Am. J. Hum. Genet. 45:348–353, 1989

Association of Genetic Variation of the Transforming Growth Factor-Alpha Gene with Cleft Lip and Palate

Holly H. Ardinger,* Kenneth H. Buetow,‡ Graeme I. Bell,§ Januz Bardach,† Duane R. VanDemark,† and Jeffrey C. Murray*

*Division of Medical Genetics, Department of Pediatrics, and †Department of Otolaryngology, University of Iowa, Iowa City; ‡Fox Chase Cancer Center, Philadelphia; and §Howard Hughes Medical Institute, Departments of Biochemistry and Molecular Biology and of Medicine, University of Chicago, Chicago

Summary

Complex segregation analysis of pedigrees having nonsyndromic cleft lip with or without cleft palate (CL/P) (Chung et al. 1986; Marazita et al. 1986) has shown that a major-locus model best explains the observed recurrence of CL/P in Caucasian families. To identify this major gene, we compared the frequencies of 12 RFLPs at five loci—epidermal growth factor, transforming growth factor-α, epidermal growth factor receptor, glucocorticoid receptor, and estrogen receptor—in both a group of 80 subjects with nonsyndromic CL/P and 102 controls. These candidate genes were selected because studies in rodents had suggested their possible involvement in palatogenesis. A significant association was observed between two RFLPs at the transforming-growth-factor-α (*TGFA*) locus and the occurrence of clefting ($P = .0047$ and $P = .0052$). This suggests that either the TGFA gene itself or DNA sequences in an adjacent region contribute to the development of a portion of cases of CL/P in humans and provides an opportunity to begin to examine the molecular events underlying lip and palate formation.

Introduction

Cleft lip with or without cleft palate (CL/P) is one of the most common birth defects, affecting 1/700–1/1,000 Caucasians. Since the risk of recurrence of CL/P within a family is 28–40-fold greater than that for the general population, genetic factors are thought to contribute to the development of this disorder. It has been suggested that nonsyndromic clefting in humans is most likely due to a combination of genetic and environmental factors as described by the multifactorial threshold model (Carter 1969; Fraser 1976).

Recently, Chung et al. (1986) and Marazita et al. (1986) reanalyzed published Danish pedigree data on CL/P by using new methods of complex segregation analysis. Several alternative models of clefting etiology were examined, including sporadic, multifactorial, major gene alone, and a mixed model. The pedigree data best fitted the model of action of a major gene with autosomal recessive inheritance modified by additional genetic and/or environmental factors. This major gene was predicted to have a frequency of 3.5% and to account for about one-third of the CL/P cases in the Danish population (Chung et al. 1986).

To identify this major gene, we performed an association study comparing genetic variation at candidate gene loci in a group of unrelated clefted individuals with that of a group of controls. Our hypothesis was that one specific mutational event in a candidate gene might underlie a substantial portion of clefting in Caucasians. Such a mutation could be identified through its nonrandom association with nearby RFLPs. The candidate genes we investigated were selected owing to their role in normal and abnormal palatogenesis in rodents (Pratt et al. 1980; Ferguson 1987). They included the glucocorticoid receptor (GRL) and estrogen receptor (ER), which share both sequence and functional homology, and epidermal growth factor (EGF) and transforming growth factor-α and their common receptor, the EGF receptor (EGFR).

Received February 3, 1989; revision received May 1, 1989.

Address for correspondence and reprints: Holly H. Ardinger, M.D., Division of Medical Genetics, Department of Pediatrics, University of Iowa, Iowa City, IA 52242.

0002-9297/89/4503-0002$02.00

Material and Methods

Patients

Eighty unrelated Caucasian patients were recruited through the University of Iowa Cleft Palate Clinic and the Iowa Regional Genetic Counseling Service. These patients, who were born between 1956 and 1987, had isolated CL/P but no other abnormalities. The distribution of cleft type was 29% bilateral CL+P, 54% unilateral CL+P, 2% bilateral CL, and 15% unilateral CL. Family histories for CL/P, as well as blood samples, were collected from each case after signed informed consent was obtained. A control group comprised (1) 63 unrelated Caucasians, without clefting, who had been born in or around the state of Iowa between 1914 and 1987 and on whom blood samples were available and (2) 39 unrelated Caucasians, born at University of Iowa Hospital during October and November, 1987, on whom placentas were available. Seventy percent of the cases were male, and 45% of the controls were male.

