RESEARCH REPORTS

Clinical

R.I. Werneck¹, F.P. Lázaro¹, A. Cobat^{3,4},
A.V. Grant^{3,4}, M.B. Xavier², L. Abel^{3,4},
A. Alcaïs^{3,4}, P.C. Trevilatto¹,
and M.T. Mira^{1*}

¹Core for Advanced Molecular Investigation, Graduate Program in Health Sciences, Center for Biological and Health Sciences, Pontifical Catholic University of Paraná, Imaculada Conceição St. 1155 – CCBS/PPGCS, Prado Velho, CEP: 80215-901, Curitiba, Paraná, Brazil; ²Tropical Medicine Core, Federal University of Pará, Belém, Brazil; ³Laboratoire de Génétique Humaine des Maladies Infectieuses, Institut National de la Santé et de la Recherche Médicale, INSERM U550, Paris, France; and ⁴Université Paris René Descartes, Faculté Médecine Necker, Paris, France; *corresponding author, m.mira@pucpr.br

J Dent Res 90(6):735-739, 2011

ABSTRACT

Despite recent advances revealing genetic factors influencing caries susceptibility, questions regarding the model of inheritance involved are yet to be addressed. We conducted a Complex Segregation Analysis on decayed teeth in a sample of homogeneous, isolated families recruited from the Brazilian Amazon. A dominant, major gene effect controlling resistance to phenotype was detected. The frequency of the resistance allele “A” was 0.63; mean numbers of decayed teeth were 1.53 and 9.53 for genotypes AA/AB and BB, respectively. These results represent a step toward a description of the exact nature of the genetic risk factors controlling human susceptibility to caries.

KEY WORDS: caries, genetics, genetic epidemiology, susceptibility, inheritance.

A Major Gene Effect Controls Resistance to Caries

INTRODUCTION

Caries, the most common oral condition and the most prevalent infectious, non-contagious disease in the world, is a chronic, multifactorial disease that exerts enormous impact on public health systems of industrialized and developing countries (Petersen, 2003; Fejerskov, 2004). Caries is an important cause of dental loss and dental pain, both conditions associated with impaired performance in school and absenteeism at work (Petersen, 2003; Fejerskov, 2004), ultimately leading to a decrease in quality of life (Petersen, 2003). Even though reports have shown that the Decayed, Missing and Filled Teeth (DMFT) index, commonly used as an estimate of caries, has been decreasing over the past few years in developed and developing countries, caries continues to affect 60 to 90% of children at school age and the majority of adults (Petersen, 2003).

It is widely accepted that the occurrence of caries depends on environmental and host-related factors (Keyes, 1960; Evans *et al.*, 1993; Featherstone, 2004; Pine, 2005; Antunes and Peres, 2006). When the biofilm is exposed to highly fermentable carbohydrates, cariogenic bacteria like *Streptococcus mutans*, *Streptococcus sobrinus*, and some species of *Lactobacillus* (Keyes, 1960) are selected. Continuous exposure to acids produced by these bacteria, associated with a limited buffering capacity of the host, leads to dental decalcification (Featherstone, 2004). The process is modified by environmental factors, such as oral hygiene, fluoride exposure, as well as socio-economic status (SES), gender (Lukacs, 2011), ethnicity, and age (Evans *et al.*, 1993; Antunes and Peres, 2006). Of note, dental caries, early dental loss, and edentulism seem to concentrate in some groups of individuals (Petersen, 2003).

Although previous studies (Deeley *et al.*, 2008; Patir *et al.*, 2008; Ozturk *et al.*, 2010; reviewed in Werneck *et al.*, 2010) point to the existence of a genetic component controlling host susceptibility to caries, no description of the genetic model involved, as provided by Complex Segregation Analyses, has been produced to date. The approach enables the identification of a major gene effect (that is, an effect important enough to be distinguished from other susceptibility genes, but not necessarily caused by one single gene) and the estimation of the allele frequency and penetrance of the deleterious allele, among other parameters. Here we present the results of a Complex Segregation Analysis for caries performed in a sample of the Colony of Santo Antônio do

DOI: 10.1177/0022034510397614

Received September 2, 2010; Last revision November 20, 2010; Accepted November 22, 2010

A supplemental appendix to this article is published electronically only at <http://jdr.sagepub.com/supplemental>.

