ORIGINAL ARTICLE

Interferon Regulatory Factor 6 (*IRF6*) Gene Variants and the Risk of Isolated Cleft Lip or Palate

Theresa M. Zucchero, B.S., Margaret E. Cooper, M.S., M.S.I.S.,
Brion S. Maher, Ph.D., Sandra Daack-Hirsch, R.N., M.S.N.,
Buena Nepomuceno, R.N., B.S.N., Lucilene Ribeiro, Ph.D., Diana Caprau, M.D.,
Kaare Christensen, M.D., Ph.D., Yasushi Suzuki, D.D.S.,
Junichiro Machida, D.D.S., Ph.D., Nagato Natsume, D.D.S., D.Med.Sc., Ph.D.,
Koh-Ichiro Yoshiura, M.D., Ph.D., Alexandre R. Vieira, D.D.S., Ph.D.,
Ieda M. Orioli, M.D., Ph.D., Eduardo E. Castilla, M.D., Ph.D., Lina Moreno, D.D.S.,
Mauricio Arcos-Burgos, M.D., Ph.D., Andrew C. Lidral, D.D.S., Ph.D.,
L. Leigh Field, Ph.D., You-e Liu, M.D., Ajit Ray, Ph.D., Toby H. Goldstein, B.S.,
Rebecca E. Schultz, B.S., Min Shi, M.S., Marla K. Johnson, B.S., B.S.E.,
Shinji Kondo, M.D., Ph.D., Brian C. Schutte, Ph.D., Mary L. Marazita, Ph.D.,
and Jeffrey C. Murray, M.D.

ABSTRACT

BACKGROUND

Cleft lip or palate (or the two in combination) is a common birth defect that results from a mixture of genetic and environmental factors. We searched for a specific genetic factor contributing to this complex trait by examining large numbers of affected patients and families and evaluating a specific candidate gene.

METHODS

We identified the gene that encodes interferon regulatory factor 6 (*IRF6*) as a candidate gene on the basis of its involvement in an autosomal dominant form of cleft lip and palate, Van der Woude's syndrome. A single-nucleotide polymorphism in this gene results in either a valine or an isoleucine at amino acid position 274 (V274I). We carried out transmission-disequilibrium testing for V274I in 8003 individual subjects in 1968 families derived from 10 populations with ancestry in Asia, Europe, and South America, haplotype and linkage analyses, and case–control analyses, and determined the risk of cleft lip or palate that is associated with genetic variation in *IRF6*.

RESULTS

Strong evidence of overtransmission of the valine (V) allele was found in the entire population data set (P<10⁻⁹); moreover, the results for some individual populations from South America and Asia were highly significant. Variation at *IRF6* was responsible for 12 percent of the genetic contribution to cleft lip or palate and tripled the risk of recurrence in families that had already had one affected child.

CONCLUSIONS

DNA-sequence variants associated with *IRF6* are major contributors to cleft lip, with or without cleft palate. The contribution of variants in single genes to cleft lip or palate is an important consideration in genetic counseling.

IRTH DEFECTS ARISE FROM THE INTER-play of multiple genetic and environmental factors. Although such complex traits are characterized by familial aggregation, recurrence rates within families are relatively low; the risk that an affected child will have a sibling who is also affected is typically less than 5 percent. Cleft lip, cleft palate, or the combination of the two is a common birth defect that varies in prevalence according to geographic origin, with populations of Asian and Amerindian ancestry having the highest rates and groups of African ancestry the lowest.[1] Isolated cleft lip or palate (i.e., a cleft occurring in the absence of any other structural or cognitive abnormalities) makes up about 70 percent of all disorders characterized by a cleft; the remaining 30 percent are accounted for by several hundred mendelian (autosomal dominant, autosomal recessive, or X-linked), chromosomal, teratogenic, and sporadic conditions that include other birth defects.

One gene contributing to isolated cleft lip or palate, *MSX1*, has been identified. It was initially targeted for investigation after a nonsense mutation was detected in the gene, which segregated with cleft lip or palate in a large family.[2] Approximately 2 percent of patients with cleft lip or palate have missense mutations in the coding sequence or in highly conserved regulatory elements of *MSX1*.[3] *MSX1* has also been implicated as a causative gene in cleft palate in case–control studies and by the appearance of cleft palate in Msx1-deficient mice.

One mendelian disorder that resembles isolated cleft lip or palate is Van der Woude's syndrome, an autosomal dominant clefting disorder in which pits in the lower lip are the only additional remarkable characteristic. The phenotype of Van der Woude's syndrome directly overlaps with that of isolated cleft lip or palate, in that clefts are typical and are accompanied by pits in the lower lip in only approximately 85 percent of cases of the syndrome.[4] Thus, 15 percent of cases of Van der Woude's syndrome may be clinically indistinguishable from isolated cleft lip or palate. We recently reported that mutations in the gene for interferon regulatory factor 6 (*IRF6*) cause Van der Woude's syndrome.[5] In searching the gene for mutations, we also identified a common polymorphic variant in which isoleucine is substituted for valine at amino acid position 274 (V274I) in the protein-binding domain of IRF6. Since the valine found at this site in IRF6 is strongly conserved among species, we hypothesized that this variant might affect gene function and contrib-

ute to cleft lip or palate. In the current study, we evaluated V274I and other polymorphisms in *IRF6* in a large number of affected subjects and their families from 10 populations around the world, as well as in control subjects, in order to determine whether they are associated with cleft lip and palate.

METHODS

SUBJECTS

The subjects we studied are summarized in Table 1. Overall, we studied 8003 persons; 6755 were members of families that included at least one subject with a cleft, whereas the remaining 1248 were from families without a history of clefting. The 1968 families with clefts had ancestry in East Asia (populations in Japan, Vietnam, China, and the Philippines), ancestry in South America (in Brazil, Colombia, and Latin America, in the Estudio Colaborativo Latino Americano de Malformaciones Congénitas [ECLAMC]), ancestry in Europe (Denmark and the United States [Iowa]), and ancestry in India. Seven of the 10 populations have been described previously, at least in part.[6-13] Persons with clefts, their family members, and unaffected controls from Japan, Vietnam, and Brazil were identified as part of our investigations of genetic and environmental causes of clefting. All persons with clefts were screened for the presence of associated anomalies or syndromes, and only those determined to have isolated cleft lip or palate were included in this study. The study was approved by the institutional review boards at the University of Iowa and the University of Pittsburgh, and written informed consent was obtained from each person included in the study. In most cases, whole blood was collected for genetic analysis and was processed with use of a DNA-isolation kit (QIAamp, Qiagen). When blood collection was not possible, DNA was purified from buccal swabs.