DNA Analysis

DNA was extracted from either leukocytes or placental tissue by using a modification of the procedure of Poncz et al. (1982). Five micrograms DNA was digested with the restriction endonuclease known to yield RFLP with the probes being studied, according to manufacturer's instructions (New England Biolabs or Bethesda Research Laboratories). DNA fragments were separated by electrophoresis on 0.8% or 1.2% agarose gels. DNA was transferred to Zetabind™ nylon filters (AMF-Cuno). Probes used are shown in table 1 and are described in the corresponding references. They were labeled by using the random primer method of Feinberg and Vogelstein (1983), with [α^{32}-P] dCTP. Following hybridization and washing, autoradiography was carried out by using intensifying screens at −80°C. The resulting films were scored for the presence or absence of the various polymorphic alleles (table 1 and fig. 1).

Statistical Analysis

Allele-frequency differences between cases and controls were evaluated by χ^2 analysis. Haplotypes were constructed for multiple RFLPs detected at the same locus where one or more of the individual RFLPs showed significant association. Haplotypes were derived by inspection of fully homozygous individuals or individuals heterozygous at a single site and were reanalyzed using logistic regression and log-linear models. A frequency bias in the haplotypes is created by not being able to identify haplotypes for all cases and controls, but we have no reason to expect such a bias is different for CL/P cases and controls.

Results

Table 2 gives the RFLP distribution in cases and controls. Rare alleles found for *EGFR/Hin*dIII, *EGFR/Pst*I, *EGFR/Mbo*I, *EGF/Hinc*II, and *TGFA/Bam*HI were excluded from analysis but account for the odd numbers of chromosomes shown in some case and control samples. Comparisons of the RFLP common allele frequencies at the GRL, ER, EGF, and EGFR loci each revealed no significant differences between cases and controls ($P > .05$).

Table 1

RFLPs Studied

Gene	Probe	Location	Enzyme	Polymorphism Size (1) Allele/(2) Allele	Reference
GRL	OB7	5q11-q13	*Bcl*I	4.5 kbp/2.3 kbp	Murray et al. 1987
ER.....	OR8	6q24-q27	*Pvu*II	1.5 kbp/0.7 kbp	Castagnoli et al. 1987
EGF.....	phEGF121	4q25-q27	*Hinc*II	8.0 kbp/4.5 kbp	Murray et al. 1986*b*
			*Sac*I	12.0 kbp/11.0 kbp	Murray et al. 1986*b*
EGFR	HER-A64	7p13-p12	*Hae*III	2.6 kbp/1.7 kbp	Smith et al. 1987
			*Hin*dIII	12.5 kbp/10 kbp	Smith et al. 1987
			*Mbo*I	2.6 kbp/2.3 kbp	Smith et al. 1987
			*Pst*I	10.0 kbp/8.0 kbp	Smith et al. 1987
			*Stu*I	20.0 kbp/13.0 kbp	Smith et al. 1987
TGFA	phTGF1-10-3350	2p13	*Taq*I	3.0 kbp/2.7 kbp	Hayward et al. 1987
	phTGF1-10-925		*Bam*HI	7.0 kbp/4.0 kpb	Murray et al. 1986*a*
			*Rsa*I	1.5 kbp/1.2 kbp	Murray et al. 1986*a*

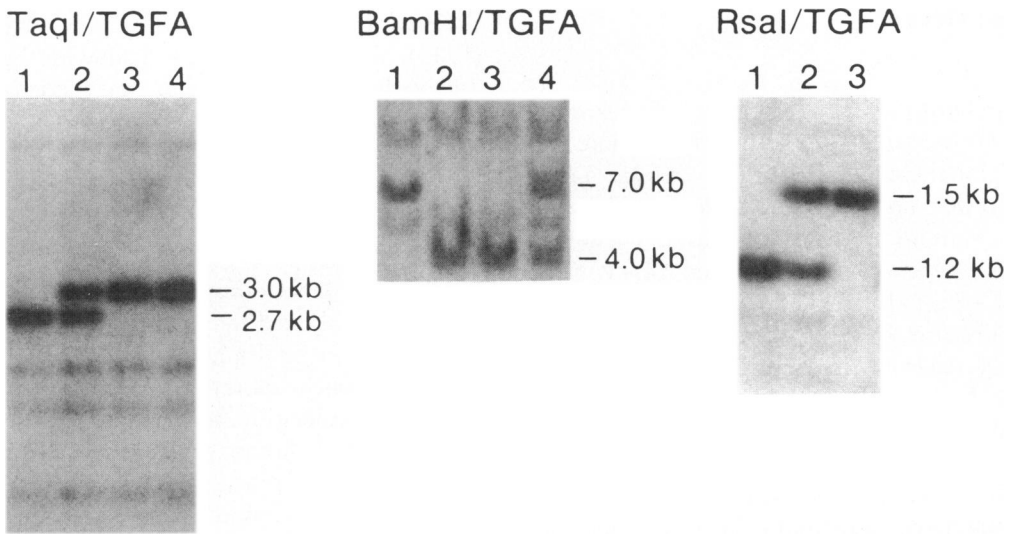

Figure 1 *TGFA* RFLPs used in the present study. Restriction enzymes used and polymorphic band sizes are indicated. Alleles designated as "1" in the text are the larger band in each pair of two polymorphic bands.

