

## Consanguineous marriages and the genetic load due to lethal genes in Kerala

BY SUSHIL KUMAR, R. A. PAI AND M. S. SWAMINATHAN

*Division of Genetics, Indian Agricultural Research Institute, New Delhi (India)*

The Hindu population throughout India is organized into castes and subcastes which are largely endogamous. The populations in the southern states of India (Andhra, Kerala, Madras and Maharashtra) are unique in the occurrence of a fair frequency of consanguineous marriages within the subcastes, and they thus provide an excellent opportunity for studying the detrimental effects of low levels of inbreeding and for estimating the genetic load. The socio-economic reasons for the high incidence of consanguineous marriages in India have been outlined by Dronamraju (1964) and Sanghvi (1966). Dronamraju & Meera Khan (1963) reported the morbidity pattern in the children of consanguineous and non-consanguineous marriages in a hospital population of Andhra Pradesh. During the present study, several hospital populations in Kerala were examined for the relative mortality and morbidity of offspring from the two types of marriages. Several authors have recently considered the effects of inbreeding on mortality and other traits determining fitness in human populations (Morton, Crow & Muller, 1956; Bök, 1957; Schull, 1958; Slatis, Reis & Hoene, 1958; Chung, Robinson & Morton, 1959; Morton & Chung, 1959; Morton, 1960; Salzano *et al.* 1962; Neel & Schull, 1962; Freire-Maia, Freire-Maia & Quelce-Salgado, 1963; Freire-Maia, Guaraciaba & Quelce-Salgado 1964; Freire-Maia & Krieger, 1963; Goldschmidt *et al.* 1963; Dewey *et al.* 1965). Morton *et al.* (1956) developed formulae for the estimation of genetic load in terms of *A* and *B* statistics which could be used for discrimination between the mutational and segregational components of loads. Their method has been used in the present paper for analysing the data on mortality.

### MATERIAL AND METHODS

The consanguinity data were collected during May and June 1966, from several hospitals in Kerala at Ernakulum, Quilon and Trivandrum. The hospitals are: Government Hospital, Ernakulum, Government Hospital and Neendakara Indo-Norwegian Hospital, Quilon; and Medical College Hospital and Mental Hospital, Trivandrum. Most of the patients received in these hospitals are from within the limits of the respective districts. The Mental Hospital at Trivandrum, however, receives patients from all over the Kerala State. Almost all in-patients during the period of visiting these hospitals were included in this study. In-patients were questioned about the consanguinity and progeny size of their parents. Morbidity data were taken from the hospitals' records. In the Mental Hospital and in the paediatric wards of other hospitals, where a statement could not be obtained from the patient due to the nature of his illness, data were collected by interviewing parents or other relatives. Mortality data were recorded on the families of patients of paediatric wards at the Trivandrum and Quilon hospitals. Only those families were considered where histories of births and deaths could be obtained by direct questioning of the mother or grandmother of the patient. Socio-economic data on caste, profession and income available in the hospital records were utilized.

## RESULTS AND DISCUSSION

*Observations on consanguinity*

Data on the relative proportions of various types of consanguineous marriages are given in Table 1. It will be noted that (1) the mean proportion of consanguinity based on 889 marriages is 0.2036; (2) the incidence of consanguineous marriages in Quilon and Trivandrum districts is significantly higher than in Ernakulam ( $\chi^2_{(1)} = 13.92$ ,  $P > 0.01$ ); (3) the most frequent types of consanguineous marriages are, first, a girl with her maternal uncle's son and, secondly, a girl with her paternal aunt's son; and (4) marriages between children of two brothers or two sisters which are socially forbidden among Hindus do occur in this area, although their relative proportions are low. No case of uncle-niece marriage which is the highest degree of consanguinity permitted in Hindu law and which is frequent in the other southern states was recorded in Kerala.

