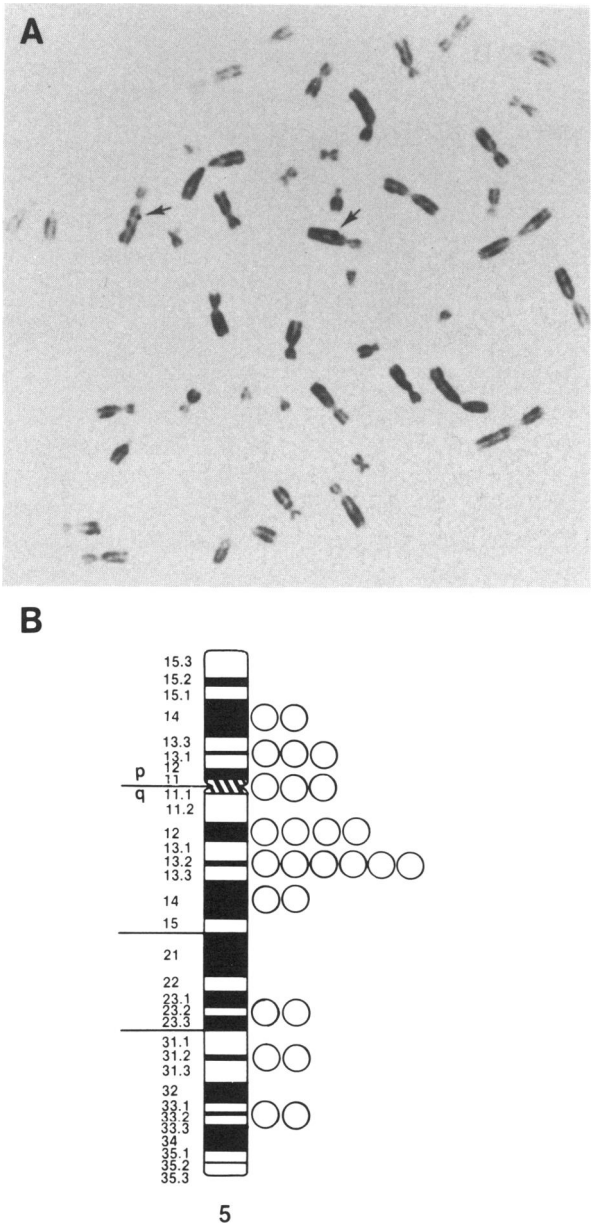

Figure 3 Localization of the functional DHFR gene by using in situ hybridization. *A*, Q-banded metaphase with labeled chromosome 5 (arrows). *B*, Ideogram illustrating grain distribution along chromosome 5.

1985), which indicated localization of the gene to band q23 of chromosome 5.

Racial Distribution of the Polymorphic DHFRP1 Pseudogene

In our previous study (Anagnou et al. 1984) we found that DHFRP1 is present in the DNA of some individuals but is absent in that of others. The presence or absence of the pseudogene creates different restriction sites in the two types of corresponding chromosomes (DHFRP1[+] and DHFRP1[−]), as shown in figure 4; this results in different restriction fragments that are specific for each type of chromosome. *Pst*I and *Bgl*II digestion identify three patterns: homozygotes for the presence of the pseudogene (+/+), homozygotes for its absence (−/−), and heterozygotes for the presence and absence of the pseudogene (+/−) (fig. 4). To determine the presence of the polymorphic DHFRP1, we analyzed 180 human chromosomes from 90 individuals belonging to five different racial groups. The results are shown in table 1. The pseudogene was present in all five racial groups, and its allelic frequency ranged from 94% in Mediterraneans to 32% in American blacks.

Discussion

In the present study, the processed polymorphic DHFRP1 was assigned to chromosome 18, DHFRP2 was assigned to chromosome 6, and the regional localization of the functional DHFR gene was determined by means of in situ hybridization to region q11.1-q13.3 of chromosome 5. Furthermore, DHFRP1 was found to be polymorphic in the five distinct human groups studied. The assignment of DHFRP1 to chromosome 18 provides this chromosome with a polymorphic marker useful for linkage analysis of unassigned genes associated with genetic diseases. The thymidylate synthase gene (coding for an enzyme that utilizes tetrahydrofolate as a cofactor) also has been recently assigned (Nussbaum et al. 1985) to chromosome 18, providing a selectable marker for this chromosome.

