*Am. J. Hum. Genet. 42:297–308, 1988*

# A Linkage and Physical Map of Chromosome 22, and Some Applications to Gene Mapping

Cécile Julier,*†[1] G. Mark Lathrop,† Abderrezak Reghis,‡ Marie-France Szajnert,* Jean-Marc Lalouel,† and Jean-Claude Kaplan*

*Institut de Pathologie Moléculaire, Institut National de la Santé et de la Recherche Médicale U129, Paris; †Howa d Hughes Medical Institute, University of Utah Medical Center, Salt Lake City; and ‡Centre National de Transfusion Sanguine, CHU Mustapha, Alger, Algeria

## Summary

**A genetic map of human chromosome 22 has been derived from physical assignments and multilocus linkage analysis. It consists of the loci for the immunoglobulin lambda light-chain variable (IGLV) and immunoglobulin lambda light-chain constant (IGLC) regions, myoglobin (MB), the *sis* proto-oncogene (SIS), and an arbitrary probe (D22S1). The first RFLPs at the loci for SIS, IGLV, and MB are described. The most likely gene order on the basis of multilocus analysis was cen-(IGLV-IGLC)-D22S1-MB-SIS. This map provides further evidence for localization of the P1 polymorphism of the P blood group to chromosome 22, close to the SIS locus. Analysis of families segregating recessive congenital methemoglobinemia (RCM), a disease in which the cytochrome b5 reductase is defective, as well as of families with cases of hereditary low levels of cytochrome b5 reductase activity, confirmed that the locus responsible for RCM is on chromosome 22. Biochemical studies had already suggested that mutation at the cytochrome b5 reductase locus (DIA1) is responsible for RCM. We found no evidence of genetic heterogeneity between the families segregating RCM and the families exhibiting cases of low cytochrome b5 reductase activity. Linkage analysis indicated that the most probable location of DIA1 lies between MB and SIS.**

## Introduction

Genetic maps of the human chromosomes will increase our understanding of the genetics of man and especially of his hereditary diseases, in most of which the affected genes or the fundamental biochemical defects, or both, are unknown. Genetic linkage maps of some chromosomes have already been published (see, e.g., Drayna et al. 1984; White et al. 1985; Leppert et al. 1986), and the availability of an increasing number of polymorphic DNA sequence markers dispersed on human chromosomes should make possible a complete linkage map of the whole genome in the near future (Botstein et al. 1980).

Received April 28, 1987; revision received August 24, 1987.

Address for correspondence and reprints: Dr. Cécile Julier, INSERM U91, Hôpital Henri Mondor, 94010 Créteil, France.

1. Present address: Institut National de la Santé et de la Recherche Médicale U91, Hôpital Henri Mondor, 94010 Créteil, France.


0002-9297/88/4202-0012$02.00

Chromosome 22 is one of the two smallest human chromosomes, corresponding to ~1.7% of the male haploid genome (Mendelsohn et al. 1973) or ~50,000 kb. We have constructed a genetic linkage map of chromosome 22 on the basis of the analysis of five loci: the immunoglobulin lambda light-chain variable (IGLV) and immunoglobulin lambda light-chain constant (IGLC) regions, myoglobin (MB), the *sis* proto-oncogene (SIS), and a locus defined by an arbitrary probe (D22S1). The preliminary map has been applied to the still controversial localization of the gene for the P1 blood-group antigen (Nielsen et al. 1984).

We have also applied this map to localization of the gene causing recessive congenital methemoglobinemia (RCM). It has been hypothesized that this disorder is due to a homozygous deficiency in cytochrome b5 reductase activity, caused by mutation(s) in the gene coding for this enzyme (the DIA1 locus) (Leroux et al. 1975; Kaplan et al. 1979). The DIA1 locus has been assigned to chromosome 22 in some reports (Fisher et al. 1977; Junien et al. 1978). In the

present study, linkage analyses of families segregating type II RCM were used to confirm the localization of the disease gene to chromosome 22, as well as to investigate possible genetic heterogeneity between type II RCM and hereditary forms of low cytochrome b5 reductase activity that show no clinical evidence of disease.

A few linkage analyses of chromosome 22 loci have been published, documenting linkage between P1 and RCM (Scott and Wright 1969), P1 and DIA1 (McAlpine et al. 1977), DIA1 and 22p (Cook et al. 1978), and P1 and transcobalamin II (Eiberg et al. 1986). Preliminary data from our own work were reported as abstracts at the Human Gene Mapping 8 conference (Julier et al. 1985*a*, 1985*b*, 1985*c*).

## Material and Methods

### Subjects

Genetic polymorphisms and allelic frequencies were determined from a sample population including 10 unrelated French subjects, 15 unrelated Algerian subjects, and 31 spouse pairs from French families provided by the Centre d'Étude du Polymorphisme Humain (CEPH). Twelve nuclear families with a total of 68 children were also selected from the CEPH panel, to test procedures and biological specimens in a subset that would provide maximum heterozygosity and large sibships. Additionally, nine pedigrees, all but one of which were of Algerian origin, were investigated. Four of the nine, including one of French origin, were ascertained through a proband having a diagnosis of type II RCM. We called these pedigrees "type D (for *d*isease) pedigrees." The remaining five pedigrees were ascertained through an individual with cytochrome b5 reductase activity lower than the 95th percentile of its normal distribution in a reference Algerian population. Since this reduced activity has been hypothesized to result from other mutations at the RCM locus (Reghis et al. 1981, 1983), we called these "type Q (for *q*uantitative) pedigrees." In the present study, we tentatively designated the disease loci as RCM.D and RCM.Q for the two types of pedigree.

### DNA Probes

The six DNA probes studied here defined five polymorphic genetic systems. Both polymorphisms detected by the arbitrary DNA sequence pMS3.18 (locus D22S1) and by the 2.3-kb noncoding sequence pCIVS2.3 (locus IGLC) have been described by other groups of investigators (Taub et al. 1983; Naylor et al. 1984). We characterized new polymorphisms for the following probes: an 8-kb *Eco*RI genomic probe, V4A, isolated from a chromosome 22 library and containing an IGLV gene corresponding to the subtype $V_\lambda 0$ (IGLV locus) (Anderson et al. 1984); an 8-kb *Eco*RI genomic probe, pHM27.B2.9, containing the first exon of the myoglobin gene and flanking sequences (Jeffreys et al. 1984; Weller et al. 1984); and two probes for the locus of the proto-oncogene c-*sis* (SIS locus), a 1-kb *Xba*I-*Pst*I v-*sis* probe (pvsis) from the simian sarcoma virus oncogene (Robbins et al. 1981) and a 1.8-kb *Bam*HI subclone (pcsisBam) from the human genomic clone L33 (Dalla-Favera et al. 1982) containing a c-*sis* sequence.