Table 1. *Relative frequencies and percentages of various types of consanguineous marriages recorded among inmates of some hospitals in Kerala*

Consanguineous marriages	<i>F</i> Value	Ernakulam		Quilon		Trivandrum		Mental Hospital Trivandrum		Mean	
		Fre- quency	Percent- age	Fre- quency	Percent- age	Fre- quency	Percent- age	Fre- quency	Percent- age	Fre- quency	Percent- age
Maternal uncle's son	1/16	2	2	24	12.7	61	12.3	3	3	90	10.1
Maternal aunt's son	1/16	—	—	—	—	3	0.6	—	—	3	0.3
Paternal uncle's son	1/16	—	—	2	1.1	5	1.0	—	—	7	0.8
Paternal aunt's son	1/16	2	2	19	10.1	42	8.5	2	2	65	7.3
Second cousin	1/64	—	—	3	1.6	6	1.2	—	—	9	1.0
Third cousin	1/256	—	—	2	1.1	5	1.0	—	—	7	0.8
Unrelated marriages	0	98	96	139	73.4	375	75.4	96	95	708	79.7
Total				102		189		497		101	889
Mean <i>F</i> and standard error				0.0024 ± 0.0012		0.0152 ± 0.0019		0.0142 ± 0.0013		0.0031 ± 0.0013	

The mean coefficient of inbreeding among the offspring of the 889 marriages studied, analysed without reference to the number of children produced by these marriages, is 0.01182. Using the fertility structure of marriages, the mean coefficient of inbreeding is 0.01056. The difference in the mean fertility of inbred and outbred marriages, indicated by the lower value of inbreeding coefficient when the analysis is based on the offspring of the marriages than when the coefficient is derived from a consideration of marriages alone, is not significant ( $\chi^2_{(1)} = 0.13$ ,  $P = 0.8-0.7$ ). The actual coefficients of inbreeding for this population should be higher than those estimated, since the effects of inbreeding in the past generations have not been considered. It is known that the incidence of consanguineous marriages in the past was much higher than today, a decreasing trend being evident during the last seven decades (Dronamraju, 1964).

The coefficients of inbreeding in sympatric populations of Christians and Muslims have also been calculated for comparison and work out for 84 Christian and 53 Muslim marriages to 0.00351 and 0.01563 respectively. For assessing the possibility of correlating levels of inbreeding with specific illnesses, estimates of inbreeding coefficient values of the parents of patients have been calculated and the scores for *F* are given in Table 2. The coefficients of inbreeding for con-

genital anatomical malformations, diabetes mellitus, susceptibility to tubercular meningitis and pulmonary tuberculosis are higher than for other illnesses or the hospital group as a whole. Analysis of the incidence of tuberculosis in twenty families indicated that the more frequent occurrence of this disease may not be solely due to a greater possibility of exposure. The small number of patients on which these values are based, however, render detailed analysis and drawing of definite inferences difficult.

Table 2. *Estimates of coefficient of inbreeding F.*  
*Estimation of inbreeding coefficients*

Groups	Number of cases encountered	F
Congenital malformations	18	0.022
Pulmonary T.B.	28	0.022
T.B. meningitis	30	0.019
Diabetes mellitus	48	0.019
Liver cirrhosis	23	0.015
Bronchial asthma	32	0.011
Anaemia	26	0.010

Table 3. *Mortality in proportion to the pregnancies in progenies of consanguineous and non-consanguineous marriages*

F	Families studied	Pregnancies	Still births + miscarriages	Infant + juvenile deaths
First cousin 1/16	83	391	59 (0.151)	72 (0.186)
Second cousin 1/64	8	34	4 (0.118)	3 (0.088)
Not related 0	185	770	231 (0.030)	67 (0.087)

*Mortality in children of non-consanguineous and consanguineous marriages*

The observed proportions of mortalities in terms of still births + miscarriages and infant + juvenile deaths (death prior to the age of 15) are given in Table 3. A significantly higher frequency of still births and miscarriages occur in the inbred progenies ( $F = 1/16$  and  $1/64$ ) as compared to outbred ( $F = 0$ ). Infant + juvenile deaths are significantly higher in progenies of first cousin marriages ( $F = 1/16$ ) than in second cousin ( $F = 1/64$ ) or control ( $F = 0$ ) marriages. The total mortality for first cousin, second cousin and unrelated marriages is 33.58, 20.58 and 11.69 % respectively. It is of interest that the still birth and miscarriage rate reported in Table 3 for the unrelated group is much below the WHO figures for India as a whole. This is probably because of the comparatively superior medical, public health and educational facilities available in Kerala, thanks in part to the work of various missionary organizations.