Highly significant associations were observed, however, for two of three RFLPs detected by the *TGFA* probes and the presence of clefting (tables 2 and 3). The *TGFA* *Taq*I 2.7-kbp allele (C2) and the *TGFA* *Bam*HI 4.0-kbp allele (A2) showed significant association when standard criteria were used with *P* values of .0047 and .0052, respectively. They approached significance when the rejection criteria were adjusted for

Table 2

RFLP Distribution and Analysis

Gene Probe	Enzyme	Allele[a]	No. of CL/P Chromosomes	No. of Control Chromosomes	χ^2	*P* Value
GRL	*Bcl*I	A1	61	50	.506	NS
		A2	97	94		
ER	*Pvu*II	A1	69	58	.113	NS
		A2	81	74		
EGF	*Hinc*II	A1	67	52	1.013	NS
		A2	89	88		
	*Sac*I	B1	142	123	.520	NS
		B2	14	9		
EGFR	*Hae*III	A1	9	5	.936	NS
		A2	147	137		
	*Hind*III	C1	144	124	.082	NS
		C2	16	15		
	*Mbo*I	E1	7	4	.130	NS
		E2	116	55		
	*Pst*I	B1	142	127	.082	NS
		B2	16	16		
	*Stu*I	D1	44	31	1.799	NS
		D2	114	111		
TGFA	*Taq*I	C1	135	186	7.497	.0047
		C2	21	10		
	*Bam*HI	A1	7	25	7.483	.0052
		A2	150	170		
	*Rsa*I	B1	44	54	.039	NS
		B2	116	148		

[a] Designations are as found in Human Gene Mapping 9 (1987).

Table 3

***TGFA* Allele and Genotype Frequencies**

Subjects Category (*N*)	Allele[a] Frequency		Genotype Frequency		
	C1	C2	C1C1	C1C2	C2C2
CL/P (78)	.865	.135	59 (76%)	17 (22%)	2 (2%)
Control (98)	.95	.05	89 (91%)	8 (8%)	1 (1%)
	A1	A2	A1A1	A1A2	A2A2
CL/P (77)	.05	.95	0 (0)	7 (9%)	70 (91%)
Control (96)	.13	.87	1 (1%)	23 (24%)	72 (75%)
	B1	B2	B1B1	B1B2	B2B2
CL/P (80)	.275	.725	6 (7.5%)	32 (40%)	42 (52.5%)
Control (101)	.27	.73	11 (11%)	32 (32%)	58 (57%)

[a] *TGFA*/*Taq*I: C1—3.0 kbp, C2—2.7 kbp; *TGFA*/*Bam*HI: A1—7.0 kbp, A2—4.0 kbp; *TGFA*/*Rsa*I: B1—1.5 kbp, B2—1.2 kbp.

the 12 test comparisons made (alpha = .05/12 = .0042). No significant difference was observed in frequency of the *TGFA Rsa*I RFLP alleles (χ^2 = 0.03, P = 0.4815) between cases and controls.

Haplotypes at the *TGFA* locus, consisting of the *Taq*I, *Bam*HI, and *Rsa*I RFLPs, were derived and compared between cases and controls (table 4). Detailed examination of the haplotype distributions showed that a significant portion of the difference between cases and controls was due to the overrepresentation of the C2A2B2 haplotype in cases. Forty-eight percent (χ^2 = 5.62) of the total Pearson χ^2 (11.64) was contributed by the C2A2B2 haplotype. A likelihood-ratio test indicated that the C2A2B2 haplotype was significantly overrepresented in cases as compared with controls (likelihood ratio χ^2 =5.737, P = 0.017). No other haplotype showed a significant individual difference between cases and controls. However, the C1A2B1 haplotype was also overrepresented in cases, but it contributed only 12% of the total Pearson χ^2. Similar results were observed when *TGFA* genotypes composed of the five observed haplotypes were compared (table 5). Of the 67 genotypes that could be derived for the cases, 16% were heterozygous and 3% were homozygous for the C2A2B2 haplotype. In the 80 control genotypes, 5% were C2A2B2 heterozygotes and 1% were homozygotes. The C2A2B2 genotypes accounted for 38% of the difference between cases and controls.