© International & American Associations for Dental Research

Prata (the Prata Colony), a community geographically isolated in the Amazonian state of Pará, north of Brazil. The population of the Prata Colony is highly exposed to caries susceptibility factors and shares very homogenous non-genetic variables (Lázaro *et al.*, 2010).

MATERIALS & METHODS

Study Population and Enrollment Strategy

The study was approved by the Research Ethics Committee of the Pontifical Catholic University of Paraná. All families included in the study were recruited from the Prata Colony, a former leper colony created in the early 1920s with the objective to isolate individuals affected with leprosy. Isolation was compulsory until 1962; however, the population of the colony remains isolated today, probably due to the strong stigma associated with leprosy, a disease still highly prevalent within the community. Previous assessment indicated very homogenous environmental and socio-economic variables and predominance of a mixed ethnic group, as reported elsewhere (Lázaro *et al.*, 2010). Family recruitment was performed by a systematic approach intended to reduce ascertainment bias. First, one household was randomly selected, and all members were invited to participate. Upon agreement, all individuals from this first nuclear family were asked to report whether there was any relative living in the colony; if so, these relatives and their families were also contacted and included in the study. The procedure was repeated until no additional relatives belonging to this first extended family were reported. Then, the next nuclear family was selected, counting two households to the left from the first included nuclear family and subjected to the same enrollment strategy. The same was applied until complete recruitment of the population sample was attained.

Phenotype Definition and Epidemiological Data Collection

Data collection was composed of a structured interview and a clinical examination, both performed by a single investigator (R.I.W.). Prior to data collection, all individuals were asked to sign an informed consent agreeing to participate in the study. The parents provided consent regarding individuals under 18 years old. The examination was conducted in the field, with natural light, tongue depressor, and gauze. A case of caries was defined according to World Health Organization guidelines (WHO, 1997). Information regarding demographic characteristics (gender, age, ethnicity), SES (socio-economic status, educational level), oral health (dietary habits, brushing habits, frequency of dental appointments), and clinical evaluation (decayed teeth, gingivitis, plaque, and leprosy status) was obtained and analyzed prior to the Complex Segregation Analysis.

Statistical Methods

To test formally for intra-observer repeatability, we re-collected data from 25 individuals during a second, independent expedition to the colony, following the same original conditions. Study participants were selected randomly from all families recruited

during the first expedition. Datasets were then compared by the Kappa test (Cohen, 1960).

The phenotype “number of decayed teeth” (DT) was investigated as a continuous trait. To bring the observed distributions closer to normality, DT was root-transformed. Prior to the Complex Segregation Analysis, the impact of non-genetic covariates on DT was investigated by univariate and multivariate linear regression analysis, as implemented in the *t* test, correlation and linear regression functions of the SPSS software (version 13.0). Covariates yielding statistically significant associations with DT in the multivariate analysis were included in the corresponding Complex Segregation Analysis.

Complex Segregation Analysis was conducted following the same regressive model applied recently for this population and for leprosy (Bonney, 1986; Lázaro *et al.*, 2010). In the first model, sporadic transmission (model I) includes only the non-genetic covariates with significant impact over disease susceptibility. Next, in addition to the significant covariates, the dependence on phenotypes of preceding relatives, which is parameterized in terms of familial correlations (model II), is included in the model, with the class D pattern of familial correlations (Lázaro *et al.*, 2010). Four types of phenotypic familial correlation were considered: father-mother (FM), father-offspring (FO), mother-offspring (MO), and sib-sib (SS), with corresponding regression coefficients denoted as “ ρ_{FM} ”, “ ρ_{FO} ”, “ ρ_{MO} ”, and “ ρ_{SS} ”, respectively. To rule out the possibility of significant familial dependency due to unmeasured shared environmental factors, our next step was to include a major gene effect in the model (models III and IV). The estimated parameters of the major gene were q (the frequency of allele A predisposing to decayed teeth), μ_{AA} , μ_{AB} , and μ_{BB} (the phenotypic means of individuals with genotypes AA, AB, and BB, respectively), and σ^2 (the common residual variance of the phenotype). Finally, two additional models including a major gene effect were considered: (i) an “absence of transmission” model (model V), in which three types of individuals—AA, AB, and BB—are specified but in which absence of parent-offspring transmission is obtained by setting $\tau_{AAA} = \tau_{ABA} = \tau_{BBA}$; and (ii) a more general transmission model (model VI), in which the three τ s are estimated (Elston and Stewart, 1971). Segregation of a major gene can be inferred if we fail to reject the Mendelian transmission of the major effect when compared with the general transmission model and if we reject the non-transmission hypothesis when compared with the general transmission model (this latter test rules out the possibility that the failure to reject the Mendelian transmission of the major effect was due to lack of power) (Demenais *et al.*, 1986).