### Somatic Cell Hybrids

Sixteen human-mouse or human–Chinese hamster somatic cell hybrids were used to confirm the physical location of probes V4A, pHM27.B2.9, and pMS3.18 on chromosome 22. Regional localization was carried out with hybrid cells (H, Saza, and Eaza) derived from a (X;22)(q13;q11.2) translocation retaining either the proximal part or the distal part of chromosome 22 and some other human chromosomes (Hors-Cayla et al. 1981).

### Southern Transfers

DNA for Southern blot analysis was isolated either from blood samples or from cultured fibroblasts or lymphoblasts, according to published methods (Grunebaum et al. 1984). The DNA was digested by the appropriate restriction enzyme, and the fragments were separated on agarose gels according to standard protocols. DNA was transferred on DBM paper (Alwine et al. 1979), or on zeta-probe membrane (BioRad) as specified by the manufacturer. The membranes were hybridized to radiolabeled DNA probes, and the final wash of the hybridized membranes was in 0.1% SDS and 2 × SSC at 60 C (probes pCIVS2.3, pcsisBam, and pMS3.18), 2 × SSC at 42 C (probe pvsis), or 0.1 × SSC at 68 C (probes V4A and pHM27.B2.9).

### RFLPs

The pMS3.18 probe reveals polymorphism with restriction enzymes *Bgl*II and *Taq*I (Barker et al. 1984). With the probe pCIVS2.3, four insertion-deletion alleles are detected with *Eco*RI (Taub et al. 1983). We searched for RFLPs with probes V4A, pvsis, and pHM27.B2.9, which define, respectively,

**Table 1**

**Mean ± SD NADH Cytochrome b5 Reductase Activity in Heterozygotes and Normal Members of Type II RCM Pedigrees**

| Genotype | Erythrocyte-soluble Enzyme ($\mu$mol min$^{-1}$ g Hb$^{-1}$) | Lymphocyte Total Enzyme (nmol min$^{-1}$ mg protein$^{-1}$) |
|---|---|---|
| +/+ ..... | 0.93 ± 0.22 (46) | 157.2 ± 13.9 (28) |
| +/d ..... | 0.20 ± 0.21 (36) | 38.7 ± 12.8 (36) |

NOTE.—Numbers in parentheses are numbers of individuals.

the loci IGLV, SIS, and MB. For each of 17 restriction enzymes (*Hin*dIII, *Eco*RI, *Pst*I, *Bam*HI, *Xba*I, *Msp*I, *Bgl*II, *Kpn*I, *Sac*I, *Bst*EII, *Bgl*I, *Eco*RV, *Apa*I, *Xho*I, *Hae*III, *Rsa*I, and *Taq*I) two DNA samples were run simultaneously, one corresponding to 10 $\mu$g of a DNA pool from 20 unrelated individuals and the other consisting of 10 $\mu$g of DNA from one of these individuals. When a difference between the two lanes was found, we sought confirmation of polymorphism by analyzing DNA from 10 unrelated individuals that was digested with the same restriction enzyme and by screening the family sample to verify allelic exclusion.

### *Assay of NADH Cytochrome b5 Reductase Activity*

Red cell–soluble enzyme and leukocyte total enzyme activity were measured by using the method of Hegesh as described by Reghis et al. (1981). In the three type D Algerian pedigrees, both assays had been done previously (Reghis 1981); leukocyte total activity in these families allowed us to distinguish normal homozygotes from heterozygotes, the values being nonoverlapping. The nonambiguous determinations of genotype in the three disease pedigrees enabled us to establish means and SDs for erythrocyte enzyme activity in normal homozygotes and heterozygotes (table 1); these values were used for analysis of the type Q pedigrees. Although only red cell–soluble enzyme was assayed in the French type D pedigree, the results were sufficiently unambiguous that the trait could be considered codominant in this pedigree also. In the five type Q pedigrees only the red blood cell–soluble enzyme was measured.

### *P Blood-Group Typing*

The individuals were analyzed for the P blood group by using a P1 antiserum; the locus studied here contains the polymorphic gene responsible for conversion of the paragloboside to the antigen P1 (Naiki and Marcus 1975). We designated this locus P1 for mapping purposes.

### *Linkage Analysis*

Linkage analysis was performed with the computer program package LINKAGE (Lathrop et al. 1984), which permits linkage tests, recombination-frequency ($\theta$) estimates ($\hat{\theta}$'s), and tests of gene order by considering the joint segregation of two or more loci. Prior to linkage analysis we reduced the number of alleles at each locus by using an extension of Ott's method (Ott 1978) described elsewhere (Julier 1984). In most of the linkage analyses, inbreeding in the Algerian pedigrees was neglected. However, we verified that the $\hat{\theta}$ values and test of linkage were unaffected regardless of whether the closest inbreeding loop was considered in the analyses. We calculated $\chi^2$ values as the $-2\ln L$ difference between the two alternative hypotheses.

## Results

### *Physical Assignments*

The loci SIS and IGLC had already been confirmed as mapping to chromosome 22 (Tippett and Kaplan 1985). Although IGLV must map close to IGLC for functional reasons, the complex pattern observed with the IGLV probe, even under very stringent conditions, raises the possibility that some of the fragments detected with probe V4A are unlinked to IGLV itself. To confirm the localizations of D22S1 (probe pMS3.18), MB (probe pHM27.B2.9), and all the multiple hybridizing fragments of the IGLV probe to chromosome 22, we hybridized these probes and pCIVS2.3 (IGLC) to DNA from a panel of 16 somatic cell hybrids. All the probes gave results concordant with the presence or absence of chromosome 22 except in one cell line (L53N; this exception probably resulted from mistyping or from a rear-

**Table 2**

**Regional Localization of Probes V4A, pCIVS2.3, pMS3.18, and pHM27.B2.9 on Chromosome 22**

| Hybrid Cell Lines | Part of Chromosome 22 Present | Probes for Loci | | | | |
|---|---|---|---|---|---|---|
| | | IGLV | IGLC | D22S1 | MB | SIS |
| H ..... | 22pter→q11.2 | + | + | − | − | − |
| Saza ..... | 22q11.2→qter | − | − | + | + | + |
| Eaza ..... | ... | − | − | − | − | − |

Note.—The same probes as in table 2, together with pcsisBam (locus SIS), were hybridized to a Southern blot of cell hybrids containing different fractions of human chromosome 22 as the only human complement.

rangement or a translocation to the mouse of chromosome 22). The results confirmed the assignments of D22S1, MB, and the numerous V4A fragments to chromosome 22. Furthermore, hybridization of V4A and other immunoglobulin lambda–region probes with Burkitt lymphoma cell line BL2, which bears an 8;22 translocation, showed that in this cell line all the fragments that hybridize with V4A probe are deleted, together with fragments hybridizing with some C-lambda-region probes (data not shown). This result supports the contention that all fragments hybridizing to the V4A probe are clustered in a single region—i.e., in the neighborhood of the C-lambda gene cluster—and can therefore be treated as a single locus in linkage analysis.