For the analysis of Filipino subjects, 184 Filipino mothers whose children were unaffected were used as controls for the 387 subjects with cleft lip or palate (probands). Case–control comparisons with unrelated controls were also conducted for the Indian population (95 affected subjects and 40 controls) and the Chinese population (160 affected subjects and 24 controls). In addition, a collection of 1064 DNA samples representing 51 populations with different ancestry (the CEPH Human Genome Diversity Cell Line Panel) was obtained from the Centre d'Étude du Polymorphisme Humain (CEPH),

Table 1. Summary of All Families Analyzed in Tests of Association and All Control Groups.*

Regional Group	Ancestral Origin	No. of Families	Family Size		Phenotype			
			Range	Mean ±SD	Cleft Lip	Cleft Lip and Palate	Cleft Palate	Unknown Type of Cleft
					no. of families with affected proband			
Affected populations								
East Asia	Japan	115	3–4	3.0±0.1	26	73	16	0
	Vietnam	175	3–4	3.0±0.1	50	109	16	0
	China	97	3–26	10.6±4.7	33	64	0	0
	Philippines	403	3–76	15.1±15.2	116	282	5	0
South America	Brazil	303	3	3	88	161	51	3
	Colombia	107	3–19	6.7±4.4	17	85	0	5
	ECLAMC	233	2	2	62	146	25	0
Europe	Denmark	239	3–8	4.3±0.9	81	105	53	0
	Iowa	246	3	3	69	109	68	0
India	India	50	4–38	14.2±6.5	36	14	0	0
Controls†								
CEPH diversity panel	Africa	156						
	Pakistan	200						
	China	183						
	South America‡	108						
East Asia	Philippines	238						
	China	24						
India	India	40						

* ECLAMC denotes Estudio Colaborativo Latino Americano de Malformaciones Congénitas, and CEPH Centre d'Étude du Polymorphisme Humain.
† Controls were all individual unrelated subjects.
‡ These subjects were from Amerindian ancestral populations.

for the purpose of observing allele frequencies worldwide.[14]

GENOTYPING

Determination of the genotype with respect to single-nucleotide polymorphisms was performed according to one of two protocols on an automatic sequence-detection system (ABI Prism 7900HT, Applied Biosystems). The V274I polymorphism was genotyped with an allele-specific kinetic polymerase chain reaction.[10] Genotyping for all other markers was performed using TaqMan analysis (Applied Biosystems). Additional single-nucleotide polymorphisms within the *IRF6* genomic region were found by sequencing of the gene, whereas other single-nucleotide polymorphisms in the region surrounding *IRF6* were chosen on the basis of location, with use of the Celera Discovery System online research tool (Applied Biosystems). Genotyping assays were obtained from Applied Biosystems, either through the Assay-on-Demand service, in the case of 6 previously identified polymorphisms, or through the Assay-by-Design service, in the case of 29 additional assays for which we provided sequences (Table 1 of the Supplementary Appendix, available with the full text of this article at www.nejm.org). Reactions were carried out with use of standard conditions, as specified by the manufacturer, and performed in duplicate in order to confirm the results. Genotyping for the family studies was carried out at the Center for Inherited Disease Research with use of the Marshfield Genetics version 9 screening set of microsatellite repeat markers (www.cidr.jhmi.edu).

SEQUENCING

We sequenced 23 kbp of the *IRF6* genomic region of 24 subjects — 20 Filipino subjects with cleft lip or palate who had two valine (V) alleles (indicating

high risk) at the V274I site, 2 Filipino subjects with cleft lip or palate who had two isoleucine (I) alleles (indicating low risk) at the V274I site, and 2 unaffected controls of European descent. We also sequenced the exons of 160 other persons (80 Filipinos and 80 Iowans) with isolated cleft lip or palate.

Noncoding genomic regions that control the expression of genes are increasingly implicated in causing complex disease, and these regulatory regions tend to have conserved DNA sequences among species. We therefore sequenced three regions consisting of 200 to 300 bp with homology to the mouse genome (where homology is defined as greater than 75 percent identity over at least 100 bp), located approximately 81 kbp, 103 kbp, and 117 kbp 3′ of *IRF6*. We sequenced these regions in a panel of persons with cleft lip or palate from the Philippines (113 subjects) and Iowa (93 subjects), 140 control subjects from the Philippines, and 96 samples in the CEPH diversity panel. All sequencing reactions were performed with use of a cycle-sequencing kit (ABI Prism BigDye Terminator, Applied Biosystems).

STATISTICAL ANALYSIS
Preliminary Analyses
We analyzed the inheritance of each *IRF6* marker in all families, using PedCheck software[15] to test for inconsistencies due to nonpaternity or other errors. For linkage analyses, allele frequencies were estimated from the unaffected persons in early generations of the study families (founders). The parameters of the genetic model were based on the results of segregation analyses of families with isolated cleft lip or palate of the Chinese families,[16] the Indian families,[17] and the Filipino families (unpublished data).

Tests of Hardy–Weinberg Equilibrium and Case–Control Comparisons
The results of transmission-disequilibrium analysis can be biased if there are deviations from Hardy–Weinberg equilibrium (where the expected numbers of individuals of different genotypes can be predicted by the frequency of individual variance). Chi-square tests were performed to assess the Hardy–Weinberg equilibrium of V274I in the founders and persons not in the line of descent, such as spouses, in the families in each study population and thus to explore the possibility of false positives.[18] Chi-square statistics were also used to assess significance in case–control comparisons.

