Cleft Lip With or Without Cleft Palate in Shanghai, China: Evidence for an Autosomal Major Locus

Mary L. Marazita,* Dan-Ning Hu,† M. Anne Spence,‡ You-E Liu,† and Michael Melnick§

*Medical College of Virginia, Virginia Commonwealth University, Richmond; †Zhabei Eye Institute, Shanghai; ‡University of California, Los Angeles; §University of Southern California, Los Angeles

Summary

Orientals are at higher risk for cleft lip with or without cleft palate (CL ± P) than Caucasians or blacks. We collected demographic and family data to study factors contributing to the etiology of CL ± P in Shanghai. The birth incidence of nonsyndromic CL ± P (Shanghai 1980–87) was 1.11/1,000, with a male/female ratio of 1.42. Almost 2,000 nonsyndromic CL ± P probands were ascertained from individuals operated on during the years 1956–83 at surgical hospitals in Shanghai. Detailed family histories and medical examinations were obtained for the probands and all available family members. Genetic analyses of the probands' families were performed under the mixed model with major locus (ML) and multifactorial (MFT) components. The hypotheses of no familial transmission and of MFT alone could be rejected. Of the ML models, the autosomal recessive was significantly most likely and was assumed for testing three complex hypotheses: (1) ML and sporadics; (2) ML and MFT; (3) ML, MFT, and sporadics. None of the complex models were more likely than the ML alone model. In conclusion, the best-fitting, most parsimonious model for CL ± P in Shanghai was that of an autosomal recessive major locus.

Introduction

The etiology of nonsyndromic cleft lip with or without cleft palate (CL ± P) is still unclear. Early estimates of the genetic contribution to such clefts ranged from about 12%–20%, with the remainder attributed to environmental factors or gene-environment interactions (e.g., see Fogh-Andersen 1968, 1971; Ross and Johnston 1972, pp. 3–67). Estimates from more recent studies suggest that about 50% may be more realistic (Marazita et al. 1984; Chung et al. 1986).

There are significant racial differences in the incidence of clefting, with the highest rates in Orientals and Native Americans, intermediate rates in Caucasians, and lowest rates in blacks (Tanaka et al. 1969; Chung et al. 1974; Myrianthopoulos and Chung 1974; Koguchi 1975; Melnick et al. 1980; Hu et al. 1982; Marazita et al. 1986b; Melnick et al. 1986). These racial differences persist in Hawaii, where the

environment is relatively uniform among different races, and after removal of ascertainment biases (Chung et al. 1974). Comparative studies of the genetic etiology of oral/facial clefts among racial groups are of great interest and potential importance, since they could permit a determination of whether the incidence differences reflect underlying heterogeneity, or merely differences in gene frequency. The present study is a genetic analysis of almost 2,000 CL ± P proband families ascertained in Shanghai. Complex segregation analysis was utilized to test hypotheses regarding the genetic etiology of CL ± P in this study population.

Subjects and Methods

Study Population

Birth records for the years 1980–1987 from 10 hospitals in eight Shanghai city districts were used to estimate the incidence of CL ± P in Shanghai. Among the 250,372 live births studied, there were a total of 279 newborn infants with nonsyndromic CL ± P, for a birth incidence of 1.11/1,000 live births. There was no significant variation in incidence from year to year (χ^2 (7 df) = 10.03, $P > .10$). The male/female sex

ratio for cases was 1.42, as compared with a population ratio of 1.04. The sex ratio was greater for cleft lip (CL) (1.58) than for cleft lip with cleft palate (CL + P) (1.35). The unilateral/bilateral ratio for CL ± P was 2.86, with a ratio of 15.33 for CL cases versus 1.76 for CL + P. For unilateral CL ± P cases, the left/right ratio was 1.36, with a ratio of 1.71 for CL versus 1.14 for CL + P cases.

For genetic analysis, families were ascertained through almost 2,000 nonsyndromic CL ± P probands whose surgical treatments were performed at surgical hospitals in Shanghai. Extended family history information was obtained, and the status of affected family members was verified by field workers in Shanghai.

