

Genetic and Physical Mapping of the Treacher Collins Syndrome Locus with Respect to Loci in the Chromosome 5q3 Region

ETHYLIN WANG JABS,*¹ XIANG LI,* MICHAEL LOVETT,† LARRY H. YAMAOKA,‡ EUGENE TAYLOR,§
 MARCY C. SPEER,‡ CATHLEEN COSS,* RONALD CADLE,¶ BRYAN HALL,¶ KAREN BROWN,|| KENNETH K. KIDD,⊥
 GREGORY DOLGANOV,** MIHAEL H. POLYMERPOULOS,†† AND DEBORAH A. MEYERS§

*Departments of Pediatrics and Medicine, Center for Medical Genetics, and §Departments of Medicine and Epidemiology, The Johns Hopkins School of Medicine, Baltimore, Maryland; †Department of Biochemistry and McDermott Center, University of Texas Southwestern Medical Center, San Antonio, Texas; ‡Division of Neurology, Duke University Medical Center, Durham, North Carolina; ¶Department of Pediatrics, University of Kentucky, Lexington, Kentucky; ||Division of Reproductive Genetics, Albert Einstein College of Medicine and Montefiore Medical Center, Bronx, New York; ⊥Department of Genetics, The Yale University School of Medicine, New Haven, Connecticut; **Genelabs Inc., 505 Penobscot Drive, Redwood City, California; and ††NIMH Neuroscience Center at St. Elizabeth's, Laboratory of Biochemical Genetics, Washington, DC.

Received April 16, 1993; revised June 25, 1993

Treacher Collins syndrome is an autosomal dominant, craniofacial developmental disorder, and its locus (TCOF1) has been mapped to chromosome 5q3. To refine the location of the gene within this region, linkage analysis was performed among the TCOF1 locus and 12 loci (IL9, FGFA, GRL, D5S207, D5S210, D5S376, CSF1R, SPARC, D5S119, D5S209, D5S527, FGFR4) in 13 Treacher Collins syndrome families. The highest maximum lod score was obtained between loci TCOF1 and D5S210 ($Z = 10.52$; $\hat{\theta} = 0.02 \pm 0.07$). The best order, IL9-GRL-D5S207/D5S210-CSF1R-SPARC-D5S119, and genetic distances among these loci were determined in the 40 CEPH families by multipoint linkage analysis. YAC clones were used to establish the order of loci, centromere-5'GRL3'-D5S207-D5S210-D5S376-CSF1R-SPARC-D5S119-telomere. By combining known physical mapping data with ours, the order of chromosome 5q3 markers is centromere-IL9-FGFA-5'GRL3'-D5S207-D5S210-D5S376-CSF1R-SPARC-D5S119-D5S209-FGFR4-telomere. Based on this order, haplotype analysis suggests that the TCOF1 locus resides distal of CSF1R and proximal to SPARC within a region less than 1 Mb in size. © 1993 Academic Press, Inc.

arches (reviewed by Gorlin *et al.*, 1990). The major diagnostic criteria include microtia, hearing loss, midface hypoplasia, downward slant of palpebral fissures, colobomata of lower lids, and micrognathia. Treatment often involves audiologic, respiratory, and surgical intervention. Treacher Collins syndrome is the most common of the mandibulofacial dysostosis conditions and is estimated to occur in 1/25,000 to 50,000 live births. This incidence may be underestimated because, although Treacher Collins syndrome is a fully penetrant disorder, it exhibits variable expressivity, and mildly affected individuals may remain undiagnosed if they are not carefully examined. Isolation of the gene would allow for accurate post- and prenatal diagnostic testing.

The gene has been mapped to the long arm of chromosome 5 (Dixon *et al.*, 1991) and sublocalized using highly informative markers or short tandem repeat sequences to 5q31.3 → 33.3 (Jabs *et al.*, 1991a) or 5q32 → 33.2 (Dixon *et al.*, 1992). Using four markers, Dixon *et al.* (1992) presented evidence that the TCOF1 locus is flanked by D5S376 and D5S527. We present mapping data to refine the location of the Treacher Collins syndrome in relationship to loci IL9, FGFA, GRL, D5S207, D5S210, ADRB2, D5S376, CSF1R, SPARC, D5S119, D5S209, D5S527, and FGFR4 in the chromosome 5q3 region.

INTRODUCTION

Treacher Collins syndrome is an autosomal dominant, craniofacial developmental disorder affecting structures derived from the first and second branchial

MATERIALS AND METHODS

Recombinant and genomic DNAs and Southern blot hybridization. Probes were kindly provided by R. Evans of the Salk Institute for Biological Studies for locus GRL (probe OB7; Hollenberg *et al.*, 1985), G. Long of the University of Vermont for locus SPARC (probe pHVON-9-2; Villarrael *et al.*, 1989), and K. Alitalo of the University of Helsinki, Finland, for locus FGFR4 (probe HE6-2; Armstrong *et al.*, 1991). Thirteen Treacher Collins syndrome families with a total of 108 individuals and 69 meiotic events were ascertained. Thirty-eight meiotic events arose from females and 31 from males. The members of

¹ To whom correspondence should be addressed at Departments of Pediatrics and Medicine, Center for Medical Genetics, The Johns Hopkins School of Medicine, CMSC 10-04, 600 North Wolfe Street, Baltimore, MD 21287-3914. Telephone: (410) 955-4160. Fax: (410) 955-0484.

the family were examined by medical geneticists (E. W. Jabs and B. Hall). If they did not have any of the major diagnostic features of Treacher Collins syndrome (downward slant of the palepebral fissures, lower eyelid coloboma, malformed ears, and hypoplastic zygomatic arches and mandible) by clinical and radiologic examinations, they were considered unaffected. Their genomic DNAs were isolated, digested, electrophoresed, transferred, and hybridized as described in Jabs *et al.* (1991b). Genomic DNAs from 40 normal families were obtained from the Centre d'Etude du Polymorphisme Humain (CEPH).