Parameter estimation and hypothesis testing were performed with the use of classic likelihood strategies. Because of the study design—exhaustive reconstruction and data collection of randomly selected (instead of proband-based selected) extended pedigrees—there was no need for ascertainment correction (Tai and Hsiao, 2001). Complex Segregation Analysis was performed as implemented in S.A.G.E. (Statistical Analysis for Genetic Epidemiology) release 5.4.1 (Case Western Reserve University), with the segregation analysis program SEGREG (Sorant *et al.*, 1994).

Table 1. Demographic and Clinical Characteristics of the Study Population

Demographic and Clinical Characteristics			Distribution According to DT	
Variable Group	Characteristic	Value	p value Univariate	p value Multivariate
Demographic	Ethnic group, n (%)		0.32	
	Caucasian	26 (8.5)		
	Black	45 (14.6)		
	Mixed	237 (76.9)		
	Gender, n (%)		0.60	
	Male	227 (50.3)		
Oral Health	Female	224 (49.7)		
	Age (mean $\pm$ SD)	30.72 $\pm$ 16.31	0.005 [†]	0.02 [†]
	Snack habit, n (%)		0.03 [†]	0.19
	Yes	216 (58.1)		
	No	156 (41.9)		
	Toothbrushing/day, n (%)		0.25	
	Once	49 (13.2)		
	More than one time	321 (86.8)		
	Flossing every day, n (%)		0.22	
	Yes	123 (33.2)		
	No	248 (66.8)		
	Dental appointment/yr, n (%)		0.07	
SES*	None	227 (66.2)		
	One or more	116 (33.8)		
	Socio-economic Status, n (%)		0.31	
	No income	168 (56.7)		
	With income	128 (43.3)		
Clinical	Schooling, n (%)		0.95	
	Up to first grade	125 (42.4)		
	Second grade incomplete or higher	170 (57.6)		
	Gingivitis, n (%)		0.0004 [†]	0.0002 [†]
	Yes	146 (42.6)		
	No	197 (57.4)		
	Plaque, n (%)		0.08	
	Absent	96 (34.3)		
	Clinically visible	184 (65.7)		
	Number of teeth (mean $\pm$ SD)	19.21 $\pm$ 9.79	0.0001 [†]	0.07
	Leprosy status, n (%)		0.57	
	Yes	86 (23.1)		
	No	287 (76.9)		

[†]Statistical significance. *Categorical variables collapsed into binary to optimize the analysis; n = number of individuals with available information. DT = number of decayed teeth. All commercial toothpastes available in Brazil are fluoridated.

RESULTS

Family Sample and Phenotype Distribution

The kappa value for the statistical test of intra-observer repeatability was 0.83, indicating “very good agreement” between the datasets (kappa > 0.8) (Cohen, 1960).

In total, 451 individuals were enrolled in the study, distributed in 11 extended pedigrees. The male/female ratio was 1.01, mean age at enrollment was 30.72 ($\pm$ 16.31) yrs old, and mean DT was 2.44 ($\pm$ 3.34; minimum = 0, maximum = 17). There was no difference between males and females for distribution of age (mean age for males = 31.0 years; mean age for females = 30.3

years, p value = 0.68) and DT (mean DT for males = 2.33; mean DT for females = 2.60; p value = 0.47).

Analysis of Covariates

Univariate analysis revealed effects of age, snack habits, gingivitis, and number of teeth over DT (Table 1). Multivariate linear regression analysis confirmed the impact of age (p = 0.02) and gingivitis (p = 0.0002): Caries was significantly more frequent among individuals with gingivitis (mean decayed teeth = 1.44) when compared with individuals without gingivitis (mean decayed teeth = 0.98; p = 0.0004). The mean DT reached the highest value at the age class of 20 to 29 yrs old.