Linkage Calculations
Of the 10 populations in which the genotype for V274I was determined, the Colombian, Chinese, Indian, and Filipino populations included extended families and were therefore appropriate for parametric linkage analyses. In three of these — the Chinese, Indian, and Filipino populations — the Marshfield Genetics version 9 screening-set markers on chromosome 1 had been genotyped, and therefore multipoint calculations based on the V274I data from this study were appropriate. We calculated two-point logarithms of the odds (lod) scores in the extended families, using the Linkage program[19] with recent updates to speed calculations.[20-22] We calculated multipoint lod-score statistics using the descent-graph approach in SimWalk2 software[23] and multipoint nonparametric-linkage statistics using Merlin software.[24]

Transmission-Disequilibrium Testing of Allelic Association
Alleles at each *IRF6* marker were tested for association with cleft lip or palate with the use of the Family Based Association Test software.[25-27] The ECLAMC sample comprised only mother–child pairs, and so for transmission-disequilibrium calculations pertaining to that sample, we applied the likelihood-ratio test of Weinberg,[28] under the assumption that the distribution of paternal alleles was the same as that of maternal alleles. We also used the transmission-disequilibrium test to examine transmission either to unaffected siblings in the nuclear families of affected children (in the Philippines) or to control children in control triads, consisting of an unaffected child and his or her parents (in Iowa), in order to investigate the possibility of transmission distortion.[29] The identification of transmission distortion among controls would suggest that there is overall segregation distortion (more of one genotype than predicted from the parental genotypes) at this locus, which might occur if certain allelic combinations were lethal during embryonic development and which would negate the importance of our findings in affected subjects. Finally, because several populations included large extended families, all transmission-disequilibrium analyses were repeated in a sample consisting only of the nuclear families of the probands in each of the extended families.

In order to derive a summary statistic for association with the V allele across populations, a random-effects meta-analysis model, as described by

DerSimonian and Laird,[30] was used to estimate the odds ratio for the presence of the associated allele within the nuclear families of the probands. Before pooling the data, we estimated Cochran's Q statistic, which indicates the degree of heterogeneity. There was no significant evidence of heterogeneity overall (Q=9.545, P=0.30) or in the regional population groups. A random-effects model was used because it includes components of variance both within and between studies. Moreover, because it generally yields a wider confidence interval than a fixed-effects model, the random-effects model is more conservative.[31]

Analyses of Single-Nucleotide Polymorphisms and Haplotype Association

Thirty-six single-nucleotide polymorphisms (9 within *IRF6*, 10 within 100 kbp 5′ of *IRF6*, and 17 within 200 kbp 3′ of the gene) were studied in 296 triads (consisting of an affected subject and his or her parents) from the Philippines. Genotype data were available for seven of these single-nucleotide polymorphisms in *IRF6* for 108 triads from Iowa and 184 triads from Denmark. Individual values for association with cleft lip or palate were calculated with use of the Family Based Association Test software for each of the single-nucleotide-polymorphism markers typed in the Filipino and European populations. Haplotype-based transmission-disequilibrium statistics were also calculated with the haplotype version of the test.[27,32] The expectation maximization algorithm was used to determine the haplotypes and to estimate haplotype frequencies; transmission distortion was then assessed.

Attributable Risk

We estimated the attributable risk for the associated *IRF6* allele — that is, the proportion of cases of cleft lip or palate in a population that can be attributed to the V allele. The attributable risk is a function of the relative risk (RR) and the probability of exposure given that a person has the disease (P[E|D]). We calculated the estimated attributable risk (AR) according to the formula AR={(P[E|D]) (RR–1)}÷RR, using the odds ratio as an estimate of the relative risk.

RESULTS

GENE STRUCTURE AND LOCALIZATION

The structure of the *IRF6* gene with its intron–exon boundaries, the flanking regions, and the location of each single-nucleotide polymorphism analyzed is shown in Figure 1.

ALLELE FREQUENCIES AND CASE–CONTROL COMPARISONS

Only the genotypes derived from the population in Brazil showed a significant deviation from Hardy–Weinberg equilibrium (P<0.001). Brazilian populations have a high degree of admixture owing to the presence of groups with different ancestral origins (South American, Indian, African, and European) that may account for this result.[13] The samples from the CEPH diversity panel were also genotyped at the V274I polymorphism (Table 2 of the Supplementary Appendix). No Africans in this panel had the isoleucine (I) allele. The V allele is present in the chimpanzee (*Pan troglodytes*), indicating that
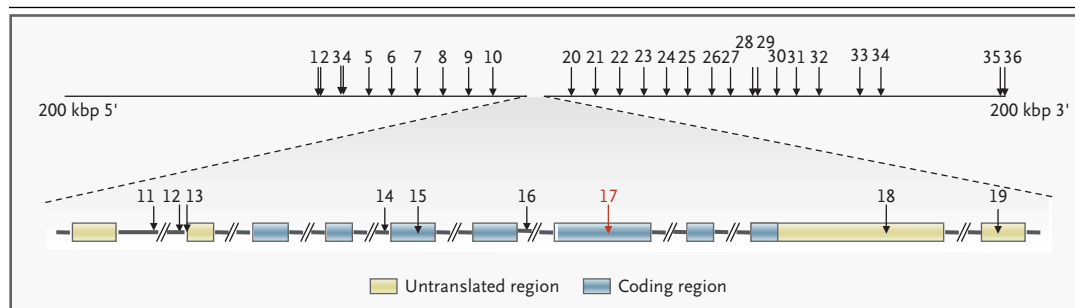
Figure 1. Diagram of the *IRF6* Gene Locus.

Blue denotes coding regions, and yellow untranslated regions. The V274I polymorphism (numbered 17) is shown in red. Black arrows and numbers indicate the locations of the single-nucleotide polymorphisms; these were used to determine genotypes in parent–child triads in the Filipino population, the population of European descent (Danish or Iowan), or both.

the V allele is probably the ancestral allele in *Homo sapiens*.