Detection of short tandem sequence repeat polymorphisms. Short tandem repeat (STR) or dinucleotide repeat ((CA)_n) polymorphisms were detected using the polymerase chain reaction (PCR). An oligonucleotide primer of a pair was end-labeled at the 5' end with [γ -³²P]dATP (Sambrook *et al.*, 1989) and utilized in PCRs according to conditions previously published for IL9 (Polymeropoulos *et al.*, 1991a); CSF1R (Polymeropoulos *et al.*, 1991b); FGFA (Li *et al.*, 1992); D5S207, D5S210, D5S119, and D5S209 (Weber *et al.*, 1991); and D5S376 and D5S527 (Dixon *et al.*, 1992). PCR products were separated by electrophoresis through 6% polyacrylamide DNA sequencing gels (Biggin *et al.*, 1983).

Order by linkage analysis. Restriction fragment length polymorphisms (RFLPs) and STR-polymorphisms were scored from autoradiographs and data were entered into an ASCII file in linkage format on an IBM 486. Two-point and multipoint linkage analyses were performed using data from the Treacher Collins syndrome families and the computer programs LIPED (Ott, 1974) and LINKAGE (Lathrop *et al.*, 1984). It was assumed that Treacher Collins syndrome is an autosomal dominant condition with a 0.0001 gene frequency and 99% penetrance. Lod scores or odds ratios for linkage versus nonlinkage for sex-average and sex-specific recombination fractions were generated between the Treacher Collins syndrome locus and each polymorphic marker and also between each pair of markers. Errors in the data were eliminated by reisolating genomic DNA from a second blood sample, re-genotyping, and reanalysis of autoradiographs by two independent laboratories for individuals in whom recombinants or higher order recombination events were noted among loci. Paternity was genetically tested using markers on four different chromosomes with heterozygosities of greater than 0.70. The computer program HOMOG was used to test for heterogeneity (Ott, 1985).

A genetic map was developed using data generated from the CEPH families and the computer program CRIMAP (Lander and Green, 1987). The microsatellite markers IL9, CSF1R, D5S207, D5S210, and D5S119 were genotyped in our laboratory; genotypes from the RFLP markers SPARC and GRL were extracted from Version 6 of the CEPH database. The genotypes were generated for locus GRL in the laboratory of Kenneth K. Kidd (Yale University, CEPH collaborating laboratory No. 8) and for marker SPARC in the laboratory of Albert de la Chapelle (Helsinki, Finland, CEPH collaborating laboratory No. 77). Likelihoods for all possible orders among IL9, GRL, D5S207, D5S210, CSF1R, SPARC, and D5S119 were calculated, under hypotheses assuming equal female:male recombination fractions (L_1) and assuming sex differences in the recombination fraction (L_2). The support for the most likely order was computed by subtracting the likelihood of the second most likely order from that of the most likely order. To test the hypothesis of sex-specific differences in the recombination fraction across the map, we compared L_2 and L_1 , using a χ^2 test ($\chi^2 = 2 \ln(L_2 - L_1)$) with degrees of freedom equal to the number of recombination fractions estimated under L_2 minus the number estimated under L_1 . Following development of the genetic map, all double and higher order recombination events were discerned using the CHROMPIC options of CRIMAP. Families in which these recombination events occurred were re-genotyped for the microsatellite markers to evaluate potential data error. After error correction, the genetic map was redeveloped as described above.

Physical order by YAC contig. The subpools of YAC clones from the total human YAC libraries of Albertsen *et al.* (1990) and Burke *et al.* (1987) were screened by PCR using primer sets that generated DNA fragments specific for the 3' end of GRL (Warrington *et al.*, 1991), D5S207 (Weber *et al.*, 1991), and D5S210 (Weber *et al.*, 1991). The primers specific for the 5' end or exon 1 of GRL (forward primer

5'-CTTTCTTAAATAGGGGCTCTC-3' and reverse primer 5'-TCTGGCAGAGGAGCCGCT-3', PCR product size 177 bp) were derived from the sequence of the gene (GenBank Accession No. M60597). Primers (forward primer 5'-GTAAGTGCTAGCGATAAC-3' and reverse primer 5'-TATAGCAGACTCAGGTCCTC3-', PCR product size 345 bp) for the locus ADRB2 were also derived from the sequence of the gene (GenBank Accession No. J02960). YAC contigs were established by STS content mapping and limited YAC walking. This was accomplished by defining YAC end fragments by PCR with vector and *Alu* primers and DNA sequencing PCR products (Nelson *et al.*, 1991). The STS sequences for the left end of YAC A113A10 containing locus GRL were forward primer 5'-TCACGAGGGCTTGTA-GTAGGT-3' and reverse primer 5'-AGTCACGGCACCCAGCCAAT-3' (PCR product size 600 bp) and for the left end of YAC A117E2 containing locus D5S207 were forward primer 5'-CTTGAGATCGGG-CGTTTCGACT-3' and reverse primer 5'-CATCAGTAGTCACAGAAG-TCAC-3' (PCR product size 140 bp). The PCR amplifications (30 cycles) were carried out by denaturing at 94°C for 30 s, annealing at 55–60°C for 30 s–2 min, and extending at 72°C for 1 min. YACs were screened for chimerism by *in situ* hybridization using purified YAC DNA as described in Saltman *et al.* (1993). A YAC designated with a letter followed by three numbers and then a letter and numbers was derived from the Burke *et al.* (1987) library. A YAC designated with three numbers and then a letter and numbers was obtained from the Albertsen *et al.* (1990) library.