Subdividing the data by sex showed no significant difference in haplotype distribution between male and female cases (χ^2 = 2.11, P = 0.349) or controls (χ^2 = 1.30), P = 0.519). No significant difference was ob-

Table 4

***TGFA* Haplotype Data**

Haplotype[a]	No. of CL/P Chromosomes, By Family History			No. of Control Chromosomes
	Positive[b]	Negative	Total	
C2A2B2	10	5	15	6
C1A2B2	32	55	87	117
C1A2B1	14	15	29	25
C1A1B2	0	0	0	2
C1A1B1	2	2	4	11
Total	58	77	135	161

[a] Derived for each of an individual's two chromosomes, from alleles present from *Taq*I, *Bam*HI, and *Rsa*I digests.

[b] Refers to any relative identified as having CL/P.

Table 5

***TGFA* Genotype Data**

Genotype[a]	No. of CL/P Cases	No. of Controls
C2A2B2/C2A2B2	2	1
C1A2B2/C1A2B2	27	47
C1A2B1/C1A2B1	2	1
C1A1B1/C1A1B1	0	2
C1A2B2/C1A2B1	21	16
C1A2B1/C1A1B1	4	7
C2A2B2/C1A2B2	11	4
C1A2B2/C1A1B2	0	2
Total	67	80

[a] Derived by combining an individual's *TGFA* haplotypes (see table 3).

served in haplotype distribution ($\chi^2 = 4.50$, $P = 0.343$) when bilateral CL/P was compared with unilateral CL/P or with unilateral CL alone. However, a significant difference in haplotype distribution was observed with respect to family history of clefting ($\chi^2 = 6.42$, $P = 0.04$). Fifty-eight percent of the Pearson χ^2 resulted from the overrepresentation of the C2A2B2 haplotype in cases reporting any family history of CL/P. The haplotype distributions in cases with and without family histories were then independently compared with the control haplotype distributions. No significant difference in haplotype distribution was observed between cases without a family history and controls ($\chi^2 = 2.84$, $P = 0.417$). A highly significant difference was observed, however, between cases with a positive family history and controls ($\chi^2 = 13.71$, $P = 0.003$). As before, the difference in the C2A2B2 haplotype frequency contributed the largest component to the Pearson χ^2 (66%).

Discussion

To identify the major gene proposed by Chung et al. and Marazita et al. to be involved in a portion of CL/P cases in Caucasians, we considered two possible approaches. Linkage studies require a number of multigenerational families or multiple affected sib pairs, are most successful when the disorder under study is homogeneous, and do not imply causality between the marker allele being studied and the disease. Because of difficulties in obtaining large numbers of families with nonsyndromic CL/P, we used an association study to look for the prevalence, across a well-defined population, of a given disorder (CL/P) among individuals with different genetic markers. Such studies look for the disorder and a particular marker to be found together with a greater frequency than expected by chance alone, implying a causal relationship between the marker and the disorder in question. The classic paradigm for this are the studies that have shown the association of particular HLA antigens with a higher risk of developing diabetes mellitus.

Our study demonstrates a significant association between (1) the 2.7-kbp *Taq*I and the 4.0-kbp *Bam*HI fragments of the *TGFA* probe and (2) nonsyndromic CL/P. This association suggests that an abnormality in this gene may underlie a predisposition for clefting in some individuals or that an as yet unidentified clefting gene may be tightly linked to the *TGFA* locus.

The development of the secondary palate, or roof of the mouth, involves proliferation and differentiation of palatal epithelial cells as well as programmed cell death (Pratt et al. 1980). The occurrence of cell death within the medial epithelial lamina between the fusing palatal processes assists in the removal of the medial palatal epithelium. EGF/TGF-α and glucocorticoids are believed to regulate the proliferation and differentiation of palatal epithelial cells both in vitro and in vivo. Moreover, the continued presence of EGF inhibits the fusion process; TGF-α likely has similar effects. These latter biological studies suggest that mutations in the *TGFA* gene might contribute to the development of CL/P, especially for those mutations that affect the timing of the tissue-specific expression of this gene.

While currently no single *TGFA* allele, haplotype, or genotype is completely consistent with the Chung et al. (1986) model predictions, it is interesting to note that the frequency for the C2A2B2 haplotype in controls is .037, which is similar to the .035 predicted by the Chung et al. model. The homozygote frequency for this haplotype in cases, however, is significantly less than predicted by the autosomal recessive major-locus model of Chung et al. The Iowa population may have a higher proportion of cases not attributable to the major locus. Segregation analysis of pedigree data from our Iowa CL/P cases may determine the contribution that a major locus makes to clefting in this population.