The polymorphic DHFRP1 was detected in the DNA of individuals from all five racial groups studied. The allelic frequency did not differ significantly among Mediterraneans, Asian Indians, Chinese, or Southeast Asians. In American blacks the pseudogene exhibited its lowest frequency. It is conceivable that the allelic frequency found in the American blacks in the present study might be falsely high, since this population carries alleles of white origin at a frequency of ~20% (Reed 1969). Thus, these data are not sufficient to support the hypothesis that this low frequency in American blacks might represent another Black-specific DNA polymorphism previously described for certain loci, such as those of α-globin (Goodburn et al. 1984), β-globin (Kan and

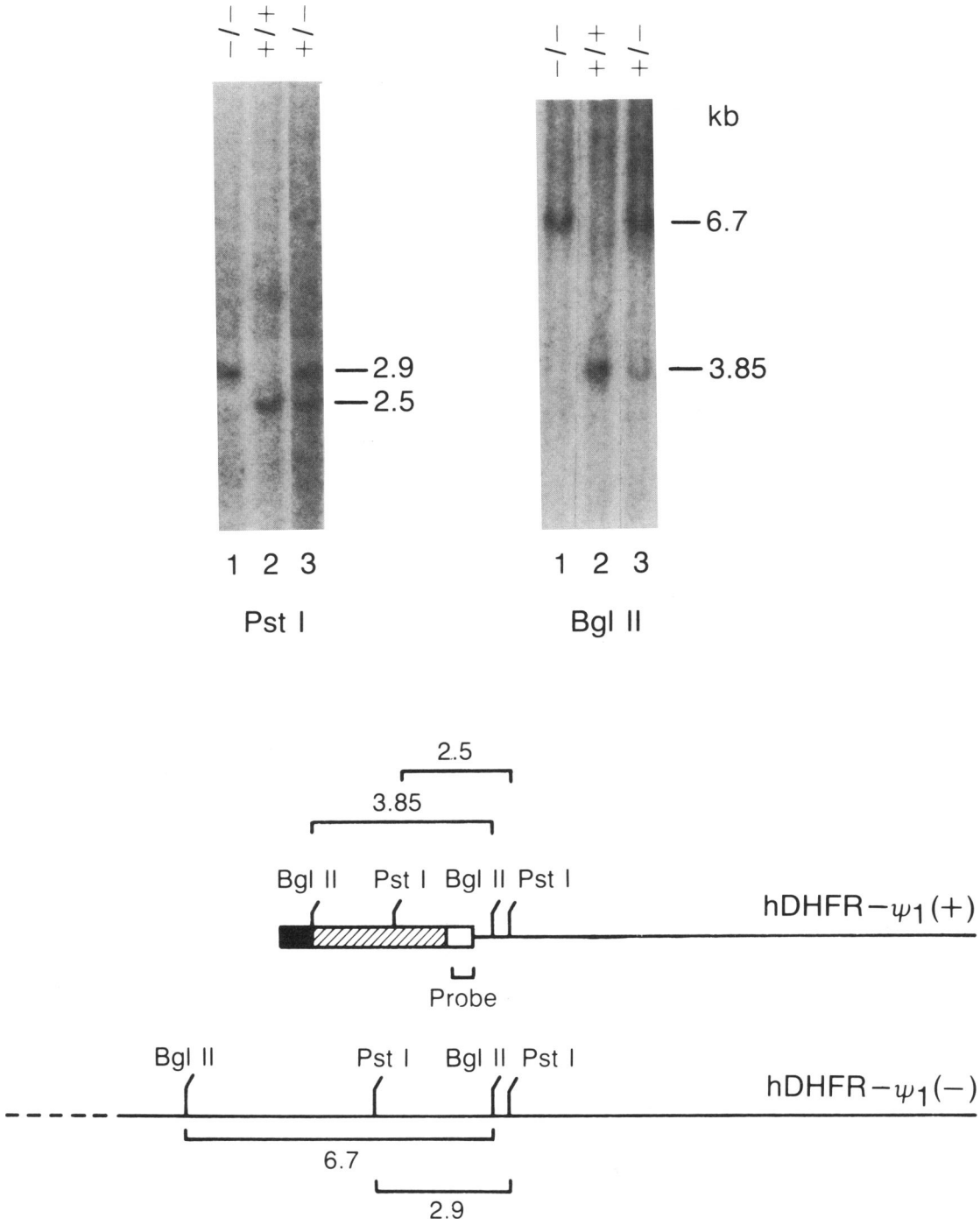

Figure 4 Polymorphism of *DHFRP1* (hDHFR-ψ_1) in humans. This processed pseudogene is present in the DNA of some individuals (DHFRP1 [+]) and is absent in others (DHFRP1 [−]). Digestion of DNA with *Pst*I or *Bgl*II and hybridization with the specific 0.42-kb *Xba*I-*Eco*RI probe (see diagram and fig. 1) generates specific fragments for chromosomes positive (2.5 and 3.85 kb, respectively) and negative (2.9 and 6.7 kb, respectively) for the pseudogene. Thus, screening of DNA for the presence (+) or absence (−) of the gene identifies homozygotes (+/+) for its presence, heterozygotes (+/−), and homozygotes for its absence (−/−).

Table 1

Racial Distribution of the Polymorphic Processed DHFRP1 Pseudogene

Racial Group	No. of Chromosomes with DHFRP1 Gene/Total No. of Chromosomes (%)
Mediterraneans	36/38 (94.7)
Asian Indians	31/40 (77.5)
Chinese	23/34 (67.6)
Southeast Asians	16/28 (57.1)
American blacks	13/40 (32.5)

Dozy 1980), growth hormone (Chakravarti et al. 1984), insulin (Bell et al. 1984), and albumin (Murray et al. 1984). On the other hand, limited investigations of DNAs from several homogeneous African-blacks populations, such as Mbuti Pygmies (Zaire) and Biaka Pygmies (Central African Republic), indicate allelic frequency of this pseudogene that is comparable with that seen in the five groups considered in the present paper (N. P. Anagnou, unpublished data). Thus, the presence of the pseudogene in all racial groups suggests that the molecular event of transposition of this "perfect" pseudogene occurred prior to the divergence of the human races. Furthermore, the perfect homology of the pseudogene to the functional DHFR gene indicates that this event must have occurred very recently in evolution.