RESULTS

Linkage Analysis between TCOF1 and Other Chromosome 5q3 Loci

We performed two-point linkage analysis between the Treacher Collins syndrome locus (TCOF1) and 12 loci (Table 1). The highest lod score was obtained between the D5S210 and the TCOF1 loci ($Z = 10.52$), and the TCOF1 locus was estimated to be 2 cM (confidence limits ± 7 cM) or 9 cM in either direction from D5S210. There was no recombination between the TCOF1 and GRL loci. For each locus D5S207, D5S210, D5S376, CSF1R, SPARC, D5S119, or D5S209 there were three or fewer obligate recombinants with the TCOF1 locus. More than three obligate recombinations were observed between the TCOF1 locus and locus IL9, FGFA, or D5S527. There were more meiotic recombinations occurring in females than in males as evidenced by the higher female than male recombination fractions between locus TCOF1 and chromosome 5q3 loci. No evidence for heterogeneity was obtained by HOMOG ($P = 0.5$).

Two-point linkage analysis among chromosome 5q3 loci revealed no recombination between D5S207 and GRL ($Z = 7.30$) and SPARC and D5S119 ($Z = 3.91$). Locus CSF1R was tightly linked to loci D5S210 ($Z = 13.82$, $\hat{\theta} = 0.03$) and D5S376 ($Z = 12.03$, $\hat{\theta} = 0.03$). Multi-point linkage analysis with any 2 of the 12 loci and the disease locus to establish the position of the Treacher Collins syndrome gene with respect to these markers did not reveal a best order with significant odds.

Genetic Map of Chromosome 5q3 Loci

Identification of loci that flank locus D5S210 at a genetic distance of at least 9 cM would determine loci that flank the region containing the TCOF1 locus. To determine such loci, the best order and the genetic distances

TABLE 1
Two-Point Lod Scores between TCOF1 and 5q3 Loci

Locus	H	FLP	Z	$\hat{\theta}$	$Z_{t,m}$	$\hat{\theta}_f$	$\hat{\theta}_m$
IL9	0.80	(CA) _n	2.16	0.19	2.50	0.25	0.08
FGFA	0.88	(CA) _n	2.12	0.17	3.05	0.34	0.05
GRL	0.49	BclI	4.19	0.00	4.19	0.00	0.00
D5S207	0.69	(CA) _n	2.70	0.13	3.39	0.23	0.00
D5S210	0.75	(CA) _n	10.58	0.02	10.80	0.03	0.00
D5S376	0.74	(CA) _n	5.58	0.07	5.58	0.07	0.07
CSF1R	0.86	(CA) _n	6.26	0.11	6.35	0.12	0.07
SPARC	0.44	TaqI	3.00	0.08	3.37	0.12	0.00
D5S119	0.49	(CA) _n	1.72	0.15	1.91	0.22	0.09
D5S209	0.71	(CA) _n	2.00	0.07	2.11	0.09	0.00
D5S527	0.89	(CA) _n	0.98	0.27	1.31	0.35	0.16
FGFR4	0.30	EcoRI	0.05	0.43	0.05	0.44	0.42

Note. H, heterozygosity; FLP, restriction or PCR fragment length polymorphism; Z, maximum lod score, sex average; $Z_{t,m}$, sex difference; $\hat{\theta}$, recombination fraction at Z; $\hat{\theta}_f$, $\hat{\theta}$ females; $\hat{\theta}_m$, $\hat{\theta}$ males.

among markers for loci IL9, GRL, D5S207, D5S210, CSF1R, SPARC, and D5S119 were analyzed in all 40 CEPH pedigrees. Two-point lod scores and maximum likelihood estimates of the recombination fraction for sex-average and sex-specific differences were obtained among these loci (data not shown). No recombination was noted between D5S207 and D5S210 with a maximum lod score of 72.55. The recombination fractions among loci were consistent with those obtained in our two-point linkage analysis for the same loci in Treacher Collins syndrome families. For example, the lack of recombination between D5S207 and D5S210 ($Z = 12.47$) was also found when two-point linkage analysis was performed in the Treacher Collins syndrome families.

From our CEPH data, the order IL9-D5S207/D5S210-CSF1R-SPARC-D5S119 was supported over the next most likely order with odds >1000:1, regardless of whether likelihoods were calculated assuming equal female:male recombination fractions or allowing for differences in the female:male recombination frequencies (Table 2). Evidence in favor of sex-specific differences in recombination was significant ($\chi^2_4 = 17.75$; $P = 0.001$). The most likely position for GRL was between IL9 and D5S207/D5S210, although this placement was supported with odds of only 200:1 over the next most likely position between D5S207/D5S210 and CSF1R. Loci IL9 and SPARC flank D5S210 at a genetic distance equal to or greater than 9 cM and, therefore, these loci flank the TCOF1 locus.