*Estimates of the magnitude of genetic load*

The data presented in Table 3 have been used to estimate the magnitude of genetic loads in the population studied in units of lethal equivalents, where a lethal equivalent is defined as a group of mutant genes which would cause, on the average, one death, if dispersed in different individuals and made homozygous (Morton *et al.* 1956). By the genetic load theory of Morton *et al.* (1956) the proportion of survivors,  $S$ , is expected to be  $S = e^{-(1.2A + BF)}$  where  $F$  is the inbreeding coefficient. In the present study, estimates of  $A$  and  $B$  have been calculated from the simple relationship,  $S = 1 - A - BF$ . Table 4 gives the estimates of  $A$ ,  $B$  and  $B/A$ .

The total load,  $A + B$ , is between 3 and 4 lethal equivalents per gamete, or 6 to 8 lethal equivalents per zygote.

The value of  $A$  is 0.0565 for still births + miscarriages, 0.0762 for infant + juvenile deaths and 0.1328 for total mortality. These fall in the range of values obtained by Morton *et al.* (1956) on the basis of their analysis of the data of Sutter & Tabah (1952) on French populations and for the data of Arner (1908) on an American population. This correspondence is remarkable, considering that the calculation of  $A$  value (i) includes both the genetic and environmental components of mortalities, (ii) depends on the demographic system of the population, and (iii) is subjected to considerable sampling errors.  $A$  values given in Table 4 are smaller than those reported for Japanese (Neel & Schull, 1962) and Brazilian populations (Freire-Maia *et al.* 1963;

Table 4. *Estimates of the A and B statistics and of the B/A ratios*

Deaths	$A$	$B$	$B/A$
Still births + miscarriages	0.0565	1.64388	29.0938
Infant + juvenile deaths	0.0762	1.6887	22.1614
Total mortality	0.1328	3.3325	25.0941

For  $F = 0$  and  $1/16$ ,  $B = 3.5023$ , and for  $F = 1/64$  and  $1/16$ ,  $B = 1.6209$ ,  $\chi^2 = 0.6056$  ( $P = 0.5 - 0.3$ ).

Freire-Maia, 1963; Freire-Maia & Krieger, 1963). The  $B$  value of 3.3325 for total child mortality between the late foetal and early adult stages shows that the effect of inbreeding in the Kerala population described is significant. The ratio  $B/A$  is 25.1 which is comparable to the ratio of 18.12 for the French population of Sutter & Tabah (1952) obtained by Morton *et al.* (1956). Relatively much lower inbreeding responses for mortality were reported by Neel & Schull (1962) in two Japanese populations ( $B/A = 4.63$  and  $4.68$ ), by Freire-Maia (1963) for Brazilian white ( $B/A = 1.2$ ) and Brazilian Negro ( $B/A = 13.8$ ) populations, by Morton *et al.* (1956) for the American populations ( $B/A = 10.75$  and  $7.94$ ) examined by Arner (1908) and Bemiss (1958).

#### SUMMARY

1. The Hindu populations throughout India are organized into castes and subcastes which are largely endogamous. The populations in the southern states of India are unique in the occurrence of a high rate of consanguineous marriages within the endogamous groups. Several hospital populations in a southern state of India (Kerala) have been examined for the rate of consanguinity and the genetic load due to lethal genes.

2. The incidence of consanguineous marriages is about 20% and the estimate of mean coefficient of inbreeding 0.01056.

3. The most frequent types of consanguineous marriages are of a girl with her (a) maternal uncle's son, and (b) paternal aunt's son. The uncle-niece marriage is not preferred in Kerala.

4. The frequency of foetal and infant deaths is significantly higher in inbred progenies than in the outbred. The estimates of total mortality for first cousin, second cousin, and unrelated marriages are 33.58, 20.58 and 11.69% respectively.