For the genetic analyses using the mixed-model approach (see Statistical Methods), the large multigenerational kindreds were broken into their component nuclear families, then the mode of ascertainment for each nuclear family was specified. The following types of nuclear families were formed, for a total of 2,255 nuclear families (9,828 individuals): (1) 1,952 nuclear families composed of the probands, their parents, and their siblings (identified by single ascertainment, with the probability of an affected individual becoming a proband, π, estimated at .01 by the method of Gladstien et al. (1978); (2) 138 families composed of probands, their spouses, and their children (complete ascertainment); and (3) 165 nuclear families with other (nonproband) affected members related to the probands (ascertainment through pointers). Nuclear families with no affected members contribute little to a segregation analysis and were therefore not included (we have verified this in an unpublished data set by analyzing the data with and without such "unaffected" nuclear families). According to this scheme, some probands are included in the data set twice—in one nuclear family as a parent and in another nuclear family as a child. To remove this bias, we conditioned the probabilities of children's phenotypes on the parental phenotypes.

The trait analyzed was "CL ± P or not"; that is, all forms of CL ± P were considered affected. The Shanghai birth incidence of CL ± P was 1.11/1,000 live births, with a male/female ratio of 1.42. Therefore, for purposes of the genetic analyses, the female-specific incidence was 0.92/1,000 and the male-specific incidence was 1.30/1,000.

Statistical Methods

Because clefts of the lip and palate have historically been said to follow a multifactorial/threshold (MFT)

pattern, the goal of the genetic segregation analysis of these data was to test the MFT and Mendelian major-locus (ML) hypotheses simultaneously. To this end, we analyzed the data under the unified mixed model (Morton and MacLean 1974; Lalouel and Morton 1981; Lalouel et al. 1983), which assumes that an individual's genotype is composed of a multifactorial component and a major-locus component (Lalouel et al. 1983). Further descriptions of the model and its underlying assumptions can be found in the work of Morton and MacLean (1974) and Lalouel et al. (1983). The parameters of interest for the present analysis are given in the Appendix. Likelihoods were calculated and maximum-likelihood estimates of the parameters were obtained using the computer program POINTER (Morton et al. 1983).

Recently, Iselius and Morton (1991) reported that non-Mendelian transmission probabilities are not correctly implemented in POINTER. Using POINTER for significance tests of any departures from Mendelian expectations for the transmission probabilities may not lead to correct conclusions. We therefore estimated the transmission probabilities by using the regressive model for segregation analysis programmed in the REGD module of SAGE, version 2.0 (Sorant et al. 1989; ©1989 by R.C. Elston, Inc.).

For hypothesis tests, parameter estimates and likelihoods were obtained under each model, with various restrictions. For example, the POINTER model with $d = t = q = 0$ corresponds to a hypothesis of no major locus. Hypothesis tests were based on the likelihood-ratio criterion that compares each restricted model to the general, unrestricted model. In addition, the Akaike information criterion (AIC; Akaike 1974) was calculated for each model. The AIC for any model equals $- 2(\log \text{likelihood}) + 2(\text{number of estimated parameters})$. The model with the smallest AIC is the most parsimonious of the best-fitting models for the data.

Results

Table 1 presents the results of complex segregation analysis of the data. The likelihood-ratio criterion was used for hypothesis tests, comparing the likelihood of each restricted hypothesis to that of the general, unrestricted hypothesis (1a in table 1).

For the most general hypothesis (1a in table 1: ML, MFT, and sporadics), two parameters converged to boundary values. Therefore the df available for the hypothesis tests were reduced. The hypothesis of no

Table 1

Results of Complex Segregation Analysis of CL ± P in the Families of Almost 2,000 Surgical Probands from Shanghai

HYPOTHESIS	VALUES OF PARAMETERS ± SE[a]						
	d	t	q	X	H^2	χ^2	AIC
1. Complex models:							
a. ML, MFT, and sporadics	[0]	4.24	.05	0[b]	0[b]	. . .	−15,233
b. ML and MFT	[0]	4.12	.05	. . .	0[b]	.03	−15,232
c. ML and sporadics	[0]	4.24 ± 8.8	.06 ± .01	.60 ± .10	. . .	.30	−15,232
2. No familial transmission							
($q = H^2 = 0$)	. . .	. . .	[0]	. . .	[0]	280.51	−14,959
3. ML ($H^2 = 0$)	[0]	2.30 ± .90	.05 ± .08	. . .	[0]	1.33	−15,236
4. MFT ($q = 0$)	. . .	. . .	[0]	. . .	.72 ± .05	57.95	−15,180

NOTE.—Numbers in square brackets represent parameters not estimated; these parameters were set to the value inside the brackets.
[a] See Appendix for descriptions of parameters.
[b] Parameter converged to a boundary value.

familial transmission could be clearly rejected (df = 3; $P \ll .001$). The hypothesis of MFT transmission could also be clearly rejected, (df = 2; $P < .001$).