Table 2 reports results of hybridization of the probes c-sisBam, pMS3.18, V4A, pCIVS2.3, and pHM27.B2.9 to DNA from hybrids containing partial chromosome 22. The SIS locus—at 22q12.3→q13.1 as confirmed by in situ hybridization (see review in Tippett and Kaplan 1985) and therefore distal to the breakpoint at 22q11.2 in the somatic hybrid line H—was included as a control. The probes for loci IGLV and IGLC mapped proximally to the breakpoint at 22q11.2, whereas the probes for MB and D22S1, as well as that for SIS, mapped distally to it. Studies of the translocation breakpoint in Burkitt lymphoma cell lines having translocations t(8;22) have demonstrated that the V-lambda genes map proximally to the C-lambda genes in 22q11 (Emanuel et al. 1984; Hollis et al. 1984). The physical locations of the loci studied here are shown next to the ideogram in figure 1.

### *Identification of New RFLPs*

The strategy outlined above led to the characterization of polymorphisms at the SIS, IGLV, and MB loci with the probes pvsis, V4A, and pHM27.B2.9, respectively. Our results, together with the polymorphisms at loci D22S1 and IGLC that have been described by others, are summarized in table 3; polymorphisms at the SIS locus, the MB locus, and the IGLV locus are illustrated in figure 2. At the IGLV locus, preliminary analysis of the pooled DNA sample versus the single individual DNA sample with V4A suggested polymorphisms with many enzymes, including *Hin*dIII, *Bam*HI, *Bgl*II, *Bst*EII, *Bgl*I, *Taq*I, and *Kpn*I (data not shown). The analysis of unrelated individuals confirmed the existence of RFLPs for all these enzymes.

The IGLV polymorphisms revealed by *Taq*I, *Hin*dIII, and *Kpn*I were studied further. With *Taq*I we found six distinct phenotypic patterns among 55 individuals analyzed, each consisting of one or two fragments with strong signals (fig. 2C). The data are compatible with a three-allele system in Hardy-Weinberg equilibrium ($\chi^2_1 = 0.61$), with alleles corresponding to major fragments of 12.9, 9.2, and 4.2 kb, respectively. Likewise, three *Hin*dIII patterns could be interpreted in terms of two alleles that correspond to polymorphic fragments of 2.5 and 2.3 kb, respectively (data not shown). These alleles were also in Hardy-Weinberg equilibrium ($\chi^2_1 = 0.31$). Furthermore, the presence or absence of a 3.6-kb fragment defined a second *Hin*dIII allelic series (data not shown). Although the absence of the 3.6-kb fragment correlated with the presence of a faint band of 12.3 kb, this polymorphism was poorly defined and avoided for scoring families. *Kpn*I revealed polymorphism involving three alleles, each characterized by a major polymorphic fragment (of 19.0, 17.5, and 14.1 kb, respectively [data not shown]).

The four allelic series revealed by these three enzymes exhibited strong linkage disequilibrium. Among 35 unrelated individuals whose phase could be determined, the *Hin*dIII 2.5-kb fragment and the *Taq*I 12.9-kb fragment were in complete linkage disequilibrium. Only three of the six possible haplotypes

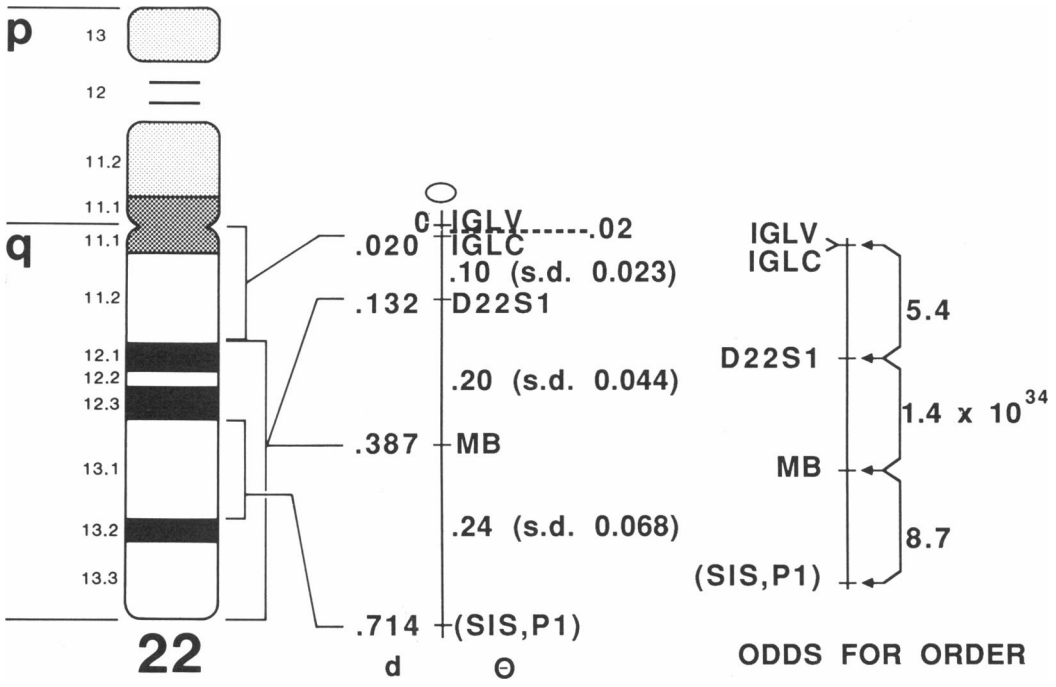


**Figure 1** Regional localization of probes for loci on chromosome 22, the genetic map of five loci, and the odds against permutations of the best-supported gene order. $\hat{\theta}$ Values (uppercase theta) within each interval, and SDs (s.d.) are based on combined male and female meioses. Map scale within the linkage group (d = distance) is in Morgans, measured from the IGLV locus.

were observed for the first *Hin*dIII series and for the *Taq*I series. These are defined by the 2.5-kb *Hin*dIII and 12.9-kb *Taq*I fragments ($N$ = 41), the 2.3-kb *Hin*dIII and 9.2-kb *Taq*I fragments ($N$ = 20), and the 2.3-kb *Hin*dIII and 4.2-kb *Taq*I fragments ($N$ = 9). All other individuals could be interpreted in terms of these three haplotypes and were thus fully characterized in terms of the *Taq*I allelic system alone. The absence of the 3.6-kb *Hin*dIII fragment was positively associated with the presence of the 12.9-kb *Taq*I fragment.