Physical Order of Loci

The physical order was then determined for loci in the chromosome 5q3 region. The consensus physical map based on radiation hybrid and *in situ* hybridization analyses placed the following loci in the order IL9-FGFA-GRL-D5S207/D5S210-ADRB2-CSF1R-SPARC-D5S119-D5S209-FGFR4 (Warrington *et al.*, 1991, 1992; Ryan *et al.*, 1992; Saltman *et al.*, 1993; Westbrook *et al.*, 1992). For the markers in common between our

genetic map and previously published physical maps, the orders were consistent.

We determined the relative order of GRL, D5S207, and D5S210 with respect to the centromere and telomere by constructing a YAC contig. Radiation hybrid mapping data suggested that D5S207 was centromeric of D5S210 (J. Wasmuth, 1992, Univ. of California, Irvine, pers. comm.). By using primer pairs specific for the 3' ends of the GRL gene and D5S207, we have established an overlap between YACs that contain the GRL locus and YACs that contain the D5S207 marker. The left end of the YAC A113A10 containing the GRL locus detected YAC A117E2, which contains the D5S207 locus, and the left end of A117E2 detected A113A10. Neither of these YACs contained locus D5S210 or the STS for the 5' end of GRL. This demonstrates that the order of these three loci is centromere-5'-GRL-3'-D5S207-D5S210-telomere.

We refined the location of D5S376. Dixon *et al.* (1992) had reported that locus D5S376 is centromeric to the

TABLE 2
Best Order of 5q3 Loci Based on Multipoint Linkage Analysis

Male θ	Female θ	Sex average θ	Loci	Odds
0.06	0.11	0.08	IL9] >1000:1
			GRL	
0.00	0.02	0.01	D5S207/D5S210] >200:1
0.04	0.10	0.07	CSF1R] >1000:1
0.06	0.12	0.10	SPARC] >1000:1
0.07	0.02	0.04	D5S119] >1000:1

Note. θ , recombination fraction or genetic distance (Morgans) between loci.

locus D5S527. We have found that locus D5S376 is located within two YAC clones (167B12 and 335D9) that also contain locus ADRB2. It also was previously known that ADRB2 is located between D5S207/D5S210 and CSF1R by both radiation hybrid (Warrington *et al.*, 1992) and *in situ* hybridization (Saltman *et al.*, 1993) analyses. Therefore, the order of loci with respect to D5S376 is centromere–D5S210–D5S376–CSF1R–telomere.

The order of loci centromere–CSF1R–SPARC–telomere was previously supported by radiation hybrid analysis (Warrington *et al.*, 1991, 1992). However, FISH data using CSF1R and SPARC cosmids had ordered SPARC centromeric of CSF1R (Saltman *et al.*, 1993). To resolve this discrepancy further FISH studies were performed using a CSF1R YAC B233G3 and the SPARC cosmids described in Saltman *et al.* (1993). These data clearly demonstrated the order to be centromere–CSF1R–SPARC–telomere (Lovett, unpublished data). The two different FISH results could be due to the different DNA content in the CSF1R YAC and cosmid. The order generated by using FISH and the CSF1R YAC is most likely to be correct because it concurs with the radiation hybrid data, and the large insert size of the nonchimeric YAC increases the specificity of the FISH probe. Therefore, we established the order of centromere–5'–GRL–3'–D5S207–D5S210–D5S376–CSF1R–SPARC–telomere.

Mapping TCOF1 Relative to the Other Loci in This Region by Haplotype Analysis

We refined the location of the TCOF1 locus by haplotype analysis using the order of markers IL9–FGFA–GRL–D5S207–D5S210–D5S376–CSF1R–SPARC–D5S119–D5S209, which was based on previously published data and our physical map. By analysis of families with obligate meiotic recombinants in affected and unaffected members and assuming that the fewest number of recombinations had occurred among these loci, the TCOF1 locus appears to be located between loci CSF1R and SPARC (Figs. 1 and 2). The size of this region has been determined previously to be 680 kb by *in situ* hybridization (Saltman *et al.*, 1993) and 910 kb by radiation hybrid (Warrington *et al.*, 1991) analyses.

However, if we cautiously interpret our data and consider obligate meiotic recombinants in only affected individuals and not unaffected individuals to exclude the rare possibility that nonpaternity would not be detected by standard screening methods and exclude unaffected individuals who are so mildly affected that they cannot be diagnosed by standard clinical or radiologic examination, then the location of the TCOF1 locus could be refined only to a region distal of CSF1R. Furthermore, because some of the loci were uninformative with respect to their segregation in these families, it is possible (but less likely) that double recombination events may have occurred among markers in this region. If this were the case, then the gene could not be definitely localized between two loci. If more than one meiotic recombin-

ation event occurred in the region centromeric to D5S376 in individuals 20 and 22 in family 7, the gene may reside between loci FGFA and D5S207, between D5S207 and D5S376, or distal to CSF1R (Fig. 1). In family 9 the gene may be located proximal to FGFA, between FGFA and D5S207, or distal to D5S376 (Fig. 2). Taking the data of both families together, the region of overlap that may contain the gene could be between FGFA and D5S207 or distal to CSF1R (see outlined left lower corner of Fig. 2). Each of the regions between two markers has been estimated by *in situ* hybridization to be less than 680 kb in size (Saltman *et al.*, 1993).