RESULTS OF LINKAGE ANALYSIS

The two-point lod scores (between two adjacent markers) for linkage with *IRF6* were negative or weakly positive in each of the populations at each value of the recombination fraction tested. The highest lod value was 0.636 (recombination fraction, 0.20) in the Indian families. Analysis of the Indian families gave a significant nonparametric-linkage result (P=0.004), and that of the Filipino families gave a result of borderline significance (P=0.06). The multipoint lod scores for the chromosome 1 microsatellite markers plus V274I were uniformly negative, whereas the multipoint heterogeneity lod values were weakly positive (the highest multipoint heterogeneity lod at *IRF6* was 0.382, in the Chinese families). Further details of the linkage analysis are available in Table 3 of the Supplementary Appendix. A recent meta-analysis of all genome scans in families containing persons with cleft lip or palate (including the Filipino, Chinese, Indian, and Colombian families included in our study) found significant evidence in favor of linkage to the region containing *IRF6* (P=0.02).[33]

RESULTS OF TRANSMISSION-DISEQUILIBRIUM TESTS OF ASSOCIATION

Transmission-disequilibrium calculations were performed for the entire data set as well as for data on subgroups defined according to the phenotype in affected probands. (Probands were characterized as having cleft lip alone, cleft lip and cleft palate, cleft lip alone or cleft lip with cleft palate, or cleft palate alone. These subgroups are summarized in Figure 1 of the Supplementary Appendix.) The populations were also classified for analysis on the basis of ancestral origin as Asian, South American, European, or Indian. (Detailed results of transmission-disequilibrium testing for each group are available in Table 4 of the Supplementary Appendix.) In the South American and Asian groups, the results were highly significant (P<0.001) for the association between cleft lip or palate and the V allele and also between cleft lip alone and the V allele, but not between cleft palate alone and the V allele. In the Indian and European groups, there was a nonsignificant trend toward a positive association between cleft lip with cleft palate and the V allele and between cleft lip alone and the V allele, but not between cleft

palate alone and the V allele. Members of the Indian and European populations were rarely heterozygous at the V274I site. Strong evidence of over-transmission of the V allele was found in the entire population data set (P<10^−9).

When we included only the nuclear families of probands from each of the extended families in the transmission-disequilibrium analysis, we saw the same pattern of results — that is, an association between the V allele and affected status (except for families in which the proband had cleft palate alone [data not shown]). The analysis of the ECLAMC population with the likelihood-ratio test showed no evidence of transmission distortion with respect to V274I. Figure 2 summarizes the odds ratios estimated in the nuclear families of the probands in each population and each regional population group.

There were no parent-specific differences in the patterns of association (data not shown). Transmission-disequilibrium testing carried out for samples of controls (made up of unaffected siblings from the probands' families) disclosed no significant transmission distortion (data not shown), except in the Chinese population, in which unaffected status was slightly associated with the I allele. That is, we found no evidence of an overall bias toward the V allele.

For 296 Filipino parent–child triads in which the child had cleft lip or palate, we performed transmission-disequilibrium analysis and calculations of the degree of overtransmission of V274I, 8 additional single-nucleotide polymorphisms within the *IRF6* gene, 10 single-nucleotide polymorphisms that flank *IRF6* at a distance of up to 80 kbp in the 5′ direction, and 17 single-nucleotide polymorphisms at distances from 10 to 200 kbp in the 3′ direction (Fig. 1). Nine of these markers (labeled 12, 16, 17 [V274I], 19, 22, 25, 27, 28, and 29 in Figure 1) were significantly associated with clefting (P<0.01). Figure 3A summarizes the results by showing the degree of overtransmission, expressed as a percentage above the expected 50 percent for an allele, of each of the 36 single-nucleotide polymorphisms analyzed. Figure 3B shows the P values obtained from the transmission-disequilibrium analysis for each of the 36 single-nucleotide polymorphisms, graphed according to physical location in relation to *IRF6*. (Numerical association data for all markers can be found in Table 5 of the Supplementary Appendix.)

RESULTS OF HAPLOTYPE-TRANSMISSION ANALYSIS

In addition to obtaining transmission-disequilibrium results for individual single-nucleotide polymorphisms, we carried out transmission-disequilibrium analysis using haplotype data. Haplotypes constructed from two or more single-nucleotide polymorphism loci can increase the information content (heterozygosity) of a locus and enhance the power of the study by generating a large number of families that can be analyzed. Figure 4 summarizes the results for all haplotypes with a frequency of more than 1 percent in either the Filipino population or the population of European ancestry (Danish or Iowan). Data were obtained from 296 Filipino parent–child triads, and haplotypes are shown for nine single-nucleotide polymorphisms that defined the haplotypes. The haplotype consisting of all the common alleles at each single-nucleotide polymorphism (estimated haplotype frequency, 46 percent) had the most significant transmission distortion (P<0.001). The European haplotype results, obtained in 108 Iowan and 184 Danish triads, are based on the four single-nucleotide polymorphisms that defined these haplotypes. The haplotype consisting of the common allele at each single-nucleotide polymorphism (frequency, 53.8 percent in the Iowan population and 54.0 percent in the Danish population) was significantly associated with clefting (P=0.04 in the Iowan population, P=0.006 in the Danish population).