The *Taq*I and *Kpn*I series yielded four of the nine theoretically possible haplotypes among 28 unrelated individuals with known phase. These four are defined by the 19.0-kb *Kpn*I and 12.9-kb *Taq*I fragments ($N$ = 24), the 14.1-kb *Kpn*I and 9.2-kb *Taq*I fragments ($N$ = 10), the 17.5-kb *Kpn*I and 9.2-kb *Taq*I fragments ($N$ = 16), and the 17.5-kb *Kpn*I and 4.2-kb *Taq*I fragments ($N$ = 6). All other individuals could be interpreted in terms of these haplotypes.

### *Linkage Analysis*

The phenotypes of individual members of 21 kindreds were tested for the polymorphic loci previously described (data not shown). In the CEPH families, offspring were tested only when at least one parent was a double heterozygote for any two of the genetic polymorphisms studied. Linkage analysis proceeded as follows: (1) $\hat{\theta}$ values were obtained from all pairwise linkage tests; (2) tests of gene order were performed for the DNA markers; (3) linkage of the P1 locus to chromosome 22 markers was investigated, and a new map was generated to include this locus; and (4) the placement of the RCM locus was evaluated by means of location scores.

*Pairwise linkage tests.*—The results of pairwise linkage tests for all the loci considered are summarized in table 4. Since these tests yielded no evidence of significant sex differences in $\theta$ values, subsequent analyses were performed assuming equal $\theta$ values in males and females.

As expected, the IGLV and IGLC loci are tightly linked, with a single recombinant detected. The SIS locus gives a lod score of 1.67 with the P1 blood group, with a $\hat{\theta}$ of 3%. No obligate rcombination between these loci could be identified; a nonzero $\hat{\theta}$ is nevertheless possible because of dominance at the P1 locus.

Since the other marker loci tested here are known to be located on chromosome 22, many of the linkage tests reported in table 4 can be evaluated by reference

## Table 3

**Polymorphisms at Loci D22S1, IGLC, IGLV, MB, and SIS, with the Allele and Haplotype Frequencies Determined among Unrelated CEPH Individuals**

| Locus (Probe) and Enzyme | Allele or Haplotype Designations | Variable Fragments[a] (kb) | Frequencies ± SD | Hetero-zygosity |
|---|---|---|---|---|
| D22S1 (pMS3.18): | | | | |
| *Bgl*II ..... | B1 | 10 | .70 ± .04 | .42 |
| | B2 | 7 | .30 ± .04 | |
| *Taq*I ..... | T1 | 1.5 1.0 | .97 ± .02 | .06 |
| | T2 | 2.5 | .03 ± .02 | |
| *Bgl*II + *Taq*I ..... | B1T1 | | .68 ± .05 | .45 |
| | B1T2 | | .29 ± .04 | |
| | B2T1 | | .03 ± .02 | |
| IGLC (CIVS2.3): | | | | |
| *Eco*RI ..... | 1 | 8 | .88 ± .03 | .22 |
| | 2 | 13 | .05 ± .02 | |
| | 3 | 18 | .07 ± .02 | |
| | 4 | 23 | .00[b] | |
| IGLV (V4A): | | | | |
| *Taq*I ..... | T1 | 12.9 9.2 | .54 ± .05 | .59 |
| | T2 | 11.0 9.2 | .32 ± .04 | |
| | T3 | 9.2 4.2 2.1 | .14 ± .03 | |
| *Kpn*I ..... | K1 | 19.0 | ... | |
| | K2 | 17.5 10.7 | ... | |
| | K3 | 14.1 | ... | |
| *Taq*I + *Kpn*I ..... | T1K1 | | .39 ± .07 | .70[c] |
| | T1K3 | | .15 ± .05 | |
| | T2K2 | | .32 ± .04 | |
| | T3K2 | | .14 ± .03 | |
| *Hin*dIII (H0/H2) ..... | H0 | 12.3 | .31 ± .04 | .43 |
| | H2 | 3.6 | .69 ± .04 | |
| SIS (v-sis): | | | | |
| *Hin*dIII ..... | 1 | 21.0 | .66 ± .04 | .45 |
| | 2 | 14.5 6.5 | .34 ± .04 | |
| MB (pHM27.B2.9): | | | | |
| *Taq*I ..... | 1 | 1.4 | .38 ± .05 | .47 |
| | 2 | 1.1 | .62 ± .05 | |

[a] Strongly hybridizing fragments observed with V4A are underlined.
[b] Appeared once in an Algerian family but was not found in the CEPH panel.
[c] Rises to .80 if the *Hin*dIII-HO/H2 system is combined with the *Taq*I + *Kpn*I–defined haplotypes.

**Figure 2** *A*, RFLP at locus SIS as revealed with probe pvsis and enzyme *Hin*dIII. The three observed phenotypes 1, 2, and 3 can be interpreted by alleles 1 (21 kb) and 2 (14.5 and 6.5 kb) as being the corresponding genotypes 1/1, 1/2, and 2/2, respectively. Weak hybridization is noted with a 5-kb fragment, whose intensity and/or presence depends on the stringency of the last wash. This polymorphism is also detected with pc-sisBam, giving fragments of 21 and 14.5 kb for alleles 1 and 2, respectively. Published restriction maps of the SIS locus (Dalla-Favera et al. 1982; Chiu et al. 1984; Johnsson et al. 1984; Josephs et al. 1984) all exhibit a pattern compatible with allele 1—except that of Johnsson et al. (1984), which bears a *Hin*dIII site between exons 3 and 4, compatible with allele 2. *B*, RFLP at locus MB as revealed by probe pHM27.B2.9 and enzyme *Taq*I. The three observed phenotypes 1, 2, and 3 can be interpreted by alleles 1 (presence of the 1.6-kb band) and 2 (presence of the 1.3-kb band) as being the corresponding genotypes 2/2, 1/2, and 1/1, respectively. *C*, RFLPs at locus IGLV, revealed by probe V4A and enzyme *Taq*I. The six phenotypes can be interpreted by the three alleles T1, T2, and T3 as being genotypes 1:T1/T1, 2:T2/T2, 3:T3/T3, 4:T1/T2, 5:T1/T3, and 6:T2/T3.