The regional assignment of the functional DHFR gene was determined by using in situ hybridization. Our data clearly indicate that the gene is regionally located on 5q11.1-q13.3, close to the centromere, and are in agreement with results of a previous study using deletion mapping analysis (Funanage et al. 1984). The assignment of the functional DHFR gene to bands 5q11.1-q13.3 seems the most likely correct one for two reasons. First, the employment of in situ hybridization permits precise localization of the gene to a specific band. In our study, the majority of the grains were localized over the region 5q11.1-q13.3 (fig. 3) whereas only two grains, constituting background levels, were found on 5q23. Second, by using the functional DHFR gene as a selectable marker, we have segregated the two homologues of chromosome 5 from a patient with a $5q^-$ syndrome into somatic-cell hybrids between Chinese hamster-ovary cells and bone-marrow cells (Nienhuis et al. 1985). Cytogenetic analysis showed that the breakpoint was within 5q15-q33 (Nienhuis et al. 1985). If the suggested localization of the functional DHFR gene had been on 5q23 (Maurer et al. 1985), it would have been impossible to select for chromosome 5 in those hybrids by using the functional DHFR gene as a selectable marker. We therefore believe that our data and those of a previous study (Funanage et al. 1984) document the regional assignment of the functional DHFR gene to 5q11.1-q13.3.

Acknowledgments

This study was supported by grants from the AHEPA Cooley's Anemia Foundation to N.P.A. and a National Institutes of Health grant to S.E.A. We are grateful to Drs. Kenneth Kidd (Yale University) and Luca Cavalli-Sforza (Stanford University) for providing us with DNA samples from Biaka and Mbuti pygmies. We thank Mary McMahon and Jeanita Danzik for excellent secretarial services.

References

Anagnou, N. P., S. J. O'Brien, T. Shimada, W. Nash, M. J. Chen, and A. W. Nienhuis. 1984. Chromosomal organization of the dihydrofolate reductase genes: dispersion, selective amplification, and a novel form of polymorphism. Proc. Natl. Acad. Sci. USA 81:5170–5174.

