*PROBABLE ASSIGNMENT OF THE DUFFY BLOOD GROUP LOCUS TO CHROMOSOME 1 IN MAN**

BY ROGER P. DONAHUE, WILMA B. BIAS, JAMES H. RENWICK, AND VICTOR A. MCKUSICK

DIVISION OF MEDICAL GENETICS, DEPARTMENT OF MEDICINE, JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE, BALTIMORE, MARYLAND

Communicated by Bentley Glass, September 5, 1968

In this paper (an abstract of which appeared earlier[1]) we report what we believe is the first assignment of a specific gene locus to a specific autosome in man. Estimates are given of recombination values between the common blood group/serum protein loci and an element (*Un*) that controls the expression of an anomalous uncoiled region of chromosome 1 in a newly studied family (V11DE) as well as in two small pedigrees (V11CR and V11PP) reported in the literature.[2, 3] We present evidence that the Duffy blood group locus *Fy*, already believed to be linked to a congenital cataract locus[4] *Cae*, is close to this uncoiler element *Un*, and that all three are probably on chromosome no. 1. A report of another family,[5] slightly favoring linkage between *Fy* and one of the break points of a no. 1 inversion, is also considered.

Methods.—Cytogenetic study: For chromosome analysis of family V11DE, venous blood was drawn with a heparinized syringe and cultured for 72 hr in a mixture of medium 199, fetal bovine serum, and phytohemagglutinin, according to a modification of the technique of Lejeune.[6] In this system, the lymphocytes underwent mitosis and were arrested at metaphase by 0.02 μg colchicine/ml culture fluid added for the final 2 hr of culture. At culture termination, the lymphocytes were treated with 0.9% (hypotonic) sodium citrate, fixed in 3:1 ethanol-acetic acid and then placed on a slide, dried by flaming, and stained in a 4% (v/v) giemsa solution.

Polymorphic marker loci: Loci with two or more alleles were investigated in family V11DE as follows. Blood group typing was done for antigens, A, B, D, C, c, E, e, M, N, S, s, P_1, K, k, Kp^a, Kp^b, Jk^a, Fy^a, Fy^b, Lu^a, Lu^b, Le^a, and Le^b. The saline tube test and antiglobulin test were used where appropriate. The immunoglobulins Gm 1, 2, 3, 5, 6, 13, 14, and Inv 1 were typed by the agglutination inhibition method.[7] Vertical polyacrylamide gel electrophoresis[8] was used to type the serum proteins haptoglobin (Hp), transferrin (Tf), and group-specific component (Gc).

The known alleles at the Duffy blood group locus are the common[9] Fy^a and Fy^b, a relatively rare[10] Fy^x, and a "silent allele"[9] *Fy*, that is largely confined to Negroes. Anti-Fy^a (lot 13) and anti-Fy^b (lot Brent) from Spectra Biologicals were used exclusively. Extensive use of these reagents has indicated that they are monospecific. One drop of the appropriate antiserum and two drops of a washed 2% cell suspension in saline were incubated at 37°C for 30–60 min. The cells were then washed three times in saline and the Coombs test performed. All reactions were either 4+ or negative. Since weak reactions were not observed, it is assumed that the *Fy(b+)* individuals were not Fy^x. If any of the *Fy(a+b−)* individuals in family V11DE are, in fact, *Fy(a+x+)*, it is not apparent from the pedigree.

Results.—The marker chromosome with an anomalous region: The marker chromosome occurring in family V11DE is shown in Figures 1 and 2. This chromosome differs from the usual no. 1 by having a greater over-all length and a greater length of the long arm. A region extends from the centromere along part of each long arm in which the chromatids may be thinner with, occasionally, an alternating pattern of light and dark staining, suggestive of coils. Rather than

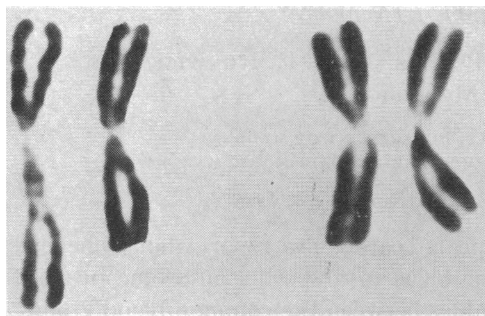

FIG. 1.—Two pairs of no. 1 chromosomes from metaphase lymphocytes of family V11DE. *Left:* from the proband's uncle who is heterozygous for the variant chromosome, shown as the leftmost member of the pair. *Right:* from the proband's sister who lacks the variant.

