INSIGHTS INTO THE GENETIC BASIS OF HUMAN DENTAL VARIATION FROM STATISTICAL MODELLING ANALYSES

PAULA J DEMPSEY¹, GRANT C TOWNSEND¹ and NICK G MARTIN²
¹ Dental School, The University of Adelaide, Adelaide, SA 5005, Australia
² Queensland Institute of Medical Research, 300 Herston Rd, Herston, QLD 4029, Australia

ABSTRACT

Understanding the genetic basis of dental variation provides a strong foundation for both anthropological studies and clinical practice. Since family pedigrees for dental traits generally have not yielded evidence of single gene transmission, researchers have resorted to other designs. Inference about genetic causes of variation from analyses of correlations between relatives have been utilised for many years, but traditional methods have had distinct inadequacies. Structural equation modelling applies models of genetic and environmental factors to covariances among relatives, and overcomes many of the disadvantages of previous methods. The improvements include: 1) separation of common or family environment from genetic factors; 2) separation of additive from non-additive genetic effects; 3) the ability to fit the models to multivariate data; 4) generation of goodness-of-fit statistics for each model; 5) statistical comparison of models; and 6) testing of significance of individual parameters. This technique has been used for analyses of permanent tooth crown size in 298 pairs of Australian twins. The analyses revealed a high proportion of genetic variation in tooth crown size, with both additive and non-additive genetic factors being statistically significant. Common and unique environmental effects on variation also were detected. Non-additive genetic effects on variation were centred on the canine and first premolar, suggesting an important role for these teeth during human evolution. Common environmental effects were most strongly apparent in the maxillary first molars, indicating a possible role of the uterine environment in the determination of the size of these teeth.

Key words: quantitative genetics, family studies, twins, tooth crown size

INTRODUCTION

The aim of this paper is to demonstrate a method of genetic and environmental analysis which is particularly well suited to the study of human dental morphology. Very few dental traits demonstrate simple Mendelian inheritance, or simple environmental determination, so it is assumed that most of these traits are the product of multiple genetic and environmental effects. Quantifying the sources of variation is one of the main aims of quantitative genetics.

Quantitative Genetic Theory

In general, phenotypic variation may be expressed as the sum of genetic and environmental variances: \( V_P = V_G + V_E \). Furthermore, the total genetic effect on a phenotype \( (V_G) \) can be divided into additive effects of alleles \( (V_A) \), dominance effects between alleles at the same locus \( (V_D) \), and epistatic interactions between loci \( (V_I) \). The total contribution of the environment may likewise be split into effects of common environment
(\(V_C\) – also called shared or family environment) and individual environment (\(V_E\) – also known as specific, unique or random environment) (Fisher, 1918; Mather and Jinks, 1982; Falconer, 1989). For example, the uterine environment is an important source of both types of environmental variance. Some aspects, such as maternal diabetes or teratogens in the blood supply, may be shared by a pair of twins; other aspects may differ between them, such as a difference in the placement of the placenta causing a difference in the rate of exchange of nutrients and wastes.

Thus, total phenotypic variation for a multifactorial trait has numerous components, and may be summarised as: \(V_P=V_A+V_D+V_I+V_C+V_E\). Additional factors which can influence phenotypic expression and correlations among relatives are assortative mating, genotype-environment correlation (CorGE), and genotype by environment interaction (GxE). Genotype-environment correlation occurs when the environments an individual experiences are not a random sample of all environments, but are influenced by, or correlated with, the individual's genotype. Genotype by environment interaction relates to the way genes and environment determine the phenotype. It describes the situation in which one genotype may be expressed the same way in two different environments, while another genotype changes. The use of variance components in genetic analyses and the modelling of the various genetic and environmental factors and interactions are described in detail elsewhere (Neale and Cardon, 1992; Hopper, 1993), so only basic information is included in this paper.

These components of variance may be used to generate heritability estimates. The proportion of phenotypic variation due to all genetic sources of variance is referred to as broad-sense heritability, and is estimated as \(V_G/V_P\). Narrow-sense heritability is the proportion of phenotypic variation due to additive genetic variance, and is estimated as \(V_A/V_P\). These estimates may vary from sample to sample, whether in different populations, different generations, even possibly in different groups of individuals within the same population and generation (due to sampling effects). Heritability estimates are dynamic, reflecting the processes which affected trait development in the given sample. A new environmental factor may decrease heritability in a sample relative to a former generation. Estimates thus apply to the population and generation in question, and their accuracy depends on the degree to which the sample was representative of the population. Wider application of results from a sample to the species level is probably better avoided, unless widespread samples reveal consistent results.