Figure 2 shows the odds ratios for the associated haplotypes in the samples from the Philippines, Iowa, and Denmark. Linkage disequilibrium was assessed for all the markers and is presented graphically and in numerical form in Figure 2 and Table 6 of the Supplementary Appendix. These results suggest that there is a long block of linkage disequilibrium extending from about 40 kbp in the 5′ direction from *IRF6* to at least 100 kbp 3′ of the gene. This is consistent with a block of linkage disequilibrium seen around *IRF6* in European samples used to construct the haplotype map HapMap, which describes the common patterns of genetic variation in humans (www.hapmap.org).
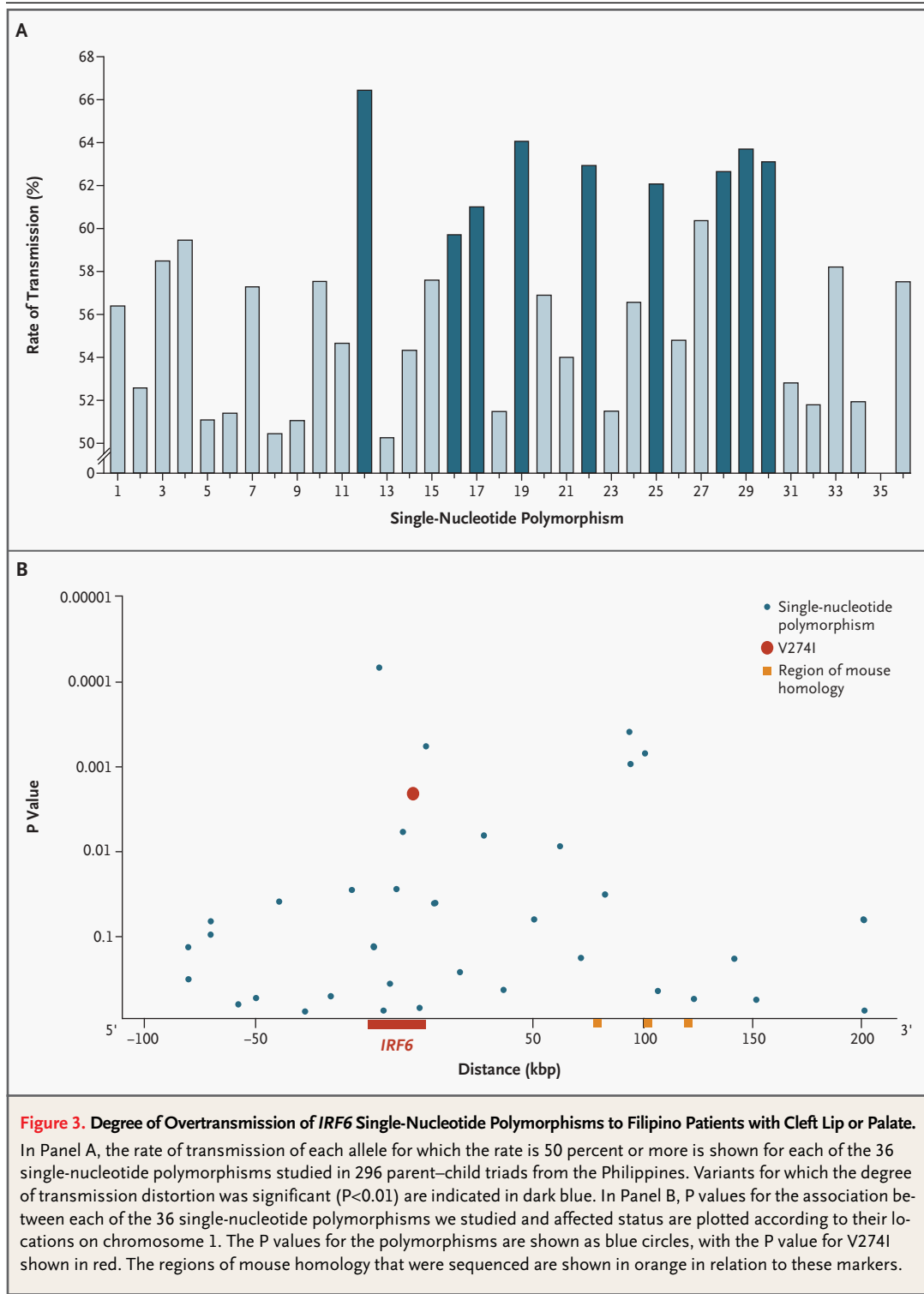
For the comparison of the Filipino subjects with cleft lip or palate with population-based controls, the estimated attributable risk was 11.6 percent. This attributable risk is based on the assumption that the risk factor is causal and is not correlated with other risk factors, so it should be interpreted cautiously.
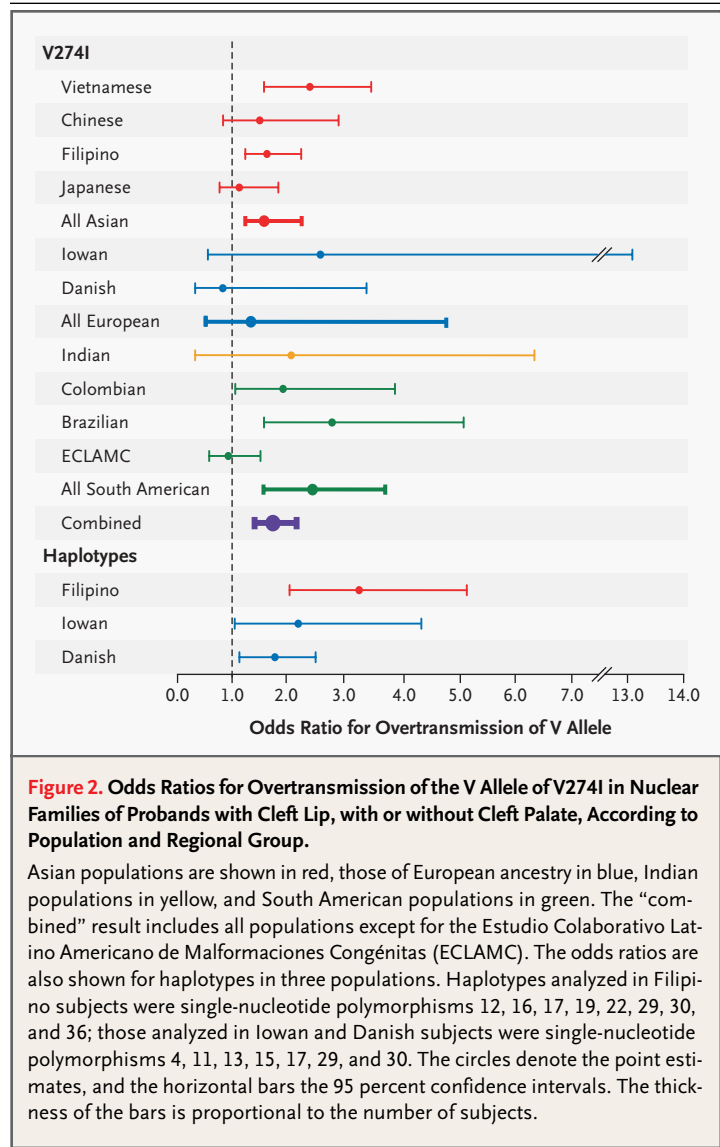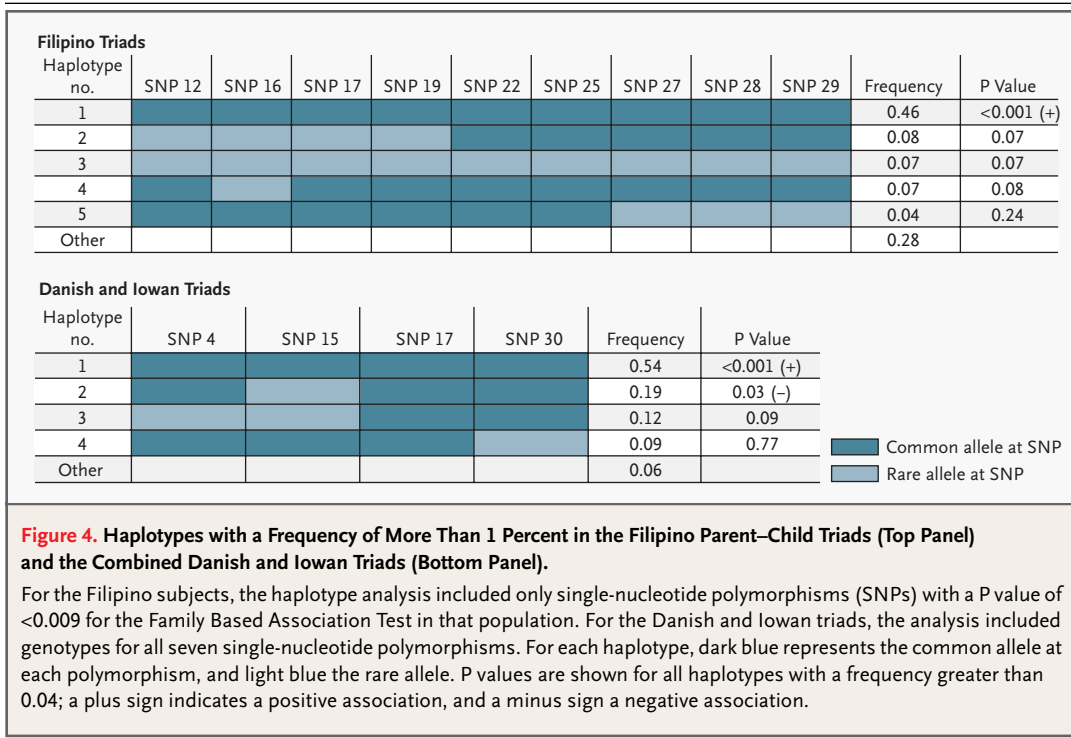