Dixon *et al.* (1992) reported D5S527 as a flanking telomeric marker of TCOF1. By haplotype analysis, D5S527 appears to be more telomeric than D5S209. Obligate meiotic recombinations between loci TCOF1 and D5S527 occurred in the same individuals as recombination events between loci TCOF1 and D5S209; in addition, there were two obligate meiotic recombinations between loci TCOF1 and D5S527 in other individuals in the Treacher Collins syndrome families (data not shown).

DISCUSSION

We have established a physical map order of IL9–5'GRL–3'–D5S207–D5S210–D5S376–CSF1R. Combining these data with previously known maps, the order of markers is centromere–IL9–FGFA–GRL–D5S207–D5S210–D5S376–CSF1R–SPARC–D5S119–D5S209–telomere. The genetic distance between IL9 and D5S119 is estimated to be 30 cM (see sum of genetic distances in Table 2), whereas the physical distance is estimated to be 10.5 Mb by radiation hybrid and 7 Mb by *in situ* hybridization analyses (Warrington *et al.*, 1991, 1992; Saltman *et al.*, 1993; Westbrook *et al.*, 1992). No recombination was observed among loci D5S207 and D5S210, and the physical distance between loci D5S207 and D5S210 is approximately 280 kb as assessed by radiation hybrid analysis (J. Wasmuth, 1992, Univ. of California at Irvine, pers. comm.).

Linkage analysis demonstrated that the TCOF1 gene is within a region that spans 9 cM centromeric and 9 cM telomeric of D5S210 and is flanked proximally by IL9 and distally by SPARC. Our haplotype analysis within this region showed that the gene is flanked proximally by FGFA. Dixon *et al.* (1992) determined in their families that the TCOF1 locus maps between loci D5S376 and D5S527 by multipoint linkage analysis with significant odds. By pairwise linkage analysis, they determined that locus D5S376 is centromeric of locus D5S527 and that these markers were 7.6 cM apart. Our haplotype data are consistent with their results in that the region containing the TCOF1 gene appears to be telomeric of D5S376, and our data suggest that the location of the TCOF1 locus may be further refined between loci CSF1R and SPARC, a region smaller than and within the region proposed by Dixon *et al.* (1992). If this is the region of the Treacher Collins syndrome gene, it is less than 1 Mb in size.