5. The total load is between 3 to 4 lethal equivalents per gamete. The estimates of  $B$  and  $B/A$  statistics are high.

Our gratitude is due to Drs K. V. Krishna Das and M. Thangavelu for the facilities provided to us at the Medical College, Trivandrum and to Mr S. Ramanujam for helpful criticism and discussions.

## REFERENCES

- ARNER, G. B. L. (1908). Consanguineous marriages in the American population, vol. 1, 101 p. New York: Columbia University.
- BEMISS, S. M. (1958). Report of influences of marriages of consanguinity upon offspring. *Trans. Am. Med. Assoc.* **2**, 319-425.
- BÖÖK, J. A. (1957). Genetical investigation in a north Swedish population. The offspring of first cousin marriages. *Ann. Hum. Genet.* **21**, 191-221.
- CHUNG, C. S., ROBINSON, O. W. & MORTON, N. E. (1959). A note on deaf mutism. *Ann. Hum. Genet.* **23**, 357-66.
- DWEY, W. J., BARRAI, I., MORTON, N. E. & MI, M. P. (1965). Recessive genes in severe mental defects. *Am. J. Hum. Genet.* **17**, 237-256.
- DRONAMRAJU, K. R. (1964). Mating systems of the Andhra Pradesh people. *Cold Spring Harb. Quant. Biol.* **29**, 81-84.
- DRONAMRAJU, K. R. & MEERA KHAN, P. (1963). The frequency and effects of consanguineous marriages in Andhra Pradesh. *J. Genet.* **58**, 387-401.
- FREIRE-MAIA, N. (1963). The load of lethal mutations in white and Negro Brazilian populations. II. Second survey. *Acta genet.* **13**, 199-225.
- FREIRE-MAIA, N. & KRIEGER, H. (1963). A Jewish isolate in southern Brazil. *Ann. Hum. Genet.* **27**, 31-9.
- FREIRE-MAIA, N., FREIRE-MAIA, A. & QUELCE-SALGADO, A. (1963). The load of lethal mutations in white and Negro Brazilian populations. I. First survey. *Acta genet.* **13**, 185-98.
- FREIRE-MAIA, N., GUARACIABA, S. M. R., M. Z. S. H. & QUELCE-SALGADO, A. (1964). The genetic load in the Bauru Japanese isolate in Brazil. *Ann. Hum. Genet.* **27**, 329-39.
- GOLDSCHMIDT, E., COHEN, T., BLOCH, N., KELETI, L. & WARTSKI, S. (1963). Viability studies on Jews from Kurdistan. *The Genetics of Migrant and Isolate Populations*, pp. 183-95, ed. by E. Goldschmidt. Baltimore: Williams and Wilkins Co.
- MORTON, N. E. (1960). The mutational load due to detrimental genes in man. *Am. J. Hum. Genet.* **12**, 348-64.
- MORTON, N. E. & CHUNG, C. S. (1959). Formal genetics of muscular dystrophy. *Am. J. Hum. Genet.* **11**, 360-79.
- MORTON, N. E., CROW, J. F. & MULLER, H. J. (1956). An estimate of the mutational damage in man from data on consanguineous marriages. *Proc. Natn. Acad. Sci. U.S.A.* **42**, 855-63.
- NEEL, J. V. & SCHULL, W. J. (1962). The effect of inbreeding on mortality in two Japanese cities. *Proc. Natn. Acad. Sci. U.S.A.* **48**, 575-82.
- SALZANO, F. M., MARCALLO, F. A., FRIERE-MAIA, N. & KRIEGER, N. (1962). Genetic load in Brazilian Indians. *Acta. Genet.* **12**, 212-18.
- SANGHVI, L. D. (1966). Inbreeding in rural areas of Andhra Pradesh. *Ind. J. Genet.* **26A**, 351-65.
- SCHULL, W. J. (1958). Empirical risks in consanguineous marriages: sex ratio, malformation and viability. *Am. J. Hum. Genet.* **10**, 294-343.
- SLATIS, H. M., REIS, R. H. & HOENE, R. E. (1958). Consanguineous marriages in the Chicago region. *Am. J. Hum. Genet.* **10**, 446-64.
- SUTTER, J. & TABAH, L. (1952). La mortalité, phénomène biometrique. *Population* **7**, 69-94.