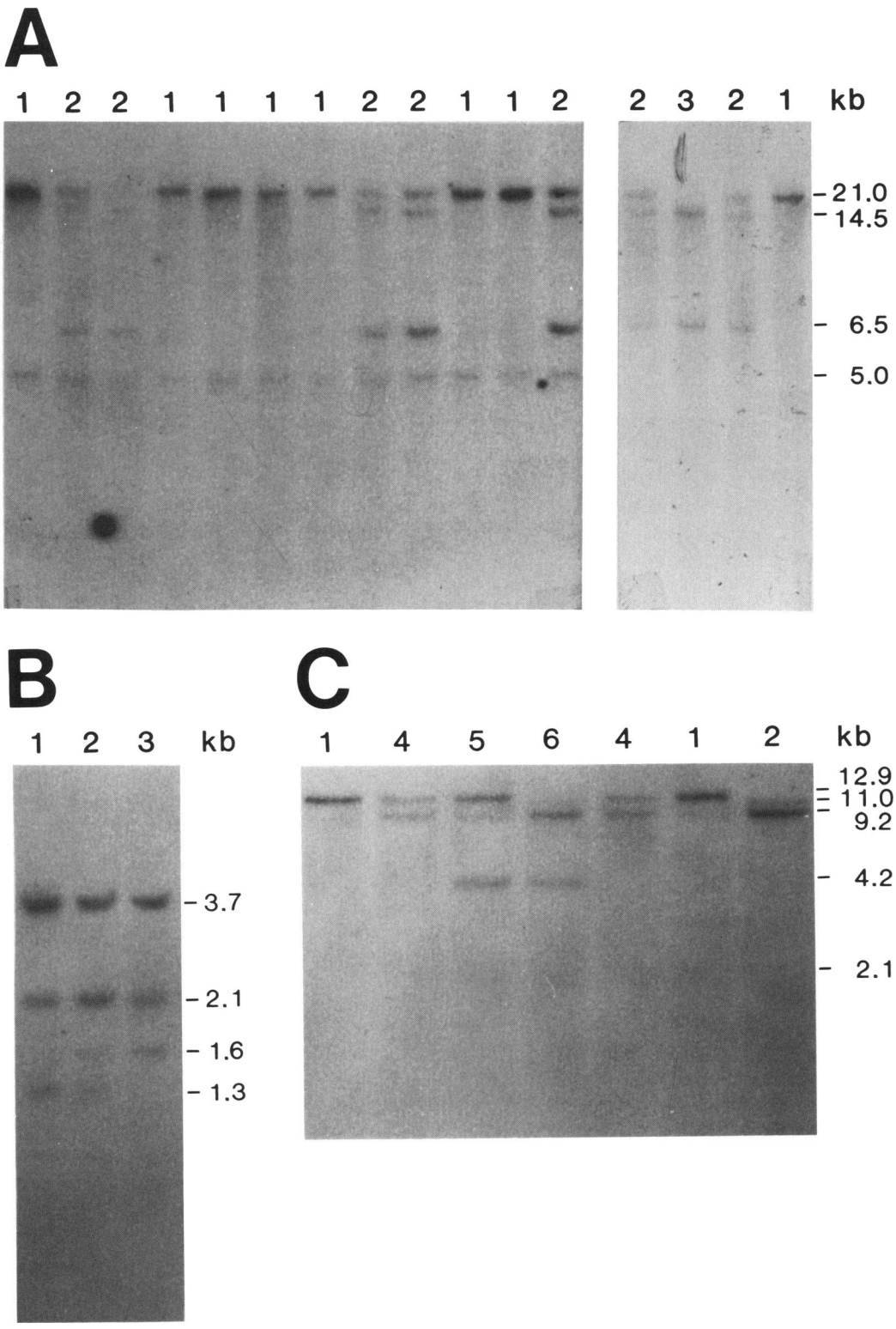
A
1 2 2 1 1 1 1 2 2 1 1 2
2 3 2 1 kb
21.0
14.5
6.5
5.0
B
1 2 3 kb
3.7
2.1
1.6
1.3
C
1 4 5 6 4 1 2 kb
12.9
11.0
9.2
4.2
2.1

**Table 4**

**Pairwise Linkage Analysis of All Loci**

| Loci | $\hat{\theta} \pm$ SD | $\chi^2_1$ | Lod Score | $\chi^2_1$ Sex[a] |
|---|---|---|---|---|
| IGLV-IGLC | .02 ± .017 | 37.80 | 8.21 | 0.78 |
| IGLV-D22S1 | .08 ± .021 | 4.40 | 0.96 | 1.20 |
| IGLV-MB | .29 ± .063 | 1.56 | 0.34 | 0.56 |
| IGLV-SIS | .38 ± .087 | 0.30 | 0.07 | 2.56 |
| IGLV-P1 | .63 ± .130 | 0 | 0 | 0 |
| IGLV-RCM.D | .30 ± .077 | 1.53 | 0.33 | 0.55 |
| IGLC-D22S1 | .16 ± .053 | 7.66 | 1.66 | 2.21 |
| IGLC-MB | .43 ± .180 | 0.02 | 0 | 0.42 |
| IGLC-SIS | .50 ± .031 | 0 | 0 | 0.33 |
| IGLC-P1 | .25 ± .120 | 0.69 | 0.15 | 0.03 |
| IGLC-RCM.D | .25 ± .130 | 1.11 | 0.24 | 0 |
| D22S1-MB | .14 ± .042 | 10.63 | 2.31 | 1.31 |
| D22S1-SIS | .30 ± .088 | 0.66 | 0.14 | 1.57 |
| D22S1-P1 | .51 ± .160 | 0 | 0 | 0.17 |
| D22S1-RCM.D | .06 ± .043 | 9.22 | 2.00 | 1.25 |
| MB-SIS | .08 ± .058 | 4.54 | 0.99 | 2.10 |
| MB-P1 | .09 ± .064 | 3.97 | 0.86 | 2.48 |
| MB-RCM.D | .11 ± .075 | 4.49 | 0.97 | 0.50 |
| SIS-P1 | .03 ± .049 | 7.68 | 1.67 | 0.00 |
| SIS-RCM.D | .50 ± .360 | 0 | 0 | 0.00 |
| P1-RCM.D | .23 ± .210 | 0.19 | 0.04 | 0.06 |

[a] Based on the likelihood-ratio statistic comparing the hypothesis of different $\theta$ values for the two sexes vs. the hypothesis of equal $\theta$ values.

to a $\chi^2$ table for 1 df. Several of the lod scores are significant when evaluated this way.

*Multilocus linkage analysis.*—To construct a genetic map of chromosome 22, we directed our attention first to the five DNA genetic systems IGLV, IGLC, D22S1, MB, and SIS. Table 5 summarizes all possible three-point tests. The three-point analysis shows that the loci D22S1, SIS, and MB are very unlikely to reside between IGLV and IGLC. The odds against an interior position are 284:1 for D22S1, $6.7 \times 10^7$:1 for SIS, and $8.7 \times 10^7$:1 for MB. Tight linkage between IGLV and IGLC is consistent with the physical analysis presented above. Four-locus analyses allowed us to eliminate the following two other orders: (1) location of SIS between (IGLV-IGLC) and D22S1 (odds of $3.3 \times 10^7$:1) and (2) location of MB between (IGLV-IGLC) and D22S1 (odds of $2.6 \times 10^3$:1). The relative order of IGLV and IGLC with respect to the other loci could not be determined on the basis of linkage methods because only a single recombinant was observed between these loci.