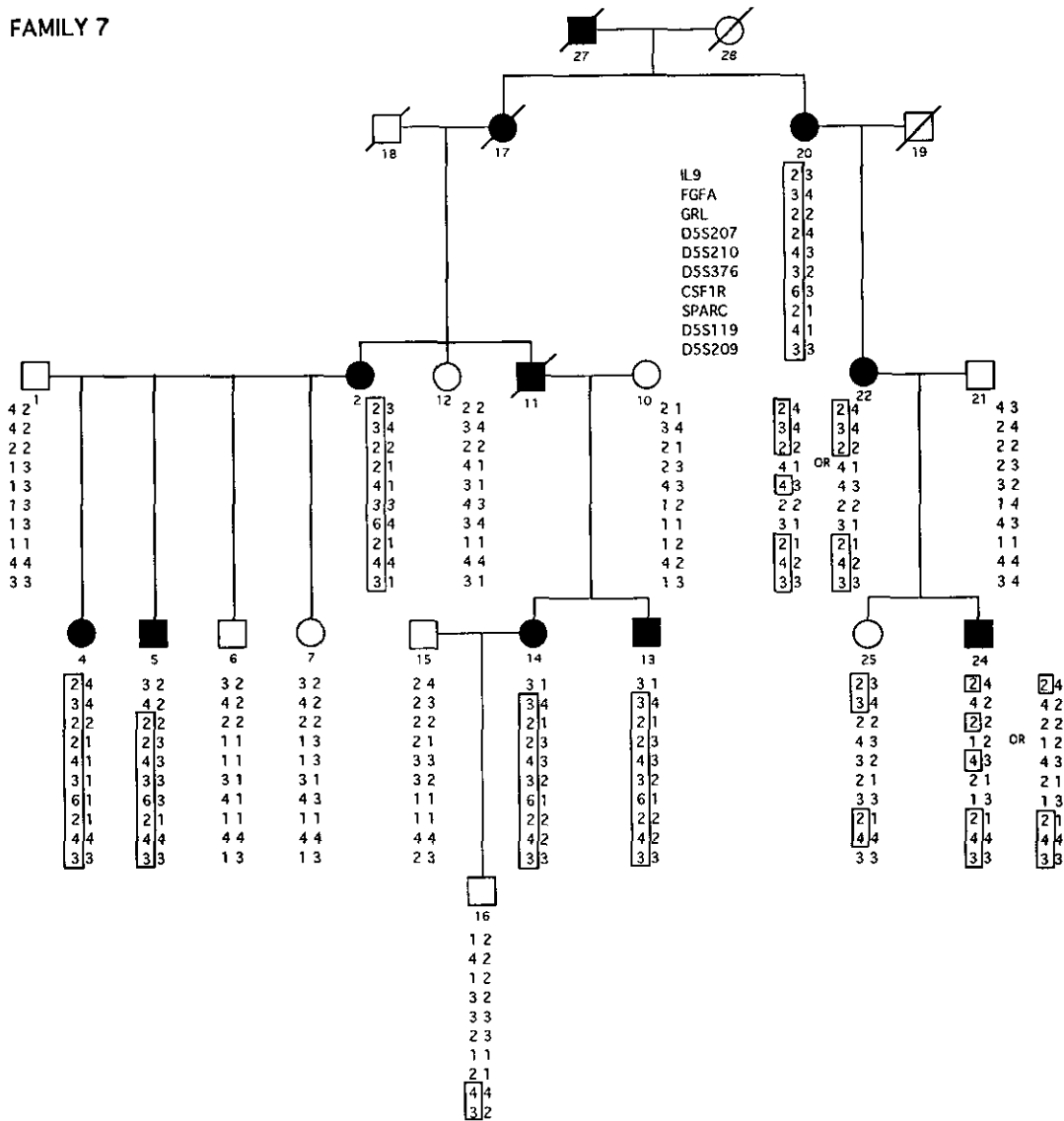


FIG. 1. Haplotype analysis of family 7. Affected members of the pedigree are indicated by black symbols. For each member, the polymorphic alleles of each locus are assigned and ordered according to the physical map of this region. The haplotype contributed by the affected parent is on the left and the unaffected parent is on the right. For affected members the alleles of loci that segregate with the disease or that are uninformative and may possibly segregate with the disease are boxed. The boxes denote the largest possible region that may contain the disease gene. In the case of individuals 22 and 24, two alternative interpretations that are dependent on two (right haplotypes) or more (left haplotypes) meiotic recombinations are given. For unaffected individuals, the boxed areas include only alleles known to have been inherited from an affected parent and denote the minimal region from which the disease locus can be excluded.

By two-point linkage analysis, the genetic distances of 11 cM between loci CSF1R and TCOF1 and 8 cM between loci SPARC and TCOF1 are surprising if the physical distance between CSF1R and SPARC is less than 1 Mb, because our data suggest that the TCOF1 locus is between the CSF1R and SPARC loci. Our multipoint linkage analysis based on all 40 CEPH pedigrees demonstrated that CSF1R and SPARC loci are 10 cM apart, corroborating the high recombination rate in this region. It is interesting that genetic and physical distances differ dramatically in this nontelomeric region. Furthermore, the high recombination rate in the region between CSF1R and SPARC and the increased informativeness of the D5S210 marker over that of the CSF1R and SPARC markers in our families may be the reasons that

the highest lod score was obtained between loci TCOF1 and D5S210 and not between TCOF1 and locus CSF1R or SPARC even though the latter loci are physically closer to one another than the former. Whether this high recombination rate in this region is a clue to the mechanism of how the Treacher Collins syndrome mutation(s) is induced remains to be answered when the gene is identified.

Candidate genes for Treacher Collins syndrome in the 5q3 region included FGFA (reviewed in Jaye *et al.*, 1986), SPARC (Villarreal *et al.*, 1989), and FGFR-4 (Korhonen *et al.*, 1992). Although these genes are expressed in mammalian facial tissues, they have been excluded as the Treacher Collins syndrome gene because of the presence of recombination between these genes and the