Table 2. Complex Segregation Analysis of Quantitative Phenotype Number of Decayed Teeth, Accounting for Age and Gingivitis Status

Model	Q_A	μ_{AA}	μ_{AB}	μ_{BB}	σ^2	$\rho_{Sib-Sib}$	$\beta_{Root\ Age}$	$\beta_{Gingivitis}$	τ_{AAB}	τ_{ABB}	τ_{BBB}	-2lnL + C
I. Sporadic	(0)	2.55	[$=\mu_{AA}$]	[$=\mu_{AA}$]	11.14	(0)	-0.08	1.52	-	-	-	108
II. Familial Correlation Sib-Sib	(0)	2.44	[$=\mu_{AA}$]	[$=\mu_{AA}$]	11.08	0.13	-0.06	1.51	-	-	-	104
III. Major Gene and Familial Correlation Sib-Sib	0.63	1.49	[$=\mu_{AA}$]	9.44	3.87	0.10	-0.16	0.63	(1)	(0.5)	(0)	4
IV. Major Gene	0.63	1.53	[$=\mu_{AA}$]	9.53	3.88	(0)	-0.15	0.60	(1)	(0.5)	(0)	4
V. Absence of Transmission	0.63	1.52	[$=\mu_{AA}$]	9.55	3.83	(0)	-0.18	0.59	0.63	0.63	0.63	8
VI. General Transmission	0.54	1.52	[$=\mu_{AA}$]	9.54	3.81	(0)	-0.18	0.62	1.00	0.52	0.41	0

“-” Non-relevant parameter in the model.

“() ” Fixed parameter for hypothesis.

“[] ” Parameter fixed to the same value as the preceding estimated parameter.

“Q” Frequency of leprosy-predisposing allele “A”.

“ μ ” Genotypic mean for genotype (AA, AB, or BB), adjusted for covariate effects.

“ σ^2 ” Residual variance of the phenotype.

“ ρ ” Familial correlation.

“ β ” Covariable regression coefficients.

“ τ ” τ_{AAB} , τ_{ABB} , τ_{BBB} probabilities of transmitting “a” for individuals AA, AB, and BB, respectively.

C = -1364.40, corresponding to twice the logarithm of the likelihood (2lnL) of the best-fitting model (model VI).

Complex Segregation Analysis

Results of the Complex Segregation Analysis are shown in Table 2. Since correlations between spouses (ρ_{FM}), father-offspring (ρ_{FO}), and mother-offspring (ρ_{MO}) were not observed on Complex Segregation Analysis, these parameters were not included in Table 2. There was evidence of familial correlation, since the sporadic model without familial correlation was rejected against the model that included sib-sib correlation [models I vs. II, $\chi^2(1df) = 4.56$, $p = 0.03$]. The inclusion of a dominant major effect to sib-sib correlations notably improved the fitness of the model [models II vs. III, $\chi^2(3df) = 100$, $p = 10^{-21}$]. Interestingly, removal of residual sib-sib correlations had no significant impact on the fitness [models III vs. IV, $\chi^2(2df) = 0.62$, $p = 0.73$]. Finally, the transmission of the dominant major effect was compatible with the Mendelian hypothesis [IV vs. VI, $\chi^2(3df) = 4.16$, $p = 0.24$], and the hypothesis of no transmission was rejected [V vs. VI, $\chi^2(2df) = 8.2$, $p = 0.01$]. In summary, there was strong evidence for the presence of a major gene controlling DT, following a dominant model with an estimated frequency of the resistance allele “A” of 0.63.

DISCUSSION

The objective of this study was to conduct a Complex Segregation Analysis for caries in a collection of multiplex, multigenerational families recruited at the former leper Colony of Santo Antônio do Prata, located in the Amazonian state of Pará, north of Brazil. Environmental, socio-economic, educational, and demographic variables are very homogenous throughout the colony (Lázaro *et al.*, 2010); in addition, characteristics such as a high cariogenic diet, low standards of oral hygiene, and the absence of fluoride in the water, known to have a major influence on caries development, make the Prata colony particularly suitable for genetic epidemiologic analysis of caries susceptibility. Caries experience was expressed by the “Decayed Teeth”

component of the “Decayed, Missing and Filled Teeth” classic index that captures information not only about caries but also about other oral conditions such as periodontitis, trauma, and even over-treatment; a Complex Segregation Analysis on such a complex phenotype would be highly prone to cryptic confounding effects, and therefore very difficult to interpret.