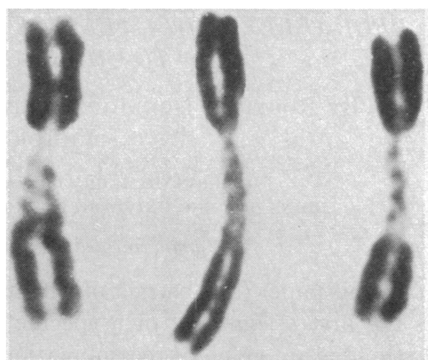

FIG. 2.—Variant no. 1 chromosomes from three metaphase cells of the proband of family V11DE. The relatively uncoiled chromatids in the paracentric region can be seen. This distinctive region is not usually so clearly visible.

the increase in length being by insertion of extra chromosomal material,[11] it probably arises directly from an alteration in the coiling structure. The agent that we call the uncoiler (*Un*) element controlling the morphology of this region could be a single locus, or small inversion, or some other, as yet unknown, factor. Since the unusual heterozygous genotype at the *Un* region affects only one of the two no. 1 homologues, we think it likely that *Un* is situated somewhere on chromosome 1 itself.

Family data: Four families (Fig. 3) with a marker chromosome 1 are used, three of which have been previously reported. Of these three, families V11CR[2] and V11PP[3] appear to possess the same marker as in our pedigree, while in family V12LE[5] the marker probably reflects a pericentric inversion.

The largest family, V11DE, is shown in Figure 3. In this Caucasian family the proband is clinically normal and the variant chromosome is not associated with an abnormal phenotype. Of 18 individuals in three generations, 10 are heterozygous for the variant and a further one is presumed to be heterozygous, since his son carries the anomalous chromosome. Segregation thus appears to be Mendelian. The full marker data are recorded in Appendix I.[12]

Linkage analysis: The genetic map distance between any two loci on a chromosome is estimated from the observed proportion of recombinants. The true recombination fraction θ can take any value between 0 (complete linkage) and 0.5 (no linkage, indicating free recombination between the two loci, whether they are far apart on the same chromosome or are on separate chromosomes). To determine whether or not any one of the polymorphic blood group and blood protein loci is linked to the uncoiler element (*Un*) or an inversion break point (*Invn*), the odds for linkage have been calculated using all four families. The ratio of the likelihood of obtaining these pedigrees if there is linkage at a particular value (θ) compared to the corresponding likelihood if there is no linkage ($\theta = {}^1/_2$), is the standardized likelihood ratio. These odds were calculated for

