

## SHORT COMMUNICATION

# Localization of *CYP2F1* by Multipoint Linkage Analysis and Pulsed-Field Gel Electrophoresis

ALLEN E. BALE,\* ANNA L. MITCHELL,\* FRANK J. GONZALEZ,† AND O. WESLEY McBRIDE\*

\*Laboratory of Biochemistry and †Laboratory of Molecular Carcinogenesis, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892

Received August 27, 1990

*CYP2F1*, which encodes a P450 enzyme capable of metabolizing several mono-oxygenase substrates with the highest activity toward ethoxycoumarin, has been mapped to human chromosome 19 by somatic cell hybrid studies. The *CYP2A* and *CYP2B* subfamilies are known to lie within 350 kb of each other on chromosome 19. To determine the locations of *CYP2F1* with respect to *CYP2A* and *CYP2B*, multipoint linkage analysis and pulsed-field gel electrophoresis were performed. No recombinants were found between *CYP2F1* and *CYP2B* in more than 50 meioses, and one recombinant was found between *CYP2F1* and *CYP2A* (50 meioses). Pulsed-field gel electrophoresis showed the three loci to lie within a 240-kb region. Multipoint analysis of the haplotyped *CYP* cluster with four other chromosome 19 loci yielded the order *D19S7-D19S9-APOC2-D19S8-CYP*, although four other orders could not be excluded. The odds were  $4.4 \times 10^3$  against the order proposed by the HGM10 consortium, *D19S7-D19S9-CYP-D19S8-APOC2*. © 1991 Academic Press, Inc.

Cytochrome P450s are enzymes involved in the oxidative metabolism of a variety of endogenous and exogenous lipophilic substances. Genes encoding the P450 enzymes have been categorized into families and subfamilies on the basis of global protein sequence similarities (Nebert *et al.*, 1989). Members of different families display less than 37% sequence similarity, and members of different subfamilies are not more than 65% similar. Enzymes within a single subfamily can have distinct activities, and the genes encoding them map close together on the same chromosome. In the human *CYP2* family the *CYP2D* subfamily maps to chromosome 22 (Gonzalez, 1989). Both the *CYP2C* and the *CYP2E* subfamilies map to chromosome 10, but because they map to two different chromosomes in rodents, they are probably located in different regions of human chromosome 10

(Gonzalez, 1989). In contrast, the *CYP2A* and *CYP2B* subfamilies map within 350 kb of each other on chromosome 19 (Miles *et al.*, 1989).

A recently described P450 gene subfamily, designated *CYP2F*, has been localized to human chromosome 19 and mouse chromosome 7 by somatic cell hybrid studies (Nhamburo *et al.*, 1990). The one member of this subfamily that has been characterized, *CYP2F1*, encodes a P450 capable of metabolizing several common mono-oxygenase substrates with the highest activity toward ethoxycoumarin. There is evidence for a second highly similar gene adjacent to *CYP2F1*. The purpose of this study was to localize the *CYP2F* subfamily with respect to *CYP2A* and *CYP2B* and to determine the position of *CYP2F* among other known markers on chromosome 19.

All linkage studies involved the 40 CEPH families. RFLP typing of *CYP2A*, *CYP2B*, and *CYP2F* was performed by standard Southern blot analysis (Reed and Mann, 1985) under high stringency conditions. The probe-enzyme combinations used to detect RFLPs were full-length *CYP2A3* (Yamano *et al.*, 1989) using *SstI* (Wainwright *et al.*, 1985), the coding region of *CYP2B7* cDNA (Yamano *et al.*, 1989) using *BamHI* (Miles *et al.*, 1988), and full-length *CYP2F1* cDNA using *XbaI* (Nhamburo *et al.*, 1990). Results for other chromosome 19 RFLPs were obtained from the CEPH database version 3. Errors in *D19S8* data (incorrect typing of 1416-11) and *D19S9* data (reversal of typing for 1333-11 and 1333-12) were corrected in cooperation with Collaborative Research, Inc. (Weisenbach, personal communication). Pairwise linkage analysis was performed with the MLINK option of the LINKAGE package (Lathrop *et al.*, 1985), and CILINK was used for five-point analysis. For localization of the *CYP* cluster with respect to other chromosome 19 markers, haplotypes were constructed from alleles at the three *CYP* loci. The single

**TABLE 1**  
**Two-Point Lod Scores for *CYP2A* vs *CYP2B* vs *CYP2F***

Locus pair	Recombination fraction					$Z_{\max}$	$\theta_{\max}$
	0.0	0.1	0.2	0.3	0.4		
<i>CYP2A</i> - <i>CYP2B</i>	15.12	11.98	8.66	5.23	2.04	15.12	0.0
<i>CYP2A</i> - <i>CYP2F</i>	$-\infty$	7.24	5.41	3.36	1.32	8.15	0.02
<i>CYP2B</i> - <i>CYP2F</i>	18.50	15.43	12.00	8.14	3.85	18.50	0.0

**TABLE 2**  
**Two-Point Lod Scores for Haplotyped *CYP2A*, *2B*, and *2F* Data vs Four Chromosome 19 Markers**

Locus	Recombination fraction					$Z_{\max}$	$\theta_m, \theta_f$
	0.0	0.1	0.2	0.3	0.4		
<i>APOC2</i>	$-\infty$	10.76	8.96	6.25	3.17	11.85	0.15, 0.00
<i>D19S7</i>	$-\infty$	10.60	10.47	7.74	3.81	11.51	0.09, 0.23
<i>D19S8</i>	14.68	12.35	9.74	6.78	3.41	14.68	0.00, 0.00
<i>D19S9</i>	$-\infty$	4.94	4.41	3.17	1.55	5.15	0.06, 0.15

individual in this study showing recombination within this cluster was excluded from the multipoint analysis.