The goal of Complex Segregation Analysis is to detect and discriminate between and among the different factors causing familial resemblance, ultimately aiming to demonstrate a major gene effect. Our Complex Segregation Analysis found the existence of a major gene effect transmitted following a dominant model with the frequency of the resistance allele “A” estimated at 0.63. Mean DT was estimated at 1.53 for individuals with genotypes AA and AB and 9.53 for BB individuals. The analysis was adjusted by age and gingivitis, both associated with DT in our study population sample: DT increased with age only up to the 20- to 29-year-old age class and decreased after that; patients with gingivitis had higher mean DT. From a clinical point of view, gingivitis is a result of plaque accumulation on the tooth surfaces owing to inadequate oral hygiene. Gingivitis is a significant predictor of higher numbers of caries lesions and is compatible with the finding that plaque occurrence is also associated with caries (Julián *et al.*, 2006). Number of teeth present in the mouth was, not surprisingly, associated with DT in the univariate analysis; however, the effect was totally explained by age, as demonstrated in the multivariate analysis.

Although Complex Segregation Analysis is ideal for identifying the genetic model behind a specific phenotype, it has classic limitations. Different strategies of adjustment for non-familial variables may result in discordant findings across independent studies; our study population was heavily and homogeneously exposed to caries risk factors, minimizing the impact of non-familial variability and increasing the power of the analysis. Also, our segregation analysis does not preclude other non-genetic explanations for the familial nature of the caries risk in the Prata population, such as passing of microbiota

from caregivers to infants; however, the risk of such confounding is minimized by the observation of Mendelian transmission of the trait (models IV vs. VI and V vs. VI), unlikely to be seen for non-genetic variables. Finally, a more precise inference about the exact nature of the major gene effect observed is impaired by the inability of the Complex Segregation Analysis to distinguish between one single, co-dominant, or dominant gene with very strong effect or several co-dominant or dominant genes with milder effect that will play additively on the risk (Jarvik, 1998). For further investigation, molecular, DNA-based studies are necessary.

The results of the Prata Colony Complex Segregation Analysis on caries add to the evidence of a genetic component controlling the development of caries and lay groundwork for future genetic studies. The parameters generated in this Complex Segregation Analysis may be used for powerful, model-based linkage analysis, followed by high-density association mapping using the already-mapped Prata families, as used in previous studies (Mira *et al.*, 2004). To date, only one combined parametric and non-parametric linkage analysis has been conducted to detect genomic regions containing candidate genes for caries (Vieira *et al.*, 2008). Unfortunately, no data from the Complex Segregation Analysis performed by the authors were made available.

The description of the exact nature of the genetic component of the complex mechanism controlling susceptibility to human caries may ultimately lead to a better understanding of the physiopathological basis of this complex, chronic, multifactorial, and common disease and Public Health concern, with potential impact over preventive strategies and, consequently, over public health systems worldwide.

ACKNOWLEDGMENTS

We thank all families for agreeing to participate in the study. We also thank the local team from the Federal University of Pará, as well as the leaders and health workers of the Prata Colony, for their help in guiding our team through the community, socially and geographically. The Core for Advanced Molecular Investigation laboratory of PUCPR is supported by grants from The Brazilian Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and by the INSERM-CNPq joint program, project 490844/2008-1 (ASCIN/CNPq 61/2008 – Convênios Bilaterais Europa). Renata I. Werneck was supported by CAPES/PDEE and CAPES/PROSUP scholarships. Some of the results in this paper were obtained by using the Statistical Analysis for Genetic Epidemiology (S.A.G.E.) package, which is supported by a US Public Health Service Resource Grant (RR03655) from the National Center for Research Resources. This article is based on a thesis submitted to the Graduate Program in Health Sciences, Center for Biological and Health Sciences, Pontifical Catholic University of Paraná, in partial fulfillment of the requirements for the PhD degree.