MARKER CHROMOSOME WITH ANOMALOUS REGION

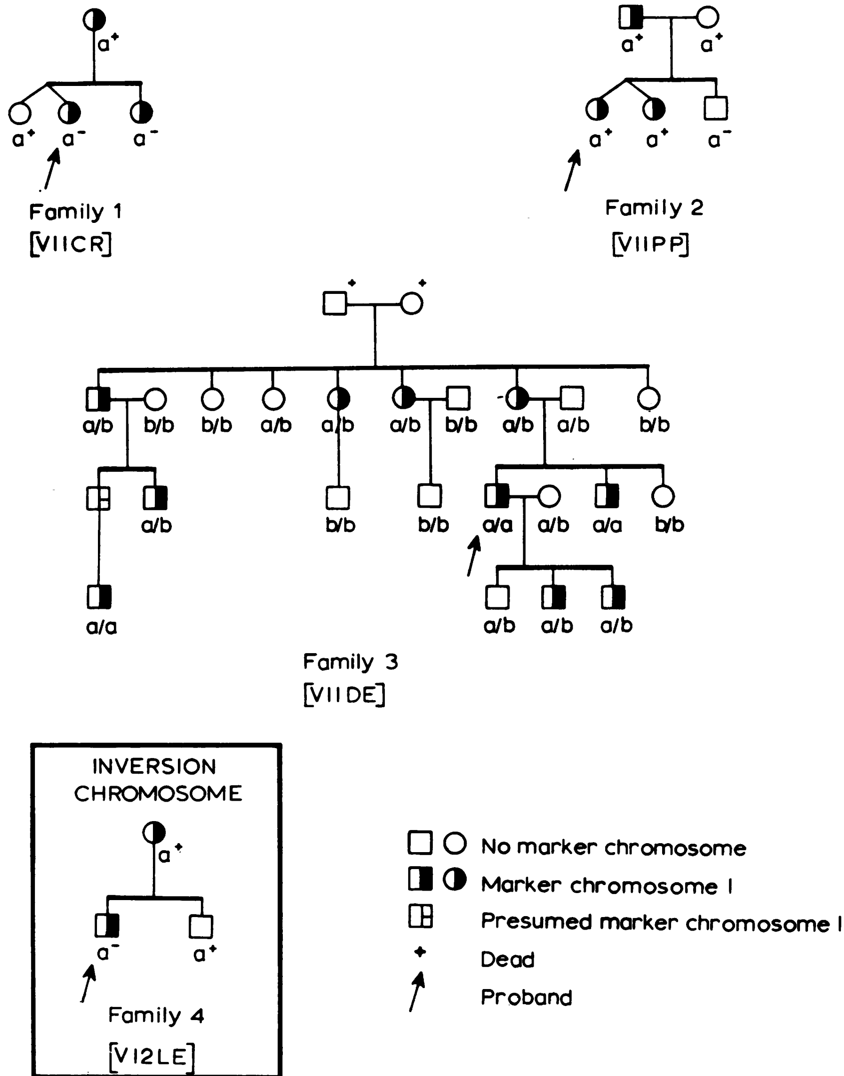

FIG. 3.—The four families used to test for linkage. In our family V11DE, genotypes at the Duffy blood group locus are shown with *a* and *b* representing, respectively, the Fy^a and Fy^b alleles. In families V11CR,[2] V11PP,[3] and V12LE,[5] Duffy phenotypes are indicated for the presence ($a+$) or absence ($a-$) of the Fy^a allele. Those individuals indicated as deceased were not tested.

each pedigree, with computer assistance, over θ values between 0 and 0.5. Each of these likelihoods (odds), when combined with an appropriate prior probability, is a measure of the final probability that the corresponding θ value is the true

value. For convenience, the logs of the standardized odds are sometimes used and these are known as "lods."

Linkage was sought between the marker loci and either *Un* or *Invn* on chromosome 1. With the exception of that for *Fy*:*Un*, none of the sets of likelihoods suggests linkage. Families V11CR and V11PP (Fig. 3) indicate a consistent tendency for the Fy^a allele to be transmitted with the marker chromosomal phenotype and in families V11DE and V12LE, the Fy^b allele is apparently so transmitted. The weight of this evidence for linkage can be assessed most readily in terms of the confidence with which it allows us to assign the *Fy* locus to chromosome 1. The Bayesian procedure used is more fully discussed elsewhere.[13]