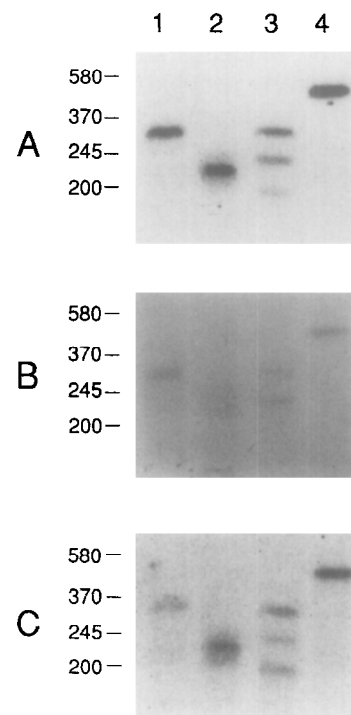
DNA samples for pulsed-field gel electrophoresis (PFGE) were prepared from peripheral blood leukocytes by standard techniques (Sambrook *et al.*, 1989), digested with *NotI*, *MluI*, *SalI*, and *KspI*, and electrophoresed with parameters set for separation of fragments in the range 200 kb to 2 Mb according to the manufacturer's instructions (Bio-Rad).

Pairwise analysis of *CYP2A*, *CYP2B*, and *CYP2F* (Table 1) revealed only a single recombinant between *CYP2A* and *CYP2F* in approximately 50 informative meioses. For *CYP2A* vs *CYP2B* and *CYP2B* vs *CYP2F* there were no recombinants in more than 50 meioses. The *CYP* cluster showed no recombination with *D19S8* after correction of erroneous typing in the CEPH database (Table 2). Significantly positive lod scores indicated linkage between *CYP* and *D19S9*, *D19S7*, and *APOC2*, although all showed some recombination. The most likely order of these loci was

*D19S7*-*D19S9*-*APOC2*-*D19S8*-*CYP* (Fig. 1), but four other orders could not be excluded with odds greater than 100:1 (Table 3). The odds were  $4.4 \times 10^3$  against the order proposed by the HGM10 consortium (LeBeau *et al.*, 1989), namely, *S7*-*S9*-*CYP*-*S8*-*APOC2*.

Odds	1119:1	11.5:1	6.4:1	1.5:1
M	0.00	0.01	0.06	0.00
	—D19S7—	D19S9—	APOC2—	D19S8—CYP—
F	0.23	0.01	0.08	0.00
SA	0.07	0.01	0.07	0.00

**FIG. 1.** Most likely map generated by multipoint analysis of haplotyped *CYP* data and four chromosome 19 markers. M, F, and SA = male, female, and sex-averaged recombination distances, respectively. Odds = relative likelihood of inverting adjacent pairs of loci.



**FIG. 2.** Pulsed-field gel analysis of *CYP2A* (A), *CYP2B* (B), and *CYP2F* (C). Lane 1 represents *KspI* digests; 2, *SalI* digests; 3, *MluI* digests; and 4, *NotI* digests.

TABLE 3

## Odds of Five-Point Orders for Chromosome 19 Loci\*

Order	Odds
D19S7-D19S9-APOC2-D19S8-CYP	1:1
D19S7-D19S9-APOC2-CYP-D19S8	1.5:1
D19S7-D19S9-D19S8-APOC2-CYP	6.4:1
D19S7-APOC2-D19S9-D19S8-CYP	11.5:1
D19S7-APOC2-D19S9-CYP-D19S8	49:1

\* Odds against all other possible orders were greater than 100:1.

To determine the physical distance between the three *CYP* loci, probes for each of the three were sequentially hybridized to the same Southern blot prepared from a PFGE gel (Fig. 2). All three probes hybridized to identical *NotI* (580 kb), *SalI* (245 kb), and *KspI* (370 kb) bands. *MluI* produced a partial digest with bands at 370, 245, and 200 kb hybridizing to *CYP2A* and *CYP2F* and only the former two bands hybridizing to *CYP2B*. These data indicate that members of the *CYP2A*, *CYP2B*, and *CYP2F* subfamilies lie within 245 kb of each other on chromosome 19. *CYP2B* is separated from *CYP2A* and *CYP2F* by an *MluI* site, but the exact order and distance between these loci cannot be established with certainty. Observation of a single recombinant between *CYP2B* and *CYP2F* is not inconsistent with the maximum physical distance between these loci as estimated from PFGE.