FAMILY 9

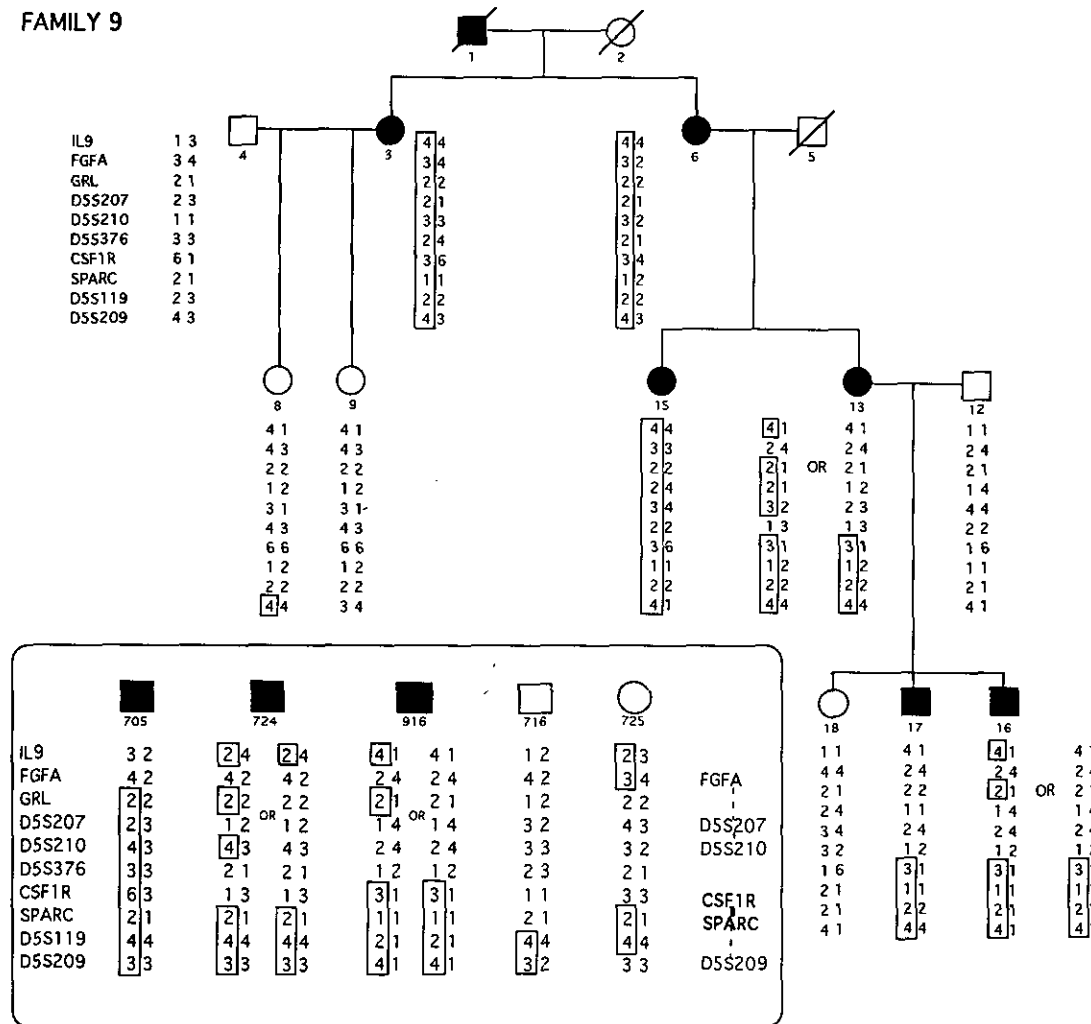


FIG. 2. Haplotype analysis of family 9. The analysis is performed as explained in Fig. 1. Note that different areas are boxed if a single (right haplotypes) or more (left haplotypes) meiotic recombination events are predicted in individuals 13, 16, and 17. (Outlined Lower Left) The haplotypes for the most informative individuals (denoted by family number followed by the individual number) from both families are presented. If we assume more than one meiotic recombination event resulted in the haplotypes of individuals 724 and 916, then the gene may reside between FGFA and D5S207 or CSF1R and SPARC. However, it is more likely that the least number of recombination events occurred and that the gene is present between CSF1R and SPARC. To eliminate errors in diagnosis if we do not incorporate the haplotype information provided by the unaffected individual 725, then the gene may reside distal of CSF1R. The clinical and radiologic examination by two geneticists of individuals 716 and 725 revealed no evidence that they were affected. Therefore, the most likely location of the gene is between CSF1R and SPARC as denoted by the bold type. Less likely locations of the gene are denoted by the dashed lines on the right side of the outlined area.

TCOF1 locus. To identify other potential candidate genes, we plan to screen the chromosome 5q3 region, especially between loci CSF1R and SPARC, for cDNAs from a fetal craniofacial cDNA library using YAC and cosmid contigs from this region.

ACKNOWLEDGMENTS

We thank Kay Huebner and John Wasmuth for helpful discussions, Denis Le Paslier for providing the YAC clones, Margaret A. Pericak-Vance for computer time, Albert de la Chapelle for his genotype data for locus SPARC, Elisabeth Schweine and Alex Kazantsev for technical assistance, and Mahin Golabi for patient material. The work was supported by NIH Grants DE10180, HD24061, RR00052, and RR00722 (E.W.J.); HG00368 (M.L.); NS2663005 (M.P.-V.); MH39239 (K.K.K.) and the Muscular Dystrophy Association (L.H.Y.).

REFERENCES

Albertsen, H. M., Abderrahim, H., Caan, H. M., Dausset, J., Le Paslier, D. L., and Cohen, D. (1990). Construction and characterization of a yeast artificial chromosome library containing seven haploid human genome equivalents. *Proc. Natl. Acad. Sci. USA* **87**: 4256-4260.

Armstrong, E., Hastabacka, J., Partanen, J., Huebner, K., and Alitalo, K. (1991). RFLPs in the fibroblast growth factor receptor-4 locus (FGFR4) in 5q33-qter. *Nucleic Acids Res.* **19**: 5096.

Biggin, M. D., Gibson, T. J., and Hong, G. F. (1983). Buffer gradient gels and ³⁵S label as an aid to rapid DNA sequence determination. *Proc. Natl. Acad. Sci. USA* **80**: 3963-3965.

Burke, D. T., Carle, G. F., and Olson, M. V. (1987). Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science* **236**: 806-812.

Dixon, M. J., Haan, E., Baker, E., David, D., McKenzie, N., Williamson, R., Mulley, J., Farrall, M., and Callen, D. (1991). The gene for Treacher Collins syndrome maps to the long arm of chromosome 5. *Am. J. Hum. Genet.* **49**: 17-22.