The usual approximation that all values of the recombination fraction θ are equally likely for a linkage between two random sites on a chromosome is not valid for a long chromosome in man. To find suitable prior probabilities applicable on chromosome 1, it is easier to deal with map intervals directly, using a formula such as that of Carter and Falconer[14] to translate recombination fractions to map intervals. The safest prior probabilities for such map intervals would be those appropriate to an assumption of ignorance about the positions of the *Un* element and *Fy* on the anomalous chromosome, and they would follow the usual triangular distribution $2\ (l - w)/lL$.[15] This is maximal for short distances and minimal for long ones, l and L being the map lengths of chromosome 1 and of the whole autosomal complement, respectively. The prior probability density of a "map interval" in excess of the length of chromosome 1 (i.e., when *Fy* is on another chromosome) is uniformly $1/L$, the sum of such probabilities being $1 - (l/L)$ or about 10/11 in the present context. Chromosome 1 is here crudely treated as being 3 Morgans (M) (300 units) long out of a total autosomal map (averaged for males and females) of 33 M, using indirect evidence as discussed by Renwick and Schulze.[16] For all such "map intervals" corresponding to nonlinkage, the recombination fraction θ is expected to be one-half, so the effect of the standardization of the likelihood ratios is to leave the prior chances for nonlinkage unmodified by the data. The standardized likelihood ratios for various lengths of the *Fy*:*Un* map interval are then used to modify (by multiplication) the corresponding prior probabilities. The data-modified values are given in Table 1 and are graphed in Figure 4. The factor r merely allows the total probability of all hypotheses to remain 1. By planimetry, the area under the full curve of Figure 4 is found to be a factor 93 times larger than the area of the triangle of the original prior distribution, calculated on the same scale. The odds that *Fy* is on chromosome 1 have therefore been increased from 1:10 to 93:10 by the *Un* data.

If *Fy* and *Un* are on chromosome 1, the best estimate for the distance between them is 2.5 map units (0.025 M, see Fig. 4). The limits of the 95 per cent probability region are 0 and 21 map units.

Slight support comes from the inversion pedigree (family V12LE[5]) with its weak hint of a linkage between *Fy* and one of the break points (*Invn*) of the inversion. The data on the *Fy*:*Invn* linkage relationship increase the odds favoring the assignation of *Fy* to chromosome 1 to 100:10, equivalent to a direct probability of 0.91.

Table 1. *Derivation of the final odds for assigning the Fy locus to chromosome 1.*

(1) Map interval w (Morgans)	(2) Recombination fraction (θ)	(3) Standardized likelihoods V11DE, V11CR, and V11PP	(4) Points on prior probability curve $r(6\text{-}2w)$	(5) Points on final probability curve (see Fig. 4) (5) = (3) × (4)
0	0	318	6.0r*	1908r
0.05	0.05	1030	5.9r	6077r
0.1	0.1	692	5.8r	4014r
0.2	0.2	158	5.6r	865r
0.31	0.3	24	5.4r	129r
0.44	0.4	3	5.1r	15r
0.95	0.4975	1	4.1r	4r
3.00	0.5	1	0 r	0r

	Fy-off-1	:	*Fy*-on-1
Observational odds (relative area under prior and final probability curves)	1	:	93
Prior odds	10	:	1
Final odds	10	:	93

* r is a simple scaling factor that allows the prior probabilities to sum to 1.

Discussion.—The anomalous (relatively uncoiled) region of chromosome 1 may be an exaggerated version of the usual secondary constriction[17] whose frequency of occurrence within the cells of an individual can be modified by altering cell culture or cell-fixing conditions.[18] The variant chromosome, however, is transmitted in simple Mendelian fashion and is present in all of the tested cells of the carriers. This chromosome appears to be the same as those previously found segregating in families ascertained through abnormal individuals[2, 3, 11, 19] or in a single abnormal patient.[20] The lack of association between this chromosome and any clinical abnormality in our family, as well as the absence of convincing