The extreme proximity of three P450 subfamilies is somewhat surprising in view of their different functions and limited sequence similarities (Yamano *et al.*, 1990). This study provides strong evidence that this gene cluster developed by tandem duplication.

Our multipoint mapping results show a marked discrepancy with the physical and linkage consortium maps produced at the 10th International Workshop on Human Gene Mapping (LeBeau *et al.*, 1989). Our proposed order of loci is identical to that previously reported by Mitchell *et al.* (1989). However, the entire region now spans only 19 cM sex averaged—7 cM in males and 38 cM in females—whereas this region previously encompassed 38 cM sex averaged. This difference can be attributed entirely to our previous estimate of a 21-cM distance between *CYP2A*, *CYP2B*, *CYP2F*, and any other member of the cluster of loci, and it was created by mistyping of three individuals from two kindreds. This graphically illustrates the dramatic influence that a few errors in the CEPH database can exert toward greatly expanding the apparent size of the human genetic map. The discrepancy between the HGM10 physical and linkage maps is difficult to explain and suggests either additional RFLP typing errors in the CEPH database version 3 producing an erroneous linkage map or errors

in analysis of somatic cell hybrids leading to an erroneous physical map. With additional physical mapping of chromosome 19, an unambiguous answer will eventually be found.

## REFERENCES

- GONZALEZ, F. J. (1989). The molecular biology of cytochrome P450's. *Pharmacol. Rev.* **40**: 243-288.
- LATHROP, G. M., LALOUEL, J.-M., JULIER, C., AND OTT, J. (1985). Multilocus linkage analysis in humans: Detection of linkage and estimation of recombination. *Am. J. Hum. Genet.* **37**: 482-498.
- LE BEAU, M. M., RYAN, D., AND PERICAK-VANCE, M. A. (1989). Report of the committee on the genetic constitution of chromosomes 18 and 19. *Cytogenet. Cell Genet.* **51**: 338-357.
- MILES, J. S., BICKMORE, W., BROOK, J. D., McLAREN, A. W., MOEHAN, R., AND WOLF, C. R. (1989). Close linkage of the human cytochrome P450IIA and P450IIB gene subfamilies: Implications for the assignment of substrate specificity. *Nucleic Acids Res.* **17**: 2907-2917.
- MITCHELL, A. L., BALE, A. E., GONZALEZ, F., AND MCBRIDE, O. W. (1989). Mapping of cytochrome p450 IIA, IIB, and IIF subfamilies on chromosome 19 by linkage analysis. *Cytogenet. Cell Genet.* **51**: 1045.
- NEBERT, D. W., NELSON, D. R., ADESINK, M., COON, M. J., ESTABROOK, R. W., GONZALEZ, F. J., GUENGERICH, F. P., GUNSALUS, I. C., JOHNSON, E. F., KEMPER, B., LEVIN, W., PHILLIPS, I. R., SATO, R., AND WATERMAN, M. R. (1989). The P450 superfamily: Updated listing of all genes and recommended nomenclature for the chromosomal loci. *DNA* **8**: 1-13.
- NHAMBURU, P. T., KIMURA, S., MCBRIDE, O. W., KOZAK, C. A., GELBOIN, H. V., AND GONZALEZ, F. J. (1990). The human 2F subfamily: Identification of a cDNA encoding a new cytochrome P450, cDNA-directed expression, and chromosome mapping. *Biochemistry* **29**: 5491-5499.
- REED, K. C., AND MANN, D. A. (1985). Rapid transfer of DNA from agarose gels to nylon membranes. *Nucleic Acids Res.* **13**: 7207-7221.
- SAMBROOK, J., FRITSCH, E. F., AND MANIATIS, T. (1989). "Molecular Cloning: A Laboratory Manual," pp 6.53-6.59, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- SANTISTEBAN, I., POVEY, S., SHEPARD, E. A., AND PHILLIPS, I. R. (1988). The major phenobarbital-inducible cytochrome P-450 gene subfamily (P450IIB) mapped to the long arm of human chromosome 19. *Ann. Hum. Genet.* **52**: 129-135.
- WAINWRIGHT, B. J., WATSON, E. K., SHEPARD, E. A., AND PHILLIPS, I. R. (1985). RFLP for a human cytochrome P-450 gene at 19q13.1-qter (HGM8 provisional designation I). *Nucleic Acids Res.* **13**: 4610.
- YAMANO, S., TATSUNO, J., AND GONZALEZ, F. J. (1990). The 2A3 gene product catalyzes coumarin 7-hydroxylation in human liver. *Biochemistry* **29**: 1322-1329.
- YAMANO, S., NHAMBURU, P. T., AOYAMA, T., MEYER, U. A., INABA, T., KALOW, W., GELBOIN, H. V., MCBRIDE, O. W., AND GONZALEZ, F. J. (1989). cDNA cloning and sequence and cDNA-directed expression of human P450 IIB1: Identification of a normal and two variant cDNAs derived from the 2B locus on chromosome 19 and differential expression of the IIB and mRNAs in human liver. *Biochemistry* **28**: 7340-7348.