Of the autosomal dominant and recessive ML models, the recessive model was significantly more likely, so recessive was assumed for the more complex models and is the only ML model presented in table 1. When τ_2 (see Appendix) was estimated, it was not significantly different from the Mendelian expected value of .5 ($\chi^2 = 2.2$, with 1 df; therefore $P > .1$), so τ_2 was fixed at its expected value for the more complex models. Because of a reported problem with the estimation of transmission probabilities in POINTER (Iselius and Morton 1991), REGD was also used to estimate the τ's under an autosomal recessive ML model. The likelihood of a hypothesis with the τ's fixed at their Mendelian expected values was compared with the likelihood when the τ's were estimated. There was no significant difference between the likelihoods, confirming the POINTER result that the transmission probabilities were consistent with their Mendelian expectations.

Evaluation of the remaining complex models led to the conclusion that the ML-alone model was the best-fitting model in the data: (1) The three complex models—ML and sporadics; ML and MFT; and ML, MFT, and sporadics—were all approximately equally likely (see χ^2 values in table 1). The ML-alone model had approximately the same likelihood as these three models. (2) When an attempt was made to estimate the parameters under the most general model, X and H^2 both converged to 0, leaving only the parameters of the major locus. (note that no standard errors could

be estimated for the parameters in the most complex models because some parameters converged to boundary values) (3) Furthermore, the ML alone model had the smallest AIC value.

In addition to the analysis of the entire data set, POINTER analyses were done on subsets of the data to determine whether there was any heterogeneity in the results and in parameter estimates that was based on the severity of the proband's phenotype. Two comparisons were made given the following divisions of the data: (1) families in which the proband had only CL versus CL + P; and (2) families in which the proband had a unilateral cleft versus a bilateral cleft. Each subset of the families was analyzed in the same way as the entire data set. For each of the two comparisons, there was no significant heterogeneity in either the results or the parameter estimates. Therefore the heterogeneity results are not presented in detail here.

From the above results, it can be concluded that the best-fitting model from this analysis of almost 2,000 Shanghai CL ± P families was that of an *autosomal recessive major locus*. It is the most parsimonious and has the smallest AIC value of the four models with approximately the same likelihood (all of which included an ML component), and, when all parameters were estimated jointly, parameters X and H^2 converged to 0.

Discussion

Clefts of the lip and palate have been depicted and documented in China for centuries. The earliest recorded surgical cleft-lip repair is described in the *Chin*

Annals, the official history of the Chin Dynasty in China, circa A.D. 390 (translation in Khoo 1966). Clefts of the lip and palate remain an important public health problem in China. A better understanding of the etiology of CL ± P in China would have an impact on the management of the large number of cases born each year. The results of our study of almost 2,000 modern families of Chinese probands with nonsyndromic CL ± P indicate that an autosomal major locus is sufficient to explain the data.

The genetic model most often proposed for nonsyndromic CL ± P in the early 1970s was that of MFT inheritance (e.g., see Carter 1976). While the MFT model could reasonably account for the epidemiological findings in CL ± P, published reports suggesting that the MFT model best explained the data did not include statistical tests of the MFT hypothesis and its predictions (Carter et al. 1982; Hu et al. 1982). Marazita et al. (1986a) reexamined the data of Carter et al. (1982) and found that the MFT model could be rejected in favor of a mixed model (i.e., a single ML with multifactorial modification). Melnick et al. (1986) reanalyzed the Chinese families studied by Hu et al. (1982) and also found that an MFT model could be rejected, with evidence in favor of a single recessive ML for CL ± P.

Large, well-designed family studies of CL ± P reported by Chung et al. (1974, 1986), Melnick et al. (1980), Demenais et al. (1984), Marazita et al. (1984), and Nemana et al. (1992) also failed to support the MFT model. Chung et al. (1974) and Demenais et al. (1984) could not discriminate between single-locus and MFT models for CL ± P in their respective Hawaiian and French data sets. Melnick et al. (1980), Marazita et al. (1984), and Chung et al. (1986) analyzed the data of about 2,000 Danish kindreds and found that the MFT model could be rejected and that the data were consistent with an ML in at least a portion of the kindreds. Nemana et al. (1992) obtained evidence for an ML alone in a data set of 331 CL ± P families from Madras. Hecht et al. (1991b) analyzed 79 CL ± P U.S. Caucasian families and concluded that the best explanation for the data was a dominant ML with reduced penetrance.