- Dixon, M. J., Dixon, J., Raskova, D., Le Beau, M. M., Williamson, R., Klinger, K., and Landes, G. M. (1992). Genetic and physical mapping of the Treacher Collins syndrome locus: Refinement of the localization to chromosome 5q32-33.2. *Hum. Mol. Genet.* **1**: 249-253.
- Gorlin, R. J., Cohen, M. M., and Levin, L. S. (1990). "Mandibulofacial Dysostosis in Syndromes of the Head and Neck," 3rd ed., Oxford Univ. Press, New York.
- Hollenberg, S. M., Weinberger, C., Ong, E. S., Cerelli, G., Oro, A., Lebo, R., Thompson, E. B., Rosenfeld, M. G., and Evans, R. M. (1985). Primary structure and expression of a functional human glucocorticoid receptor cDNA. *Nature* **318**: 635-641.
- Jabs, E. W., Li, X., Coss, C. A., Taylor, E. W., Meyers, D. A., and Weber, J. L. (1991a). Mapping the Treacher Collins syndrome locus to 5q31.3 → q33.3. *Genomics* **11**: 193-198.
- Jabs, E. W., Coss, C. A., Hayflick, S. J., Whitmore, T. E., Pauli, R. M., Kirkpatrick, S. J., Meyers, D. A., Goldberg, R., Day, D. W., and Rosenbaum, K. N. (1991b). Chromosomal deletion 4p15.32 → p14 in a Treacher Collins syndrome patient: Exclusion of the disease locus from and mapping of anonymous DNA sequences to this region. *Genomics* **11**: 188-192.
- Jaye, M., Howk, R., Burgess, W., Ricca, G. A., Chiu, I. M., Ravera, M. W., O'Brien, S. J., Modi, W. S., Maciag, T., and Drohan, W. N. (1986). Human endothelial cell growth factor: Cloning, nucleotide sequence, and chromosome localization. *Science* **233**: 541-545.
- Korhonen, J., Partenen, J., and Alitalo, K. (1992). Expression of FGFR-4 mRNA in developing mouse tissues. *Int. J. Dev. Biol.* **36**: 323-329.
- Lander, E. S., and Green, P. (1987). Construction of multilocus genetic linkage maps in humans. *Proc. Natl. Acad. Sci. USA* **84**: 2363-2367.
- Lathrop, G. M., Lalouel, J.-M., Julier, C., and Ott, J. (1984). Strategies for multipoint linkage analysis in humans. *Proc. Natl. Acad. Sci. USA* **81**: 3443-3446.
- Li, X., Jaye, M., Crumley, G., and Jabs, E. W. (1992). Dinucleotide repeat polymorphism in the human fibroblast growth factor acidic (FGFA) gene on chromosome 5. *Hum. Mol. Genet.* **1**: 216.
- Lovett, M., Kere, J., and Hinton, L. M. (1991). Direct selection: A method for the isolation of cDNAs encoded by large genomic regions. *Proc. Natl. Acad. Sci. USA* **88**: 9628-9632.
- Nelson, D. L., Ballabio, A., Victoria, M. F., Pieretti, M., Bies, R. D., Gibbs, R. A., Maley, J. A., Chinault, A. C., Webster, T. D., and Caskey, C. T. (1991). Alu-primed polymerase chain reaction for regional assignment of 110 yeast artificial chromosome clones from the human X chromosome: Identification of clones associated with a disease locus. *Proc. Natl. Acad. Sci. USA* **88**: 6157-6161.
- Ott, J. (1974). Estimation of the recombination fraction in human pedigrees: Efficient computation of the likelihood for human linkage studies. *Am. J. Hum. Genet.* **26**: 588-597.
- Ott, J. (1985). "Analysis of Human Genetic Linkage," The Johns Hopkins Univ. Press, Baltimore, MD.
- Polymeropoulos, M. H., Xiao, H., Rath, D. S., and Merrill, C. R. (1991a). Dinucleotide repeat polymorphism at the human interleukin 9 gene. *Nucleic Acids Res.* **19**: 688.
- Polymeropoulos, M. H., Xiao, H., Rath, D. S., and Merrill, C. R. (1991b). Dinucleotide repeat polymorphism at the human c-fms protooncogene for the CFS-1 receptor (CSF1R). *Nucleic Acids Res.* **19**: 1160.
- Ryan, S. G., Dixon, M. J., Nigro, M. A., Kelts, K. A., Markand, O. N., Terry, J. C., Shiang, R., Wasmuth, J. J., and O'Connell, P. (1992). Genetic and radiation mapping of the hyperekplexia region on chromosome 5q. *Am. J. Hum. Genet.* **51**: 1334-1343.
- Saltman, D. L., Dolganov, G. M., Warrington, J. A., Wasmuth, J. J., and Lovett, M. (1993). A physical map of fifteen loci on human chromosome 5q23-33 by two-color fluorescence *in situ* hybridization. *Genomics* **16**: 726-732.
- Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). "Molecular Cloning: A Laboratory Manual," Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Villarrael, X. L., Mann, K. G., and Long, G. L. (1989). Structure of human osteonectin based upon analysis of cDNA and genomic sequences. *Biochemistry* **28**: 6483-6491.
- Warrington, J. A., Bailey, S. K., Armstrong, E., Aprelikova, O., Alitalo, K., Dolganov, G. M., Wilcox, A. S., Sikela, J. M., Wolfe, S. F., Lovett, M., and Wasmuth, J. J. (1992). A radiation map of 18 growth factor, growth factor receptor, hormone receptor, or neurotransmitter receptor genes on the distal region of the long arm of chromosome 5. *Genomics* **13**: 803-808.
- Warrington, J. A., Hall, L. V., Hinton, L. M., Miller, J. N., Wasmuth, J. J., and Lovett, M. (1991). Radiation hybrid map of 13 loci on the long arm of chromosome 5. *Genomics* **11**: 701-708.
- Weber, J. L., Polymeropoulos, M. H., May, P. E., Kwitek, A. E., Xiao, H., McPherson, J. D., and Wasmuth, J. J. (1991). Mapping of human chromosome 5 microsatellite DNA polymorphisms. *Genomics* **11**: 695-700.
- Westbrook, C. A., Neuman, W. L., McPherson, J., Camper, S., Wasmuth, J., Plaetke, R., and Williamson, R. (1992). Report of the second international workshop on human chromosome 5 mapping. *Cytogenet. Cell Genet.* **61**(4): 224-231.