REFERENCES

- Antunes JLF, Peres MA (2006). Fundamentos de Odontologia Epidemiologia da Saúde Bucal. 1st ed. Rio de Janeiro: Guanabara Koogan.
- Bonney G (1986). Regressive logistic models for familial disease and other binary traits. *Biometrics* 42:611-625.
- Cohen J (1960). A coefficient of agreement for nominal scales. *Educ Psychol Meas* 20:37-46.
- Deeley K, Letra A, Rose E, Brandon C, Resick J, Marazita M, *et al.* (2008). Possible association of amelogenin to high caries experience in a Guatemalan-Mayan population. *Caries Res* 42:8-13.
- Demenais F, Lathrop M, Lalouel J (1986). Robustness and power of the unified model in the analysis of quantitative measurements. *Am J Hum Genet* 38:228-234.
- Elston R, Stewart J (1971). A general model for the genetic analysis of pedigree data. *Hum Hered* 21:523-542.
- Evans R, Lo E, Darvell B (1993). Determinants of variation in dental caries experience in primary teeth of Hong Kong children aged 6-8 years. *Community Dent Oral Epidemiol* 21:1-3.
- Featherstone J (2004). The continuum of dental caries—evidence for a dynamic disease process. *J Dent Res* 83(Spec Iss C):C39-C42.
- Fejerskov O (2004). Changing paradigms in concepts on dental caries: consequences for oral health care. *Caries Res* 38:182-191.
- Jarvik G (1998). Complex Segregation Analyses: uses and limitations. *Am J Hum Genet* 63:942-946.
- Julián A, Barr Agholme M, Grindefjord M, Modér T (2006). Risk factors and risk indicators associated with high caries experience in Swedish 19-year-olds. *Acta Odontol Scand* 64:267-273.
- Keyes P (1960). The infectious and transmissible nature of experimental dental caries. Findings and implications. *Arch Oral Biol* 1:304-320.
- Lázaro F, Werneck R, Mackert C, Cobat A, Prevedello F, Pimentel R, *et al.* (2010). A major gene controls leprosy susceptibility in a hyperendemic isolated population from north of Brazil. *J Infect Dis* 201:1598-1605.
- Lukacs J (2011). Sex differences in dental caries experience: clinical evidence, complex etiology. *Clin Oral Invest* [Epub ahead of print, July 21, 2010] (in press).
- Mira M, Alcaïs A, Nguyen V, Moraes M, Di Flumeri C, Vu H, *et al.* (2004). Susceptibility to leprosy is associated with PARK2 and PACRG. *Nature* 427:636-640.
- Ozturk A, Famili P, Vieira A (2010). The antimicrobial peptide DEFB1 is associated with caries. *J Dent Res* 89:631-636.
- Patir A, Seymen F, Yildirim M, Deeley K, Cooper M, Marazita M, *et al.* (2008). Enamel formation genes are associated with high caries experience in Turkish children. *Caries Res* 42:394-400.
- Petersen PE (2003). The World Oral Health Report 2003: continuous improvement of oral health in the 21st century—the approach of the WHO Global Oral Health Programme. *Community Dent Oral Epidemiol* 31(Suppl 1):3-23.
- Pine C (2005). Caries risk: individual and population perspective. Proceedings of a Symposium at the 81st Annual Meeting of the International Association for Dental Research (IADR). Göteborg, Sweden, 28 June 2003. *Community Dent Oral Epidemiol* 33:239-279.
- Sorant A, Bonney G, Elston R, Bailey-Wilson J, Wilson A (1994). SAGE statistical analysis for genetic epidemiology. Release 2.5.4. In: CWRU program available from the Department of Epidemiology and Biostatistics. Cleveland, OH: Case Western Reserve University.
- Tai J, Hsiao C (2001). Effects of implicit parameters in segregation analysis. *Hum Hered* 51:192-198.
- Vieira A, Marazita M, Goldstein-McHenry T (2008). Genome-wide scan finds suggestive caries loci. *J Dent Res* 87:435-439.
- Werneck R, Mira M, Trevilatto P (2010). A critical review: an overview of genetic influence on dental caries. *Oral Dis* 16:613-623.
- World Health Organization (1997). Oral health surveys: basic methods. 4th ed. Geneva: World Health Organization.