The emerging evidence for the importance of single genetic loci in the etiology of nonsyndromic CL ± P has led to recent linkage and association studies. Several groups of investigators are analyzing candidate loci, with interesting results. Eiberg et al. (1987) have reported a tentative linkage between the F13A1 blood clotting factor located distal to HLA on chromosome 6p24-p25 and a locus for oral/facial clefts (lod score = 3.66 at a male recombination fraction of 0 and at a female recombination fraction of .26) in families with apparent autosomal dominant transmission. An interesting feature of Eiberg et al.'s report was that the evidence suggested that CP and CL ± P were allelic.

Ardinger et al. (1988, 1989) have also reported evidence for an association between the locus for transforming growth factor alpha (TGFA) and a CL ± P locus. There is no particular reason to believe that F13A1 might itself be involved in clefting, serving only as a classical genetic marker in Eiberg et al.'s study. By contrast, the TGFA association is particularly interesting because TGFA and/or other growth factors may well be important in the processes leading to a cleft. TGFA is believed to be the embryonic form of epidermal growth factor, which is believed to regulate the proliferation and differentiation of palatal epithelial cells both in vitro and in vivo (Pratt et al. 1980). However, while TGFA was chosen as a candidate for the Ardinger et al. (1988, 1989) study because of its involvement in CP in the mouse, it was found to be associated with CL ± P in humans, an etiologically distinct malformation.

The TGFA association has since been duplicated in Australian-Caucasian (Chenevix-Trench et al. 1991) and English-Caucasian (Vintiner et al. 1991) samples. However, Qian et al. (1991) found no association in a French-Caucasian sample. Further, Vintiner et al. (1991) found no evidence of linkage with TGFA in British families, nor did Hecht et al. (1991a) in a linkage study of CL ± P in 12 U.S. Caucasian families. In an effort to detect single-base-pair changes within the TGFA locus, Shiang et al. (1991) began direct sequencing of TGFA in 20 affected individuals. With one-third of the locus sequenced, no definitive mutations have yet been identified.

Each of the above-mentioned association and linkage studies was performed in Caucasian samples. Now that there is statistical evidence of an ML in the Chinese study population, we have genetic mapping studies planned in order to seek confirmation of the putative ML identified in the current study, to test the reported linkages and associations, to attempt to discover the chromosomal location of clefting gene(s), and to assess genetic heterogeneity. The relative racial and environmental homogeneity in Shanghai, as well as the large sample size available, will make this population especially useful for mapping studies.

Acknowledgments

Appendix

Parameters of the Unified Model for Complex Segregation Analysis

d = degree of dominance at the major locus

t = displacement between homozygotes at the major locus

q = gene frequency

τ_1 = probability that an individual of type AA will transmit A

τ_2 = probability that an individual of type Aa will transmit A

τ_3 = probability that an individual of type aa will transmit A

H^2 = childhood heritability

x = proportion of sporadic cases

References

Akaike H (1974) A new look at the statistical model identification. IEEE Trans Automatic Control AC-19:716–723

Ardinger HH, Beutow KH, Bell GI, Bardach J, VanDemark DR, Murray JC (1989) Association of genetic variation of the transforming growth factor-alpha gene with cleft lip and palate. Am J Hum Genet 45:348–353

Ardinger HH, Beutow KH, Bell GI, Murray JC (1988) Excess of a particular TGFA haplotype in individuals with cleft lip ± cleft palate and a positive family history. Am J Hum Genet 43 [Suppl]: A77

Carter CO (1976) Genetics of common single malformations. Br Med Bull 32:21–26

Carter CO, Evans K, Coffey R, Roberts JAF, Buck A, Roberts MF (1982) A three generation family study of cleft lip with or without cleft palate. J Med Genet 19:416–432

Chenevix-Trench G, Jones K, Green A, Martin N (1991) Further evidence for an association between genetic variation in transforming growth alpha and cleft lip and palate. Am J Hum Genet 48:1012–1013

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

Chung CS, Ching GHS, Morton NE (1974) A genetic study of cleft lip and palate in Hawaii. II. Complex segregation analysis and genetic risks. Am J Hum Genet 26:177–188

Demenais F, Bonaiti-Pelie, Briard ML, Feingold J (1984) An epidemiological and genetic study of facial clefting in France. II. Segregation analysis. J Med Genet 21:436–440

Eiberg HD, Bixler D, Fogh-Andersen P, Conneally PM, Mohr J (1987) Suggestion of linkage of a major locus for nonsyndromic orofacial cleft with F13A and tentative assignment to chromosome 6. Clin Genet 32:129–132

Fogh-Andersen P (1968) Increasing incidence of facial clefts: genetically or non-genetically determined. In: Longacre JJ (ed) Craniofacial anomalies: pathogenesis and repair. Lippincott, Philadelphia, pp 27–29