Figure 2. Odds Ratios for Overtransmission of the V Allele of V274I in Nuclear Families of Probands with Cleft Lip, with or without Cleft Palate, According to Population and Regional Group.

Asian populations are shown in red, those of European ancestry in blue, Indian populations in yellow, and South American populations in green. The "combined" result includes all populations except for the Estudio Colaborativo Latino Americano de Malformaciones Congénitas (ECLAMC). The odds ratios are also shown for haplotypes in three populations. Haplotypes analyzed in Filipino subjects were single-nucleotide polymorphisms 12, 16, 17, 19, 22, 29, 30, and 36; those analyzed in Iowan and Danish subjects were single-nucleotide polymorphisms 4, 11, 13, 15, 17, 29, and 30. The circles denote the point estimates, and the horizontal bars the 95 percent confidence intervals. The thickness of the bars is proportional to the number of subjects.

To characterize further the effect of the V allele associated with clefting at V274I, we performed a genotype transmission-disequilibrium analysis using the Family Based Association Test software for the entire data set. V/V homozygosity was significantly associated with clefting (P<0.001), whereas the V/I and I/I genotypes were negatively associated with clefting (P<0.005). This pattern is consistent with a recessive effect of the V allele. Such a recessive effect was further confirmed by the observation of a significant difference in the frequency distributions of genotypes among probands as compared with unaffected subjects (P<0.001); the frequency of the V/V genotype was increased,

Figure 3. Degree of Overtransmission of *IRF6* Single-Nucleotide Polymorphisms to Filipino Patients with Cleft Lip or Palate.

In Panel A, the rate of transmission of each allele for which the rate is 50 percent or more is shown for each of the 36 single-nucleotide polymorphisms studied in 296 parent–child triads from the Philippines. Variants for which the degree of transmission distortion was significant ($P<0.01$) are indicated in dark blue. In Panel B, P values for the association between each of the 36 single-nucleotide polymorphisms we studied and affected status are plotted according to their locations on chromosome 1. The P values for the polymorphisms are shown as blue circles, with the P value for V274I shown in red. The regions of mouse homology that were sequenced are shown in orange in relation to these markers.

Figure 4. Haplotypes with a Frequency of More Than 1 Percent in the Filipino Parent–Child Triads (Top Panel) and the Combined Danish and Iowan Triads (Bottom Panel).

For the Filipino subjects, the haplotype analysis included only single-nucleotide polymorphisms (SNPs) with a P value of <0.009 for the Family Based Association Test in that population. For the Danish and Iowan triads, the analysis included genotypes for all seven single-nucleotide polymorphisms. For each haplotype, dark blue represents the common allele at each polymorphism, and light blue the rare allele. P values are shown for all haplotypes with a frequency greater than 0.04; a plus sign indicates a positive association, and a minus sign a negative association.

and the frequencies of the the V/I and I/I genotypes were reduced, in probands as compared with unaffected controls.

We therefore calculated the rate of recurrence among siblings of an affected child according to parental genotype, using nuclear families from the entire data set in which there were one or more affected children and in which both parents' genotypes for V274I in *IRF6* were known (1493 nuclear families). We divided the parents into two groups according to whether they could have a child who was homozygous for the V allele at *IRF6* (these pairings were V/V–V/V, V/V–V/I, and V/I–V/I, accounting for 1316 families) or could not have such a child (V/I–I/I, I/I–I/I, and V/V–I/I, accounting for 177 families). Within each group, we calculated the proportion of families with one affected child or with more than one affected child. Because this was not a population-based sample, these are not true empiric relative-risk estimates, but the results may be generalized to families with a positive family history for clefting. For the first group (those whose genotypes could result in an affected child), the rate of recurrence was 9.0 percent, as compared with 5.1 percent in the second group. Although this difference was not statistically significant (P=0.08), the increased relative risk in the first group was more than three times the empirical, population-based relative risk of 2.4 percent estimated for cleft lip or palate in the Philippines.[34]

RESULTS OF SEQUENCING

The 10 exons of *IRF6* sequenced in 160 subjects with isolated cleft lip or palate had no missense, nonsense, or frame shift mutations, suggesting that persons with point mutations in *IRF6* are rarely misidentified as having isolated cleft lip or palate, rather than Van der Woude's syndrome. Sequencing of the 23 kbp that includes the entire known intron–exon structure of *IRF6* in 24 subjects (22 with cleft lip or palate and 2 without it) revealed 58 variants (sequencing results available on request from the authors), none of which seem likely to be causative. Analysis of three blocks of sequence conserved between the human and mouse genomes and downstream of *IRF6* did not reveal any changes within the most conserved regions of the blocks.