**Tooth Crown Size in the Permanent Dentition**

There have been a large number of reported investigations into genetic factors contributing to variation in tooth crown size, with Korkhaus' (1930) being one of the earliest. Most of the studies have utilised family data such as twins, triplets, full siblings, half siblings, cousins and parent-offspring pairings. Estimates for heritability from these studies range from 21% (Townsend et al, 1986) to 90% (Garn et al, 1965), although most are over 60%. In fact, statistically significant genetic variation has been reported by most researchers, including Lundström (1948), Osborne et al (1958), Kraus et al (1959), Garn et al (1965), Potter and Nance (1976), Rebich and Markovic (1976), Townsend (1978, 1992), Mizoguchi (1980), Sharma et al (1985) and Harzer (1987). Such concordance between studies led Garn (1977:67) to state that unless we discover a significant common environmental contribution, we can retain the notion that dental crown dimensions "are to the largest extent gene-determined."
Typically, researchers have assumed a polygenic pattern of inheritance for crown diameters. Support for this assumption was provided by one study involving pedigree analysis and complex segregation analysis, which failed to reveal any major gene effects (Kolakowski and Bailit, 1981). More detailed reviews of the role of genetic and environmental factors have been published by Potter (1976), Garn (1977), Kieser (1990) and Townsend (1992).

**Why We Need More Genetic Analyses of Tooth Crown Size**

Many former studies have been plagued by problems such as using methods unsuitable for distinguishing components of variation or estimating heritability. In some cases heritability was not calculated because of reservations about methodologies (Osborne *et al.*, 1958; Potter and Nance, 1976). Numerous components were assumed to be absent, particularly common environment, GxE and CorGE.

There are three main categories of genetic analysis that have been applied to tooth crown size: correlation analyses, multiple abstract variance analyses and Fisher's biometrical genetic approach (Jinks and Fulker, 1970). The first method entails calculation of correlation coefficients for pairs of related individuals, followed by substitution of the coefficients into various equations for heritability. Perhaps the best known example is Bulmer (1970) and Smith's (1974, 1975) heritability estimate: \( h^2 = 2(r_{MZ} - r_{DZ}) \).

The heritability estimate equation is severely limited in that it is valid only under a number of assumptions. For instance, the genetic variation must be additive in nature, or heritability estimates will be greater than 1.0. In fact, complex genetic factors and interactions will probably result in non-sensical values in most of the simple correlation-based estimates of heritability.

The second method (multiple abstract variance analyses) involves calculations of ratios of mean squares from analyses of variation. These represent an improvement on correlation analyses, but still do not allow testing of models for goodness-of-fit to the data (Jinks and Fulker, 1970). Fisher's biometrical genetic approach represents a major breakthrough in estimating various genetic and environmental effects, since it allows testing for, and estimation of, the components which were previously assumed absent or undetectable (ibid). The approach is considered to incorporate the other two methods, and extend beyond them.

Twin studies have been complicated further by the assertion that a number of assumptions, necessary to the interpretation of twin data, should be tested before proceeding with data analysis. Christian and co-workers (Christian *et al.*, 1974; Christian, 1979) recommended testing for differences in means, variances and environmental covariances between zygosities, and using modified formulae if significant differences were found. In light of current biometrical genetic methods, some aspects of these tests seem inappropriate, and the assumptions they test, unnecessarily severe. For instance, the proposed test for equality of environmental covariances of monozygotic twins (MZ) and dizygotic twins (DZ) is an F test of among-DZ variances divided by within-DZ variances. If the calculated ratio is not appreciably greater than 1, it is assumed to be unlikely that any substantial proportion of the variance is genetic (Corrucini *et al.*, 1990). However, the test appears to assume that the genetic variation is additive. Significant non-additive genetic variation will cause within-DZ variances to rise, and the F ratio may then suggest that there is no significant genetic variation.

Other problems with previous studies include sample sizes being too small, inaccurate or non-existent methods of zygosity determination, and even pooling of the sexes in spite of significant sexual dimorphism in tooth crown size (reviewed by Mizoguchi, 1977).
addition, most studies involved single or multiple univariate analyses, without multivariate estimates of heritability. There was also no testing of models for goodness-of-fit.

Aims

We have conducted analyses of tooth crown size involving nearly 300 pairs of twins, with more accurate zygosity determination using blood proteins or DNA, and exploration of sex differences in means and variances. The aims of this study are to find and quantify as many genetic and environmental components of variation as possible.

MATERIALS AND METHODS

Data Collection

Data were obtained from 596 subjects, incorporating 149 monozygous (MZ) twin pairs and 149 dizygous (DZ) twin pairs. Alginate impressions of the dentition were obtained, and stone models were made. Twin zygosities were based on examination of 17 blood proteins. The probability of dizygosity given concordance for all systems was less than 1%. Data collection methods were approved by the Committee on the Ethics of Human Experimentation, University of Adelaide (Approval No H/07/84), and all participants were informed volunteers.