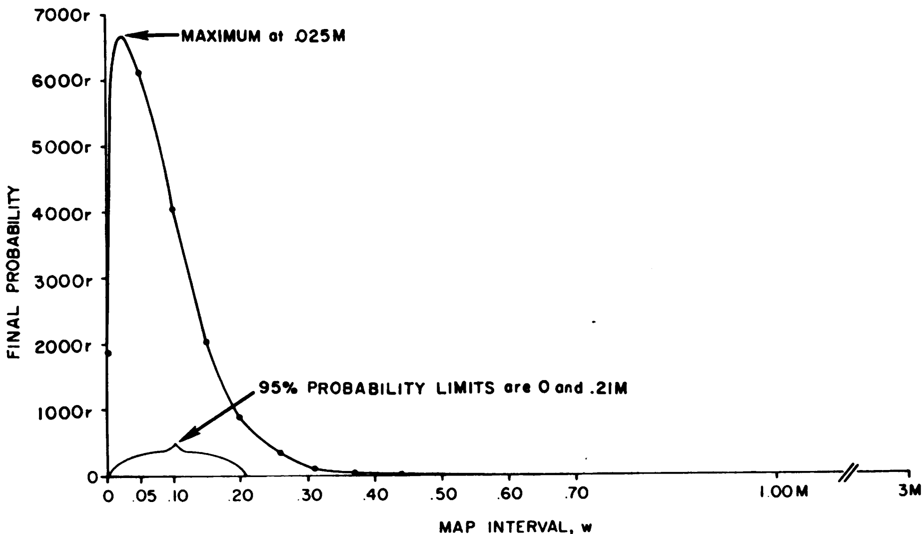

Fig. 4.—The final probability of linkage between the Duffy blood group locus and the uncoiler element. The map distance w is measured in Morgans (M).

evidence of a pathologic association in the other families, leads us to conclude that this chromosome should be labeled a variant instead of an abnormal no. 1. It has been estimated to occur in three individuals per thousand.[21]

Unexpected inheritance with reference to the Duffy system in a child with a deletion of two-thirds of the long arm of chromosome 16 led some previous workers to the suggestion that *Fy* is located on chromosome 16.[22] However, since there is evidence that alleles exist at this locus that produce antigens for which there are as yet no specific antisera,[9] an alternative explanation of the *Fy* inheritance, unrelated to mapping, is possible. In fact, since the frequencies of such alleles may be taken to be about 0.03 in the non-Negro, the chance that such an allele is present in the father exceeds the chance that *Fy* happens to be on that 1/80 part of the autosomal complement. Nevertheless, provided that the further, contramarital possibility is ignored, the odds on an assignment of *Fy* to chromosome 16 have been raised by a factor of 9 (i.e., $b/2c$ where b,c are the frequencies of Fy^b and *Fy*, respectively). If these increased odds are used, the probability that *Fy* is on chromosome 1 falls slightly from 0.91 to 0.90.

The fact that *Fy* is known not to be on certain sections (1.2%) of the autosomal complement, namely the short arms of chromosomes 13 and 18,[8, 23] gives a negligible improvement to these odds on localization. We are aware of another family[24] in which the variant 1 with the anomalous region is segregating and which supports evidence for the presence of *Fy* on chromosome 1.

If the *Fy* locus *is* on chromosome 1, then the locus *Cae* for a certain form of congenital cataract has a probability of 0.96 of also being on chromosome 1.[4] All three loci are probably close to each other, the maximum likelihood map estimates being 0 units for *Fy*:*Cae* and 2.5 units for *Fy*:*Un*. The upper limits of the 95 per cent probability intervals are 20 and 21 units, respectively, given that the loci really are on the same chromosome. We cannot yet state the linear order of these loci nor their positions on the chromosome.

Summary.—Evidence for linkage between the Duffy blood group locus *Fy* and an element *Un*, postulated as controlling the uncoiling of the paracentric region on the long arm of chromosome 1, is described. When taken in conjunction with that given in the published reports of three small pedigrees, the evidence leads to a direct probability of 0.90 that *Fy* is on chromosome 1. The congenital cataract locus *Cae*, to which *Fy* is believed to be closely linked (probability 0.96), is presumably also on chromosome 1. The *Fy* locus is estimated to be about 2.5 map units from the element controlling the uncoiling. The probability takes into account prior odds against such an assignment, including the weak indications that *Fy* may be elsewhere (on chromosome 16). The present odds that *Fy* is on chromosome 1, chromosome 16, or another chromosome are judged to be 90:1:9, respectively.

We are indebted to Dr. R. N. Schimke for assistance in obtaining blood specimens; to Mrs. R. Classon and Miss H. Herr for technical assistance; to Mrs. C. Chambers, L. Maulding, and M. Campbell for computing assistance; and to Mrs. A. Buchan for clerical assistance.