DISCUSSION

Our study of 10 populations (comprising a total of 1968 families) with isolated cleft lip or palate showed highly significant transmission disequilibrium for the V274I variant of the *IRF6* gene. Three

possible sources of bias can lead to apparent over-transmission of an allele in transmission-disequilibrium analyses of two-allele markers: departures from Hardy–Weinberg equilibrium,[18] genotyping errors, and segregation distortion.[29] In the 10 populations we investigated, only the genotypes of the affected subjects from Brazil showed any significant departure from Hardy–Weinberg equilibrium — a finding that may be explained by the relatively recent admixture of persons from different ancestral groups within that population. If there had been important genotyping errors or segregation distortion, then overall transmission distortion would be expected; this possibility was excluded by the normal transmission seen in the analysis of the unaffected children.

Single-nucleotide polymorphisms within and flanking the *IRF6* gene were analyzed to determine whether the V allele at position 274 had the maximal effect in causing cleft lip or palate, or whether other variants in linkage disequilibrium with the V allele might be more important. Alleles defined by multiple other single-nucleotide polymorphisms also showed significant transmission distortion. The most significant P value and the greatest degree of overtransmission were for a variant located in the first intron, a distance of 4 bp from the splice site, within a noncoding exon that is absent in rodents. Whether this variant itself, or another in association with it, is of functional importance is not yet known.

Additional support for the assumption that the V allele does not itself cause the defect comes from the strong disequilibrium seen between particular *IRF6* haplotypes and cleft lip or palate in the European populations, in which the I variant allele is rare. This suggests either that the V allele is not causal, or that it may share causality with variants at other sites within or near *IRF6* that show stronger transmission distortion than does the V allele. The linkage disequilibrium extends from 40 kbp in the 5′ direction to 100 kbp 3′ of *IRF6*; this location is consistent with an earlier report showing linkage disequilibrium at D1S205, which lies approximately 135 kbp 5′ of *IRF6*.[35] It is possible that more than one variant might contribute to this effect and that these variants may be different (or appear in different proportions) in the several ancestral populations we studied. In addition, a specific combination of variants on a single chromosome may be required for a person to exhibit biologic effect (cleft lip or palate).

A recent meta-analysis of all genome scans of subjects with cleft lip or palate, including the Filipino, Chinese, Indian, and Colombian families described here, found significant evidence of linkage to the region containing *IRF6*, although none of the individual studies showed evidence of linkage.[33] This suggests that even large family linkage studies may fail to find evidence of a genetic effect indicated by candidate-gene analysis and linkage disequilibrium. Similarly, this effect was found only in transmission-disequilibrium analyses and not in case–control comparisons, although there was a borderline trend toward significance in the case–control studies in some populations.

Analysis of the risk of recurrence suggests that three to six major genetic loci may contribute to clefting.[36] For *IRF6*, we found that there is an attributable risk of cleft lip or palate of about 12 percent, suggesting that this gene plays a substantial role in the causation of such defects. Our finding that the risk of recurrence is 9 percent among siblings in families with a history of cleft lip or palate and a child who could have inherited the common risk allele (as was the case for 88 percent of the families in this study) suggests the possible importance of information on V274I in genetic counseling. This is more than three times the 2.4 percent rate of recurrence found in a population-based study of cleft lip or palate in the Philippines.[34] If our results are confirmed, the *IRF6* genotype could be used to refine estimates of the risk of recurrence of this common disorder in genetic counseling.

Two genes, *MSX1* and *IRF6*, now seem to have a measurable role in the causation of cleft lip or palate. Mutations in other genes resulting in syndromes that include clefts (*TBX22*, *P63*, and *FGFR1*) will be identified in some persons who have only the cleft component of the phenotype. It will soon be practical to consider comprehensive sequence analysis or haplotype analysis of these genes, or both, in cases of isolated clefting in the absence of a family history of the associated features that can provide a syndromic diagnosis. Finding a gene-specific mutation or the associated *IRF6* haplotype could then raise the estimated familial risk of recurrence from the empirical value of 3 to 5 percent currently in use for a subsequent sibling of an affected child to much higher values.

We have demonstrated that informed candidate-gene selection can identify specific variants with a role in complex traits that may be missed by genome-wide linkage scanning or case–control

analysis. Isolated clefts are also associated with an overall lifetime increase in the risk of premature death from all causes, and haplotype associations might confer some specificity to these risks.[37] Direct identification of genes can improve genetic counseling, assist in the identification of new genetic and environmental causes of syndromes, and provide options for treatment or prevention, if the associated haplotype is correlated with efficacy.