Maximum mesiodistal (MD) and buccolingual (BL) crown diameters were recorded, following the definitions of Moorrees et al (1957). All emerged and sufficiently intact permanent teeth were included, except the third molars, yielding a maximum of 56 variables per subject. Measuring equipment comprised Mitutoyo digital vernier callipers, specially honed to produce finer points for more accurate measuring, and connected via a multi-plexer unit to an Apple IIc computer. The calipers gave readings to the nearest 0.1 mm. Data were subsequently transferred to a Sun Sparc Server 2 computer for analysis. Adjustments for age were not considered necessary, since (1) the final size of dental crowns is determined before emergence of the teeth into the oral cavity, and (2) any teeth displaying significant attrition at measurement sites were excluded from subsequent analyses.

Data Analysis

The data were tested first for evidence of GxE interactions, by looking for significant regressions of MZ pair variances on MZ pair means (Jinks and Fulker, 1970). Assortative mating with respect to the dentition was not tested, although it generally is assumed to be unlikely to occur (Potter et al, 1968; Hanihara et al, 1975). In fact, family studies have revealed that correlations between parents were not significantly different from zero (Niswander and Chung, 1965; Bowden and Goose, 1969; Townsend and Brown, 1978; El-Nofely and Tawfik, 1995). CorGE also was not tested for, since it is necessary for the particular source of environmental variation to be identifiable and measurable. It was thus assumed to be absent or trivial.

We then applied statistical models based on path analysis (Wright, 1921) to the variances and covariances calculated from the data. The software packages used for these analyses were PRELIS (Jöreskog and Sörbom, 1986), for calculation of the variance-covariance matrices, and Mx (Neale, 1995), for the modelling.

Figure 1 is a path diagram indicating all of the sources of variance and covariance for a single measurement in a pair of twins. Squares indicate measured variables, while circles represent latent (unmeasured) variables. The latent factors are A (additive genetic variation),
C (common environment), E (unique environment) and D (non-additive genetic variation). D incorporates both dominance and epistatic interaction variance, which cannot be separated when only MZ and DZ twins are used (Mather, 1974).

The path coefficients represented by lower case letters a, d, c, and e estimate the contribution of each factor to the phenotypic variation. Latent factors A, D, and C are correlated (rMZ and rDZ) between twins, with the extent of the correlations being determined by genetic theory (Falconer, 1989).

![Path diagram of four sources of variance and covariance for a pair of twins.](image)

Figure 1. Path diagram of four sources of variance and covariance for a pair of twins.

Models were constructed from the four sources by setting different combinations of paths to zero. We began with a model for unique environment only, and gradually increased the complexity of the model by adding parameters. A total of five models may be fitted to twin data: E, CE, AE, ACE, and ADE. For each model, the path coefficients were estimated and used to calculate expected variances and covariances. Expected values were then tested for goodness-of-fit with observed values using $x^2 = F(N-1)$, where F is the minimum of the fitting function and N is the number of observations on which the observed covariance matrix was based (Neale and Cardon, 1992). The degrees of freedom (df) was calculated as the number of independent statistics minus the number of free parameters.

In order to compare models with different degrees of freedom, Akaike's Information Criterion (AIC = $x^2 - 2df$) was used (Akaike, 1987). The smaller or more negative the AIC the better the parsimony and fit of a model. The statistical significance of the difference between two models was tested using the difference between the relevant $x^2$ values.

The general approach was that of retaining the simplest acceptable model, unless a more complex model was significantly better by $x^2$ (p<0.05). For univariate analyses, p=0.01 was chosen as the level of significance due to the number of tests being performed. In addition, comparisons of $x^2$ values between complex and simpler models indicate the significance of the various components (p=0.05). Heritability ($h^2$) was defined as the ratio of genetic variation to total phenotypic variation, and estimated using the parameter estimates from the accepted model.
RESULTS AND DISCUSSION

To date, univariate and three multivariate analyses of the permanent dentition have been completed. The former include mesiodistal (MD) and buccolingual (BL) diameters of the 28 permanent teeth, excluding third molars, yielding 56 variables. The multivariate analyses include the MD diameters of the eight incisors, and MD and BL diameters of the seven teeth in the maxillary right quadrant. Detailed results have been, or will be, reported elsewhere (Dempsey et al, 1995, 1996, 1998), so only summaries of the findings are presented here.

Univariate Analyses

There was no evidence for GxE interaction in any of the 56 variables. For the univariate analyses, in nearly all cases the simplest adequate model was the AE model. Significant improvements in fit were achieved by including non-additive genetic variation for the MD dimension of the four canines, both of the maxillary first premolars, and the mandibular right first premolar, as well as the BL diameter of the maxillary right first premolar. Likewise, improvement in fit was observed when common environment was included in the BL and MD dimensions of the maxillary first molars.