Anagnou, N. P., T. Papayannopoulou, G. Stamatoyannopoulos, and A. W. Nienhuis. 1985. Structurally diverse molecular deletions in the β-globin gene cluster exhibit an identical phenotype on interaction with the β^S gene. Blood 65:1245–1251.

Bell, G., S. Horita, and J. Karam. 1984. A polymorphic locus near the human insulin gene is associated with insulin-dependent diabetes mellitus. Diabetes 33:176–189.

Chakravarti, A., J. Phillips, K. Mellits, K. Buetow, and P. Seeburg. 1984. Patterns of polymorphism and linkage disequilibrium suggest independent origins of the human growth hormone gene cluster. Proc. Natl. Acad. Sci. USA 81:6085–6089.

Chen, M. J., T. Shimada, A. Davis-Moulton, M. Harrison, and A. W. Nienhuis. 1982. Intronless human dihydrofolate reductase genes are derived from processed RNA molecules. Proc. Natl. Acad. Sci. USA 79:7435–7439.

Funanage, V., T. Miyoda, P. Moses, and H. Cowell. 1984. Assignment of the human dihydrofolate reductase gene to the q11-q22 region of chromosome 5. Mol. Cell. Biol. 4:2010–2016.

Goodburn, S., D. Higgs, J. Clegg, and D. G. Weatherall. 1984. Allelic variation and linkage properties of a highly polymorphic restriction fragment in humans. Mol. Biol. Med. 2:223–238.

Harper, M., and J. Saunders. 1981. Localization of single copy DNA sequences on G-banded human chromo-

somes by *in situ* hybridization. Chromosoma **83**:431–439.

Kan, Y. W., and A. Dozy. 1980. Evolution of the hemoglobin S and C genes in world populations. Science **209**:388–391.

Masters, J., J. Yang, A. Cellini, and G. Attardi. 1982. A human dihydrofolate reductase pseudogene and its relationship to the multiple forms of specific messenger RNA. J. Mol. Biol. **167**:2–36.

Maurer, B., P. Barker, J. Masters, F. Ruddle, and G. Attardi. 1984. Human dihydrofolate reductase gene is located in chromosome 5 and is unlinked to the related pseudogenes. Proc. Natl. Acad. Sci. USA **81**:1484–1488.

Maurer, B., L. Carlock, J. Wasmuth, and G. Attardi. 1985. Assignment of human dihydrofolate reductase gene to band q23 of chromosome 5 and of related pseudogene ψHD1 to chromosome 3. Somatic Cell Mol. Genet. **11**:79–85.

Murray, J., K. Mills, C. Demopulos, S. Horung, and A. Motulsky. 1984. Linkage disequilibrium and evolutionary relationships of DNA variants (restriction enzyme fragment length polymorphisms) at the serum albumin locus. Proc. Natl. Acad. Sci. USA **81**:3486–3490.

Nienhuis, A. W., F. Bunn, P. H. Turner, T. V. Gopal, W. Nash, S. J. O'Brien, and C. Sherr. 1985. Expression of the human c-fms proto-oncogene in hematopoietic cells and its deletion in the $5q^-$ syndrome. Cell **42**:421–428.

Nussbaum, R., R. Walmsley, J. Lesko, S. Airhart, and D. Ledbetter. 1985. Thymidylate synthetase–deficient Chinese hamster cells: a selection system for human chromosome 18 and experimental system for the study of thymidylate synthetase regulation and fragile X expression. Am. J. Hum. Genet. **37**:1192–1205.

O'Brien, S. J., W. G. Nash, J. L. Goodwin, D. Lowy, and E. H. Chang. 1983. Dispersion of the *ras* family of transforming genes to four different chromosomes in man. Nature **303**:839–842.

O'Brien, S. J., J. M. Simonson, and M. Eichelberger. 1982. Genetic analysis of hybrid cells using isozyme markers as monitors of chromosome segregation. Pp. 342–370 *in* J. W. Shaw, ed. Techniques in somatic cell genetics. Plenum, New York.

Reed, T. 1969. Caucasian genes in American Negroes. Science **165**:762–768.

Shimada, T., M. J. Chen, and A. W. Nienhuis. 1984. A human dihydrofolate reductase intronless pseudogene with an *Alu* repetitive sequence: multiple DNA insertions at a single chromosomal site. Gene **31**:1–8.