REFERENCES

1. Mossey PA, Little J. Epidemiology of oral clefts: an international perspective. In: Wyszynski DF, ed. Cleft lip and palate: from origin to treatment. Oxford, England: Oxford University Press, 2002:127-58.
2. van den Boogaard MJ, Dorland M, Beemer FA, van Amstel HK. MSX1 mutation is associated with orofacial clefting and tooth agenesis in humans. Nat Genet 2000;24:342-3. [Erratum, Nat Genet 2000;25:125.]
3. Jezewski PA, Vieira AR, Nishimura C, et al. Complete sequencing shows a role for MSX1 in non-syndromic cleft lip and palate. J Med Genet 2003;40:399-407.
4. Burdick AB, Bixler D, Puckett CL. Genetic analysis in families with van der Woude syndrome. J Craniofac Genet Dev Biol 1985;5:181-208.
5. Kondo S, Schutte BC, Richardson RJ, et al. Mutations in IRF6 cause Van der Woude and popliteal pterygium syndromes. Nat Genet 2002;32:285-9.
6. Marazita ML, Field LL, Cooper ME, et al. Genome scan for loci involved in cleft lip with or without cleft palate, in Chinese multiplex families. Am J Hum Genet 2002;71:349-64.
7. Marazita ML, Field LL, Cooper ME, et al. Nonsyndromic cleft lip with or without cleft palate in China: assessment of candidate regions. Cleft Palate Craniofac J 2002;39:149-56.
8. Schultz RE, Cooper ME, Daack-Hirsch S, et al. Targeted scan of fifteen regions for nonsyndromic cleft lip and palate in Filipino families. Am J Med Genet 2004;125A:17-22.
9. Moreno LM, Arcos-Burgos M, Marazita ML, et al. Genetic analysis of candidate loci in non-syndromic cleft lip families from Antioquia-Colombia and Ohio. Am J Med Genet 2004;125A:135-44.
10. Shi M, Caprau D, Romitti P, Christensen K, Murray JC. Genotype frequencies and linkage disequilibrium in the CEPH human diversity panel for variants in folate pathway genes *MTHFR, MTHFD, MTRR, RFC1,* and *GCP2.* Birth Defects Res Part A Clin Mol Teratol 2003;67:545-9.
11. Romitti PA, Lidral AC, Munger RG, Daack-Hirsch S, Burns TL, Murray JC. Candidate genes for nonsyndromic cleft lip and palate and maternal cigarette smoking and alcohol consumption: evaluation of genotype-environment interactions from a population-based case-control study of orofacial clefts. Teratology 1999;59:39-50.
12. Field LL, Ray AK, Marazita ML. Transforming growth factor alpha: a modifying locus for nonsyndromic cleft lip with or without cleft palate? Eur J Hum Genet 1994;2:159-65.
13. Vieira AR, Orioli IM, Castilla EE, Cooper ME, Marazita ML, Murray JC. MSX1 and TGFB3 contribute to clefting in South America. J Dent Res 2003;82:289-92.
14. Cann HM, de Toma C, Cazes L, et al. A human genome diversity cell line panel. Science 2002;296:261-2.
15. O'Connell JR, Weeks DE. PedCheck: a program for identification of genotype incompatibilities in linkage analysis. Am J Hum Genet 1998;63:259-66.
16. Marazita ML, Hu DN, Spence MA, Liu YE, Melnick M. Cleft lip with or without cleft palate in Shanghai, China: evidence for an autosomal major locus. Am J Hum Genet 1992;51:648-53.
17. Ray AK, Field LL, Marazita ML. Nonsyndromic cleft lip with or without cleft palate in West Bengal, India: evidence for an autosomal major locus. Am J Hum Genet 1993;52:1006-11.
18. Schaid DJ, Jacobsen SJ. Biased tests of association: comparisons of allele frequencies when departing from Hardy-Weinberg proportions. Am J Epidemiol 1999;149:706-11.
19. Elston RC, Stewart J. A general model for the genetic analysis of pedigree data. Hum Hered 1971;21:523-42.
20. Cottingham RW Jr, Idury RM, Schaffer AA. Faster sequential genetic linkage computations. Am J Hum Genet 1993;53:252-63.
21. O'Connell JR, Weeks DE. The VITESSE algorithm for rapid exact multilocus linkage analysis via genotype set-recoding and fuzzy inheritance. Nat Genet 1995;11:402-8.
22. Terwilliger JD, Ott J. Handbook of human genetic linkage. Baltimore: Johns Hopkins University Press, 1994.
23. Sobel E, Lange K. Descent graphs in pedigree analysis: applications to haplotyping, location scores, and marker-sharing statistics. Am J Hum Genet 1996;58:1323-37.
24. Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin — rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet 2002;30:97-101.
25. Horvath S, Xu X, Laird NM. The family based association test method: strategies for studying general genotype–phenotype associations. Eur J Hum Genet 2001;9:301-6.
26. Laird NM, Horvath S, Xu X. Implementing a unified approach to family-based tests of association. Genet Epidemiol 2000;19:Suppl 1:S36-S42.
27. Rabinowitz D, Laird N. A unified approach to adjusting association tests for population admixture with arbitrary pedigree structure and arbitrary missing marker information. Hum Hered 2000;50:211-23.
28. Weinberg CR. Allowing for missing parents in genetic studies of case-parent triads. Am J Hum Genet 1999;64:1186-93.
29. Scott L, Rogus JJ. Using unaffected child trios to test for transmission distortion. Genet Epidemiol 2000;19:381-94.
30. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7:177-88.
31. Berlin JA, Laird NM, Sacks HS, Chalmers TC. A comparison of statistical methods for combining event rates from clinical trials. Stat Med 1989;8:141-51.
32. Horvath S, Xu X, Lake SL, Silverman EK,

Weiss ST, Laird NM. Family-based tests for associating haplotypes with general phenotype data: application to asthma genetics. Genet Epidemiol 2004;26:61-9.

33. Marazita ML, Murray JC, Lidral AC, et al. Meta-analysis of 13 genome scans reveals multiple cleft lip/palate genes with novel loci on 9q21 and 2q32-35. Am J Hum Genet 2004;75:161-73.

34. Murray JC, Daack-Hirsch S, Buetow KH, et al. Clinical and epidemiologic studies of cleft lip and palate in the Philippines. Cleft Palate Craniofac J 1997;34:7-10.

35. Houdayer C, Bonaiti-Pellie C, Erguy C, et al. Possible relationship between the van der Woude syndrome (vWS) locus and non-syndromic cleft lip with or without cleft palate (NSCL/P). Am J Med Genet 2001;104:86-92.

36. Schliekelman P, Slatkin M. Multiplex relative risk and estimation of the number of loci underlying an inherited disease. Am J Hum Genet 2002;71:1369-85.

37. Christensen K, Juel K, Herskind AM, Murray JC. Long term follow up study of survival associated with cleft lip and palate at birth. BMJ 2004;328:1405-8.

Copyright © 2004 Massachusetts Medical Society.