Averaging over right and left sides, the proportion of total variance attributed to additive genetic effects ranged from 59%, for the maxillary first molar MD diameter, to 92%, for the maxillary first premolar BL diameter, with most proportions being greater than 80%. The proportion of total variance attributed to common environment averaged 26% for MD and 24% for BL diameters of the maxillary first molars. Non-additive genetic variation estimates were confounded with those of additive genetic variation, and separate estimates could not be obtained, although significance of the parameter was testable. The remaining proportions of total variance, thus, were attributed to unique environmental variation and measurement error. There was a high degree of symmetry in all estimates between antimeres (right and left teeth).

Multivariate Analyses

Overall, heritability estimates from multivariate analyses were similar to those from univariate analyses (Dempsey et al, 1998). Additive genetic variation explained from 26% to 86% of total variance. Significant non-additive genetic variation was detected in MD diameters of all seven teeth, and BL diameters of all but the canine. The main effect was centred on the MD diameter of the canine and first premolar. Values ranged from 7% to 15% for BL dimensions, and 14% to 55% for MD dimensions. Common environmental variation was significant for the MD diameters of the four posterior teeth (0.04% to 17%) and the BL diameters of the molars (7% to 29%).

One theory states that non-additive genetic variation is displayed by genes which are related to selective fitness (Fisher, 1958; Kacser and Burns, 1981; Dean et al, 1988). Therefore, evidence for this type of variation may indicate that selective pressures are acting currently, or did so sometime in the past. It therefore is of interest that the canine, the most sexually dimorphic tooth in humans (Garn et al, 1964, 1967; Alvesalo, 1971; Yamada and Sakai, 1992), is the tooth which displays the strongest evidence for non-additive genetic variation.

Evidence for the presence of common environmental variation in both diameters of the maxillary first molars also is interesting. These teeth are the ones most likely to show signs of uterine environmental effects, due to their early calcification (which commences perinatally).
However, whether the effects are pre- or post-natal, there is an inconsistency in the absence of any such effects in the mandibular first molars. This phenomenon is not readily explained.

Advantages of the Biometrical Genetic Model

When applied to MZ and DZ twin pairs, the biometrical genetic modelling procedure allows testing for and (in most cases) estimation of additive genetic, non-additive genetic, common and unique environmental sources of variation. Both broad- and narrow-sense heritabilities can be generated where both additive and non-additive genetic factors are estimated.

The process can be extended beyond the modelling of variances and covariances to include the modelling of means for each of the twin groups. This allows testing for other aspects, for instance whether there are sex differences in means. Other effects, such as assortative mating, CorGE, and GxE, can be incorporated, and a variety of types of heterogeneity between the sexes can be explored (Neale and Cardon, 1992). The models can be compared, so that the most parsimonious, and the best fitting, models can be identified.

Limitations of the Biometrical Genetic Model

When applied to data from twins raised together, the method is restricted in several ways. Firstly, univariate models incorporating both C and D cannot be fitted due to negative confounding of both influences (Grayson, 1989). Caution also must be applied with multivariate models, since models will be under-identified if too many parameters are included simultaneously. This difficulty can be resolved in some cases by including data from parent-offspring pairs. Secondly, dominance and epistatic interactions usually are not separable (Mather, 1974). Thirdly, assortative mating cannot be tested, unless data is collected from the parents of the twins (Neale and Cardon, 1992). In addition, the GIGO rule (garbage in, garbage out) of computer programming applies - the results are only as good as the models applied to the data.

CONCLUSIONS

The application of biometrical genetic modelling procedures can, and is, generating new insights into the possible ontogeny and evolution of the human dentition. There is no reason why estimating heritability should be the only aim of a modern research program, when finer detail can be provided by biometrical genetic analyses. The software package used - Mx - is available free of charge on the internet (http://griffin.vcu.edu/mx), and courses in its use are offered regularly by the Institute for Behavior Genetics, Colorado, USA.

ACKNOWLEDGEMENTS

The South Australian Twin Study is supported by a grant from the National Health and Medical Research Council (NHMRC) of Australia, and by grants from the University of Adelaide. The assistance of the Australian NHMRC Twin Registry is greatly appreciated. The twins and their families are also thanked for their enthusiastic participation in the study. Professor Tasman Brown, Lhdiasy Richards, John Kaidonis, Brian Penhall, Jim Rogers, Sandy Pinkerton and Wendy Schwerdt assisted with data collection. George Travon, Stacey Cherny (Institute for Behavior Genetics, Colorado, USA) and John Pearson (QIMR) provided much needed computer support. Mike Neale (Medical College of Virginia) provided invaluable advice on model fitting procedures. We are also grateful to the staff of the Institute of Medical and Veterinary Science and Dr John Mulley of the Adelaide Medical Centre for Women and Children, for their assistance in determining twin zygosities.


REFERENCES