To provide further evidence of gene order, linkage analysis was performed with (IGLV-IGLC), D22S1, SIS, and MB jointly. After consideration of the results obtained from the three-point and four-point analyses, we had now only six different orders to test, from the 12 original possibilities. The likelihoods and odds ratios for these six orders are given in table 6.

Several alternative gene orders cannot be rejected on the basis of the results of the linkage analysis.

**Table 5**

**Results of Three-Locus Linkage Analyses for the Loci IGLV, IGLC, D22S1, MB, and SIS**

| Loci and Maximum-Likelihood Gene Order (1-2-3) | Odds against Alternative Order | |
|---|---|---|
| | 2-1-3 | 1-3-2 |
| IGLV-IGLC-D22S1 | 1.6:1 | 284.1 |
| IGLV-IGLC-MB | 1.1:1 | $8.7 \times 10^7$:1 |
| IGLC-IGLV-SIS | 1.1:1 | $6.7 \times 10^7$:1 |
| IGLV-D22S1-MB | 95.1:1 | 32.5:1 |
| IGLV-D22S1-SIS | 1.3:1 | 50.1:1 |
| IGLV-MB-SIS | 7.5:1 | 1.0:1 |
| IGLC-D22S1-MB | 4.2:1 | $2.6 \times 10^3$:1 |
| D22S1-IGLC-SIS | 1.3:1 | $3.3 \times 10^7$:1 |
| IGLC-MB-SIS | 3.7:1 | 1.3:1 |
| D22S1-MB-SIS | 3.0:1 | 9.0:1 |

**Table 6**

**Five-Point Linkage Analysis between Loci D22S1, IGLC, IGLV, SIS, and MB**

| Order | $\hat{\theta}$ | | | | −2lnL | Odds[a] |
|---|---|---|---|---|---|---|
| IGLV-IGLC-D22S1-MB-SIS .... | .025 | .096 | .218 | .166 | 1332.11 | 1.0 |
| IGLV-IGLC-D22S1-SIS-MB .... | .033 | .076 | .323 | .084 | 1336.13 | 7.5 |
| MB-IGLV-IGLC-D22S1-SIS .... | .285 | .021 | .086 | .378 | 1339.77 | 46.1 |
| SIS-IGLV-IGLC-D22S1-MB .... | .375 | .024 | .098 | .211 | 1334.90 | 4.0 |
| SIS-MB-IGLV-IGLC-D22S1 .... | .154 | .302 | .023 | .085 | 1337.25 | 13.1 |
| MB-SIS-IGLV-IGLC-D22S1 .... | .095 | .419 | .024 | .088 | 1337.60 | 15.6 |

[a] Inverses of the likelihood ratios against the most likely order (IGLV-IGLC-D22S1-MB-SIS).

Nevertheless, since the analysis of somatic cell hybrids showed that IGLV and IGLC are proximal to the other three loci, the only two possibilities for order that remain are (IGLV-IGLC)-D22S1-MB-SIS and (IGLV-IGLC)-D22S1-SIS-MB, with relative odds of 7.5:1 in favor of the former.

*The P1 blood group.*—The two-point analyses suggested close linkage of P1 with SIS and MB. Location scores for P1 were calculated with the LINKMAP program on the basis of the maximum-likelihood marker map defined above. These show that the maximum-likelihood location for P1 coincides with that of the SIS locus. The odds in favor of linkage of P1 to the chromosome 22 map are 52.5:1, equivalent to 1.72 on the lod scale. Other linkage studies (McAlpine et al. 1978) involving the P1 blood-group locus have yielded a lod score of ~2.5 between P1 and DIA1, with an estimated $\hat{\theta}$ of .10–.15. Because different loci were used in these studies, the lod scores cannot be summed. Nevertheless, the combined evidence strongly suggests localization of P1 to chromosome 22. Since no obligate recombinants were observed between P1 and SIS, the $\theta$ between these loci was fixed at .0 in subsequent analyses.

*Revision of the marker map.*—Since the preliminary map yielded odds of only 7.5:1 in favor of the order D22S1-MB-SIS versus the alternative order of D22S1-SIS-MB, we reconsidered both possibilities with the addition of P1. Under these conditions, the odds in favor of the order D22S1-MB-(SIS-P1) versus the order D22S1-(SIS-P1)-MB became 112:1 ($\chi^2_1$ = 9.4). Assuming then that the order (IGLV-IGLC)-D22S1-MB-(SIS-P1) was established, we obtained the best $\hat{\theta}$ values between these loci by finding the maximum-likelihood estimates when all the markers—IGLV, IGLC, D22S1, MB, and (SIS-P1)—were used together. For this calculation, we combined IGLV and IGLC into a single locus by ignoring the single recombinant individual between them. The genetic map thus obtained, together with relative odds for the given gene order, is illustrated in figure 1.

*Tentative localization of the RCM locus on the map.*—Linkage analysis of our four type D and five type Q families by means of the method of location scores (Lathrop et al. 1984) established that the RCM gene is located on chromosome 22 (fig. 3). In calculating location scores, we assumed the gene order and $\hat{\theta}$ values established in the previous section. The maximum-likelihood estimate of the map location for RCM places this locus between MB and (SIS-P1), with location scores of 8.88 and 2.62, respectively, for the D and Q pedigrees. Because the estimated location is identical for the two types of pedigree, a

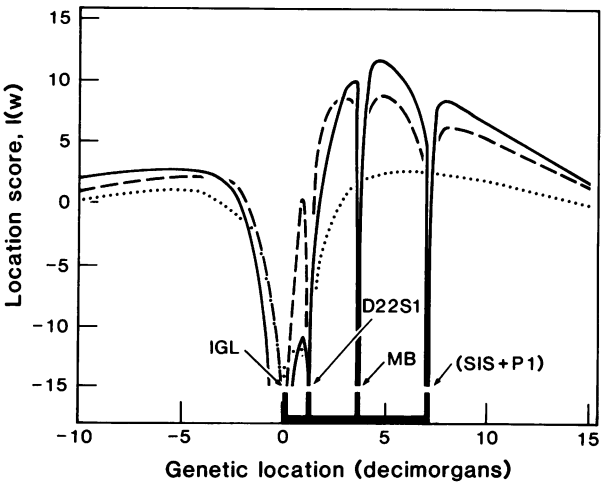


**Figure 3** Support for the linkage of the RCM locus to other markers on the chromosome 22 genetic map. Twice the natural logarithm of the likelihood ratio contrasting the hypothesis of the loci for RCM.D (— — —), RCM.Q (· · · ·), and combined D + Q (———) being at location w to the hypothesis of their being infinitely removed from the linkage group, l(w), is expressed as a function of the genetic location of RCM.

$\chi^2$-test shows no evidence of heterogeneity and the location scores can be summed to obtain a maximum of 11.5. Figure 3 shows, however, that alternative locations for the RCM gene on the map of chromosome 22 cannot be rejected.