——— (1971) Epidemiology and etiology of clefts. In: Bergsma D (ed) The third conference on the clinical delineation of birth defects, part 2: Orofacial structures. Birth Defects 70(7): 50–53

Gladstien K, Lange K, Spence MA (1978) A goodness-of-fit test for the polygenic threshold model: application to pyloric stenosis. Am J Med Genet 2:7–13

Hecht JT, Wang Y, Blanton SH, Michels VV, Daiger SP (1991a) Cleft lip and palate: no evidence of linkage to transforming growth factor alpha. Am J Hum Genet 49: 682–686

Hecht JT, Yang P, Michels VV, Buetow KH (1991b) Complex segregation analysis of nonsyndromic cleft lip and palate. Am J Hum Genet 49:674–681

Hu DN, Li JH, Chen HY, Chang HS, Wu BX, Lu ZK, Wang DZ, et al (1982) Genetics of cleft lip and cleft palate in China. Am J Hum Genet 34:999–1002

Iselius L, Morton NE (1991) Transmission probabilities are not correctly implemented in the computer program POINTER. Am J Hum Genet 49:459

Khoo B-C (1966) An ancient Chinese text on a cleft lip. Plast Reconstr Surg 38:89–91

Koguchi H (1975) Recurrence rate in offspring and siblings of patients with cleft lip and/or cleft palate. Jpn J Hum Genet 20:207–221

Lalouel JM, Morton NE (1981) Complex segregation analysis with pointers. Hum Hered 31:312–321

Lalouel JM, Rao DC, Morton NE, Elston RC (1983) A unified model for complex segregation analysis. Am J Hum Genet 35:816–826

Marazita ML, Goldstein AM, Smalley SL, Spence MA (1986a) Cleft lip with or without cleft palate: reanalysis of a three-generation family study from England. Genet Epidemiol 3:335–342

Marazita ML, Spence MA, Melnick MM (1984) Genetic analysis of cleft lip with or without cleft palate in Danish kindreds. Am J Med Genet 19:9–18

——— (1986b) Major gene determination of liability to cleft lip with or without cleft palate: a multiracial view. J Craniofac Genet Dev Biol [Suppl 2]: 89–97

Melnick M, Bixler D, Fogh-Andersen P, Conneally PM

(1980) Cleft lip ± cleft palate: an overview of the literature and an analysis of Danish cases born between 1941 and 1968. Am J Med Genet 6:83–97

Melnick M, Marazita ML, Hu DN (1986) Genetic analysis of cleft lip with or without cleft palate in Chinese kindreds. Am J Med Genet 21 [Suppl 2]: 183–190

Morton NE, MacLean CJ (1974) Analysis of family resemblance. III. Complex segregation analysis of quantitative traits. Am J Hum Genet 26:489–503

Morton NE, Rao DC, Lalouel JM (1983) Methods in genetic epidemiology. Karger, New York

Myrianthopoulos NC, Chung CS (1974) Congenital malformations in singletons: epidemiologic survey. Birth Defects 10(11): 1–58

Nemana LJ, Marazita ML, Melnick M (1992) A genetic analysis of cleft lip with or without cleft palate in Madras, India. Am J Med Genet 42:5–10

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

Qian JF, Stoll C, Feingold J, Sauvage P, May M (1991) No association of genetic variation of the transforming growth factor alpha with cleft-lip and palate. Am J Hum Genet 49 [Suppl]: A397

Ross RB, Johnston MC (1972) Cleft lip and palate. Williams & Wilkins, Baltimore

Shiang R, Lidral A, Ardinger H, Buetow KH, Murray JC (1991) Direct sequencing of TGFα DNA from patients with cleft lip and palate. Am J Hum Genet 49 [Suppl]: A418

Sorant AJM, Bonney GE, Elston RC (1989) Segregation analysis of a discrete trait under a class A regressive logistic model (REGD version 2.0). In: Elston R (ed) Statistical analysis for genetic epidemiology (S.A.G.E.) users manual. LSU Medical Center, New Orleans

Tanaka K, Fujino H, Fujita Y, Tashiro H, Sanui Y (1969) Cleft lip and palate: some evidence for the multifactorial trait and estimation of heritability based upon Japanese data. Jpn J Hum Genet 14:1–9

Vintiner J, Holder SE, Malcolm S, Winter RM (1991) Non-syndromic cleft lip and palate: association and linkage studies with the transforming growth factor-alpha gene. Am J Hum Genet 49 [Suppl]: A190