* Aided by a grant from the Hartford Foundation and in part by grants GM 10189 from the National Institutes of Health (U.S.) and G960/109B from the Medical Research Council (U.K.). The calculations were carried out in the Computing Center of the Johns Hopkins Medical Institutions, which is supported by grant FR-00004 from the National Institutes of Health and by educational contributions from the International Business Machines Corporation.

[1] Donahue, R. P., J. H. Renwick, L. de los Cobos, D. S. Borgaonkar, W. B. Bias, and V. A. McKusick, *Clin. Res.*, **16**, 296 (1968).

[2] Cooper, H. L., and R. Hernits, *Am. J. Human Genet.*, **15**, 465 (1963).

[3] Philip, J., O. Frydenberg, and V. Sele, *Cytogenetics*, **4**, 329 (1965).

[4] Renwick, J. H., and S. D. Lawler, *Ann. Human Genet.*, **27**, 67 (1963).

[5] Lele, K. P., T. Dent, and J. D. A. Delhanty, *Lancet*, **1**, 576 (1965).

[6] Lejeune, J., in *Les Chromosomes Humaines*, ed. R. Turpin and J. Lejeune (Paris: Gauthier-Villars, 1965), p. 26.

[7] Steinberg, A. G., and J. A. Wilson, *Am. J. Human Genet.*, **15**, 96 (1963).

[8] Bias, W. B., and B. R. Migeon, *Am. J. Human Genet.*, **19**, 393 (1967).

[9] Race, R. R., and R. Sanger, *Blood Groups in Man* (Philadelphia: F. A. Davis, 1962), 4th ed.

[10] Chown, B., M. Lewis, and H. Kaita, *Am. J. Human Genet.*, **17**, 384 (1965).

[11] Patau, K., E. Therman, S. L. Inhorn, D. W. Smith, and A. L. Ruess, *Chromosoma*, **12**, 573 (1961).

[12] Supplementary material (Appendix I) has been deposited with the National Auxiliary Publications Service of the American Society for Information Science, c/o CCM Information Sciences, Inc., 22 West 34th St., New York, N. Y. 10001, from which it may be obtained by ordering NAPS Document 00126 and by remitting $3.00 for photocopy or $1.00 for microfiche. Advance payment is required. Make checks payable to: ASIS-NAPS.

[13] Renwick, J. H., *Brit. Med. Bull.*, **25**, in press.

[14] Carter, T. C., and D. S. Falconer, *J. Genet.*, **50**, 307 (1951).

[15] Bell, J., and J. B. S. Haldane, *Proc. Roy. Soc.* (London), **123B**, 119 (1937).

[16] Renwick, J. H., and J. Schulze, *Am. J. Human Genet.*, **16**, 410 (1964).

[17] Ferguson-Smith, M. A., M. E. Ferguson-Smith, P. M. Ellis, and M. Dickson, *Cytogenetics*, **1**, 325 (1962).

[18] Saksela, E., and P. S. Moorhead, *Cytogenetics*, **1**, 225 (1962); Kaback, M. M., E. Saksela, and W. J. Mellman, *Exptl. Cell Res.*, **34**, 182 (1964).

[19] Yunis, J. J., and R. J. Gorlin, *Chromosoma*, **14**, 146 (1963); Koulischer, L., and J. Zanen, *Bull. Soc. Belge Ophthalmol.*, **136**, 602 (1964); Sofuni, T., and A. A. Sandberg, *Cytogenetics*, **6**, 357 (1967).

[20] Genest, P., M. A. Mortezai, and M. Tremblay, *Arch. Franc. Pediat.*, **23**, 887 (1966).

[21] Lubs, H. A., personal communication.

[22] Crawford, M. N., H. H. Punnett, and G. G. Carpenter, *Nature*, **215**, 1075 (1967).

[23] Uchida, I. A., K. N. McRae, H. C. Wang, and M. Ray, *Am. J. Human Genet.*, **17**, 410 (1965).

[24] Ying, K. L., and E. J. Ives, personal communication.