### Discussion

The present study applies physical methods of gene localization, as well as linkage analysis, to realize a preliminary genetic map of chromosome 22 that includes several polymorphic markers and the gene for RCM. Problems that we encountered in developing the map of chromosome 22 suggest the following two points that may be important in inferring genetic maps of other chromosomes from a small sample of families: (1) In our sample, it was not possible to show significant effects of either sex difference or interference (data not shown for the latter), although sex effects have been reported for other chromosomes in some large-scale studies (White et al. 1985; Leppert et al. 1986) and interference is likely to occur. Therefore we considered a model in which both effects are absent. Since this means that fewer parameters must be estimated, this type of model may be more powerful for establishing a linkage map when data are derived from a small population sample, since the influence of random fluctuations due to small sample size is likely to be reduced compared with more complex models. (2) Multilocus linkage analysis is the most efficient way to show linkage or to order loci. This degree of efficiency is important when dealing with loci of poor information content and/or with small samples; for example, the added power of four-point compared with three-point linkage analysis is illustrated by our determination of the order of MB and SIS relative to D22S1. In a three-point analysis we obtained odds of 9:1 for the order D22S1-MB-SIS versus the order D22S1-SIS-MB. When the P1 locus was added to this system ($\hat{\theta} = .00$ between P1 and SIS), the odds ratio became 112:1 for the order D22S1-MB-(SIS + P1) versus the order D22S1-(SIS + P1)-MB.

Our analysis supports the hypothesis of linkage of P1 with markers on chromosome 22. Although we cannot cumulate the location score of 1.72 (log base 10) for P1 with the lod score of ~2.5 between P1 and DIA1 that was reported by McAlpine et al. (1978), when taken together these two results provide strong evidence for the localization of P1 to chromosome 22. It has been suggested on the basis of immunological analysis of somatic cell hybrids that P is on chromosome 6 (Fellous et al. 1971). The apparent discrepancy between linkage studies and the immunological evidence might be explained both by the complexity of the P blood-group system and by the fact that the physical localization was based on serological detection of both P and P1 antigens (M. Fellous, personal communication). In fact, several genes may be involved in the expression of these antigens (Naiki and Marcus 1975).

The preliminary map derived from combination of our physical and linkage results permitted us to demonstrate linkage of the gene responsible for type II RCM to chromosome 22, a result supporting the hypothesis that the DIA1 and RCM loci are identical (Leroux et al. 1975; Kaplan et al. 1979). We find that the most likely location for the RCM locus is between MB and SIS, although the evidence presented here does not preclude other locations. The physical localization of DIA1 to q13.31→qter (Francke et al. 1982) would tend to place it distally to the SIS locus (q13.1); although that location for DIA1 does not correspond to our maximum-likelihood estimate, again the linkage analysis does not preclude it. Our studies yielded no evidence of genetic heterogeneity between type II RCM and hereditary forms of low-level cytochrome b5 reductase activity, a result supporting the Reghis et al. (1981, 1983) hypothesis that the DIA1 locus is involved in the latter phenotype also.

Despite its small size, chromosome 22 is involved in a relatively high number of inherited and acquired diseases (for review, see Kaplan et al. 1987). The preliminary map presented here could potentially help in localizing the corresponding morbid loci at the DNA level.

## Acknowledgments

We thank A. Blanchetot, D. Barker, G. Bornkamm, and K. Robbins for providing us with probes pHM27.B2.9, pMS3.18, pCIVS2.3, pcsisBam, and pvsis. We are also grateful to Drs. Nguyen Van Cong and M. C. Hors-Cayla for making available to us their somatic cell hybrids, and to CEPH for providing DNAs from their reference families.

## References

Alwine, J. C., D. J. Kemp, B. A. Parker, J. Reiser, J. Renart, G. R. Stark, and G. M. Wahl. 1979. Detection of specific RNAs or specific fragments of DNA by fractionation in

gels and transfer to diazobenzyloxymethyl paper. Methods Enzymol. 68:220–242.

Anderson, M. L. M., M. F. Szajnert, J. C. Kaplan, L. McColl, and B. D. Young. 1984. The isolation of a human IgV gene from a recombinant library of chromosome 22 and the estimation of its copy number. Nucleic Acids Res. 12:6647–6661.

Barker, D., M. Schafer, and R. White. 1984. Restriction sites containing CpG show a higher frequency of polymorphism in human DNA. Cell 36:131–138.

Botstein, D., R. White, M. Skolnick, and R. W. Davis. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms (RFLPs). Am. J. Hum. Genet. 32:314–331.

Chiu, I. M., E. P. Reddy, D. Givol, K. C. Robbins, S. R. Tronick, and S. A. Aaronson. 1984. Nucleotide sequence analysis identifies the human c-sis proto-oncogene as a structural gene for PDGF. Cell 37:123–129.

Cook, P. J., K. E. Buckton, and J. A. Robinson. 1978. Family studies with chromosomes 21 and 22. Cytogenet. Cell Genet. 22:518–520.

Dalla-Favera, R., R. C. Gallo, A. Gillongo, and C. M. Croce. 1982. Chromosomal localization of the human homolog (*c-sis*) of the simian sarcoma virus oncogene. Science 218:686–688.

Drayna, D., K. Davies, D. Hartley, J. L. Mandel, G. Camerino, R. Williamson, and R. White. 1984. Genetic mapping of the human X chromosome by using restriction fragment length polymorphisms. Proc. Natl. Acad. Sci. USA 81:2836–2839.

Eiberg, H., N. Moeller, J. Mohr, and L. S. Nielsen. 1986. Linkage of transcobalamin II (TC2) to the P blood group system and assignment to chromosome 22. Clin. Genet. 29:354–359.

Emanuel, B. S., J. R. Selden, E. Wang, P. C. Nowell, and C. M. Croce. 1984. *In situ* hybridization and translocation breakpoint mapping. Cytogenet. Cell Genet. 38:127–131.

Fellous, M., C. Billardon, J. Dausset, and J. Frezal. 1971. Linkage probable entre HLA et P. C. R. Acad. Sci. 272:3356–3359.

Fisher, R. A., S. Povey, M. Bobrow, E. Solomon, Y. Boyd, and B. Carritt. 1977. Assignment of the DIA1 locus to chromosome 22. Ann. Hum. Genet. 41:151–155.

Francke, U., P. Tetri, R. T. Taggart, and N. Oliver. 1982. Regional mapping of DIA1, ARSA and ACO2 on chromosome 22 using hybrids with a t(15;22)(q14;q13.31) translocation. Cytogenet. Cell Genet. 32: 276–277.

Grunebaum, L., J. P. Cazenave, G. Camerino, C. Kloepfer, J. L. Mandel, P. Tolstoshev, M. Jaye, H. De La Salle, and J. P. Lecocq. 1984. Carrier detection of hemophilia B by using a restriction site polymorphism associated with the coagulation factor IX gene. J. Clin. Invest. 73:1491–1495.