It is also possible that more than one mutation that can cause CL/P has occurred at the *TGFA* locus. Examination of the haplotype distributions show that, in addition to the overrepresentation of the C2A2B2 haplotype, the C1A2B1 haplotype is also found in excess in cases. If one speculates that familial aggregation of CL/P is caused by homozygosity for either of these two individual haplotypes or by heterozygosity (C2A2B2/

C1A2B1), the proportion of affecteds attributable to *TGFA* variation is much more consistent with the predictions of the Chung et al. model.

As noted previously, segregation analysis suggested that only one-third of nonsyndromic CL/P cases would be explained by a single major gene. Other cases may be due to different gene(s), to environmental agents, or to the interaction between them. These cases would be expected to be distributed randomly with respect to the *TGFA* haplotypes.

The association reported here identifies a population of individuals genetically at-risk for CL/P. It should be noted that more accurate estimates of haplotype frequencies will require both adjustment for bias (resulting as a consequence of derivation by homozygosity) and larger sample sizes. The overall contribution of the *TGFA* locus to clefting in our population remains to be determined, and confirmation in other groups is important. Demonstration of linkage of the *TGFA* gene to the CL/P phenotype by using affected sib pairs would substantiate this association, as well as indicate that this locus is indeed a major genetic risk factor for this disorder.

The identification of additional RFLPs within the *TGFA* locus—and their analysis in the study population that we have described here—would facilitate the identification of susceptibility alleles, which could then be isolated. The characterization of mutations contributing to CL/P would provide important insight into the regulation of craniofacial development.

Acknowledgments

We are grateful to Nancy Cox for her review of statistical analysis, to Shirley Maxey for her help in patient ascertainment, and to the patients and their families who made this study possible. This work was supported by NIH grants DE05506 (H.H.A.), DE08556 (J.C.M.), and DE05837 (J.B. and D.R.V.). This work was presented in part at the 1988 meetings of the Society for Pediatric Research, Washington, DC.

References

Carter CO (1969) Genetics of common disorders. Br Med Bull 25:52–57

Castagnoli A, Maestri I, Bernardi F, Del Senno L (1987) PvuII RFLP inside the human estrogen receptor gene. Nucleic Acids Res 15:866

Chung CS, Bixler D, Watanabe T, Koguchi H, Fogh-Andersen P (1986) Segregation analysis of cleft lip with or without cleft palate: a comparison of Danish and Japanese data. Am J Hum Genet 39:603–611

Feinberg AP, Vogelstein B (1983) A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal Biochem 132:6–13

Ferguson MWJ (1987) Palate development: mechanisms and malformations. Ir J Med Sci 156:309–315

Fraser FC (1976) The multifactorial threshold concept—uses and misuses. Teratology 14:267–280

Hayward NK, Nancarrow DJ, Bell GI (1987) A TaqI polymorphism for the human transforming growth factor alpha gene (TGFA). Nucleic Acids Res 15:5503

Human Gene Mapping 9 (1987) Ninth International Workshop on Human Gene Mapping. Cytogenet Cell Genet 46:1–4

Marazita ML, Spence MA, Melnick M (1986) Major gene determination of liability to cleft lip with or without cleft palate: a multiracial view. J Craniofac Genet Dev Biol 2:89–97

Murray JC, Buetow KH, Bell GI (1986a) RFLPs for transforming growth factor alpha (TGFA) gene at 2p13. Nucleic Acids Res 14:5117

Murray JC, Dehaven CR, Bell GI (1986b) RFLPs for epidermal growth factor (EGF), a single copy sequence at 4q25–4q27. Nucleic Acids Res 14:5117

Murray JC, Smith RF, Ardinger HH, Weinberger CD (1987) RFLP for the glucocorticoid receptor (GRL) located at 5q11–5q13. Nucleic Acids Res 15:6765

Poncz M, Solowiejczyk D, Harpel B, Mory Y, Schwartz E, Surrey S (1982) Construction of human gene libraries from small amounts of peripheral blood: analysis of β-like globin genes. Hemoglobin 6:27–36

Pratt RM, Yoneda T, Silver MH, Salomon DS (1980) Involvement of glucocorticoids and epidermal growth factor in secondary palate development. In: Pratt RM and Christiansen RL (eds) Current trends in prenatal craniofacial development. Elsevier/North Holland, New York, pp 235–252

Smith RF, Ardinger HH, Murray JC (1987) Multiple RFLPs demonstrated for epidermal growth factor receptor (EGFR) on chromosome 7. Nucleic Acids Res 15:6764