Hollis, G. F., K. F. Mitchell, J. Battey, H. Potter, R. Taub, and G. M. Lenoir. 1984. A variant translocation places the immunoglobulin genes 3′ to the c-Myc oncogene in Burkitt's lymphoma. Nature 307:752–755.

Hors-Cayla, M. C., C. Junien, S. Heuertz, J. F. Mattei, and J. Frezal. 1981. Regional assignment of arylsulfatase A, mitochondrial aconitase and ADH-cytochrome b5 reductase by somatic cell hybridization. Hum. Genet. 58:140–143.

Jeffreys, A. J., V. Wilson, A. Blanchetot, P. Weller, A. Geurt van Kessel, N. Spurr, E. Solomon, and P. Goodfellow. 1984. The human myoglobin gene: a third dispersed globin locus in the human genome. Nucleic Acids Res. 12:3235–3243.

Johnsson, A., C. H. Heldin, A. Wasteson, B. Westermark, T. F. Deuel, J. S. Huang, P. H. Seebury, A. Gray, A. Ullrich, G. Scrace, P. Stroobant, and M. D. Waterfield. 1984. The *c-sis* gene encodes a precursor of the beta-chain of platelet-derived growth factor. EMBO J. 3:921–928.

Josephs, S. F., C. Guo, L. Ratner, and F. Wong-Staal. 1984. Human proto-oncogene corresponding to the transforming region of simian sarcoma virus. Science 223:487–490.

Julier C. 1984. Cartographie génétique du chromosome 22 humain. Thèse de 3$^{eme}$ cycle, Université de Paris 7.

Julier, C., G. M. Lathrop, J.-M. Lalouel, and J. C. Kaplan. 1985*a*. Use of multilocus test of gene order: example for chromosome 22. Cytogenet. Cell Genet. 40:663A–664A.

Julier, C., G. M. Lathrop, J.-M. Lalouel, A. Reghis, M. F. Szajnert, and J. C. Kaplan. 1985*b*. New restriction fragment length polymorphisms on human chromosome 22 at loci SIS, MB and IGLV. Cytogenet. Cell Genet. 40:664A.

Julier, C., A. Reghis, M. F. Szajnert, J. C. Kaplan, G. M. Lathrop, and J.-M. Lalouel. 1985*c*. A preliminary linkage map of human chromosome 22. Cytogenet. Cell Genet. 40:665A.

Junien, C., M. Vibert, D. Weil, N. Van Cong, and J. C. Kaplan. 1978. Assignment of NADH-cytochrome b5 reductase to human chromosome 22. Ann. Hum. Genet. 42:233–239.

Kaplan, J. C., A. Aurias, C. Julier, M. Prieur, and M. F. Szajnert. 1987. The human chromosome 22. J. Med. Genet. 24:65–78.

Kaplan, J. C., A. Leroux, and P. Beauvais. 1979. Formes cliniques et biologiques du deficit en cytochrome b5 reductase. C. R. Soc. Biol. 173:368–379.

Lathrop, G. M., J.-M. Lalouel, C. Julier, and J. Ott. 1984. Strategies for multilocus linkage analysis in humans. Proc. Natl. Acad. Sci. USA 81:3443–3446.

Leppert, M., W. Cavenee, P. Callahan, T. Holm, P. O'Connell, K. Thompson, G. M. Lathrop, J.-M. Lalouel, and R. White. 1986. A preliminary genetic map of chromosome 13q. Am. J. Hum. Genet. 39:425–437.

Leroux, A., C. Junien, J. C. Kaplan, and J. Bamberger. 1975. Generalized deficiency of cytochrome b5 reductase in congenital methemoglobinemia with mental retardation. Nature **258**:619–620.

McAlpine, P. J., H. Kaita, and M. Lewis. 1978. Is the DIA1 locus linked to the P blood group locus? Cytogenet. Cell Genet. **22**:629–632.

Mendelsohn, M. L., B. H. Mayall, E. Bogart, D. H. Moore, and B. H. Perry. 1973. DNA content and DNA-based centromeric index of the 24 human chromosomes. Science **179**:1126–1129.

Naiki, M., and D. M. Marcus. 1975. An immunochemical study of the human blood group P1, P, Pk glycosphingolipid antigens. Biochemistry **14**:4837–4841.

Naylor, S. L., A. Y. Sakaguchi, D. Barker, R. White, and T. B. Shows. 1984. DNA polymorphic loci mapped to human chromosomes. Proc. Natl. Acad. Sci. USA **81**:2447–2451.

Nielsen, L. S., J. Mohr, and H. Eiberg. 1984. Data concerning the linkage relationship of the HLA and P systems. Cytogenet. Cell Genet. **37**:555A.

Ott, J. 1978. A simple scheme for the analysis of HLA linkages in pedigrees. Ann. Hum. Genet. **42**:255–257.

Reghis, A. 1981. La cytochrome b5 reductase. Thèse de doctorat de sciences medicales, Faculté de Médecine d'Alger (Algeria).

Reghis, A., M. Benabadji, P. Tchen, and J. C. Kaplan. 1981. Quantitative variations of red-cell cytochrome b5 reductase (NADH-methemoglobin-reductase) in the Algerian population. Hum. Genet. **59**:148–153.

Reghis, A., C. Troungos, D. Lostanlen, R. Krishnamoorthy, and J. C. Kaplan. 1983. Characterization of weak alleles at the DIA1 locus (Mustapha-1, Mustapha-2 and Mustapha-3) in the Algerian population. Hum. Genet. **64**:173–175.

Robbins, K. C., S. G. Devare, and S. A. Aaronson. 1981. Molecular cloning of integrated simian sarcoma virus: genome organization of infectious DNA clones. Proc. Natl. Acad. Sci. USA **5**:2918–2922.

Scott, E. M., and R. C. Wright. 1969. The absence of close linkage of methemoglobinemia and other loci. Am. J. Hum. Genet. **21**:194–195.

Taub, R., G. F. Hollis, P. A. Hieter, S. Korsmeyer, T. A. Waldmann, and P. Leder. 1983. Variable amplification of immunoglobulin λ light-chain genes in human populations. Nature **304**:172–174.

Tippett, P., and J. C. Kaplan. 1985. Report of the Committee on the Genetic Constitution of Chromosomes 20, 21, and 22. Cytogenet. Cell Genet. **40**:268–295.

Weller, P., A. J. Jeffreys, V. Wilson, and A. Blanchetot. 1984. Organisation of the human myoglobin gene. EMBO J. **3**:439–446.

White, R., M. Leppert, T. Bishop, D. Barker, J. Berkowicz, C. Brown, P. Callahan, T. Holm, and L. Jerominski. 1985. Construction of linkage maps with DNA markers for human chromosomes. Nature **313**:101–105.