NOTE ON PAPER BY S. KUMAR, R. A. PAI AND M. S. SWAMINATHAN

By CEDRIC SMITH

The estimation of the values of  $A$ ,  $B$  and their ratio  $B/A$  can be improved by using a weighted regression. Let us suppose that the families are divided up into classes ( $i$ ) according to their degrees of relationship ( $F_i$ ); for example, class 1 might consist of all families with first cousin parents ( $F_1 = 1/16$ ), class 2 of all families with unrelated parents ( $F_2 = 0$ ), and so on. In class  $i$  suppose that there are altogether  $n_i$  ( $> 0$ ) children, of which  $m_i$  are affected. The observed proportion of affected is  $p_i = m_i/n_i$ . The corresponding expected proportion we will denote by  $P_i$ . The hypothesis is that

$$P_i = 1 - \exp(-A - BF_i). \quad (1)$$

Suppose that we find in all classes that  $m_i < n_i$  (which will almost always be true), so that  $p_i < 1$ , and let us write

$$z_i = -\ln(1 - p_i), \quad (2)$$

$$Z_i = -\ln(1 - P_i). \quad (3)$$

Then the assumption is that the regression of  $Z_i$  on  $F_i$  is exactly linear

$$Z_i = A + BF_i. \quad (4)$$

We can reasonably estimate  $A$  and  $B$  by minimizing the weighted sum of squares

$$S = \sum_i w_i (z_i - Z_i)^2 \quad (5)$$

where the weight

$$w_i = n_i(1 - P_i)/P_i, \quad (6)$$

although the assumptions for least squares are not absolutely exactly fulfilled (e.g. we have in reality a multinomial distribution instead of a Gaussian one). There is a slight complication here in that the weight depends on the value of  $P_i$ , which again depends on the values of  $A$  and  $B$  we are trying to estimate. The calculation must therefore be done iteratively; we start with provisional values of  $A$  and  $B$ , use these to find appropriate weights  $w_i$ , use the weights to find the new values of  $A$  and  $B$ , use these new values to find new weights  $w_i$ , and so repeatedly until the new values agree with the old to the order of accuracy required.

More exactly, we proceed as follows. Given the values of  $n_i$ ,  $m_i$ , and  $F_i$  we first calculate  $p_i = m_i/n_i$  and  $z_i$  from (2). We now assume provisional values of  $A$  and  $B$ , e.g.  $A = \ln 2$  and  $B = 0$ ; we also set  $j = 1$ , and  $\epsilon =$  some arbitrary small quantity, such as 0.001. From these we calculate, for each class  $i$ , (\*)  $Z_i = A + BF_i$  (this will ordinarily be positive: if exceptionally it should be zero or negative difficulties would arise and the calculation would have to be modified);

$$\begin{aligned} P_i &= 1 - \exp(-Z_i) \\ &= 1 - \text{antilog}(-0.4343Z_i) \end{aligned}$$

$$w_i = n_i(1 - P_i)/P_i;$$

and from these

$$W = \sum w_i,$$

$$\bar{Z} = (\sum w_i Z_i)/W,$$

$$\bar{F} = (\sum w_i F_i)/W,$$

$$\Delta_{FF} = \sum w_i (F_i - \bar{F})^2,$$

$$\Delta_{FZ} = \sum w_i (F_i - \bar{F})(Z_i - \bar{Z}),$$

$$B' = \Delta_{FZ}/\Delta_{FF},$$

$$A' = \bar{Z} - B'F,$$

$$j' = j + 1.$$

We now see whether these new values  $A'$ ,  $B'$  differ appreciably from the provisional values  $A$ ,  $B$ . More precisely if

$$[j < 2] \quad \text{or} \quad [10 |A' - A| + |B' - B| > \epsilon]$$

we replace the provisional values  $A$ ,  $B$ ,  $j$  by the new values  $A'$ ,  $B'$ ,  $j'$ , and repeat the calculation from (\*) onwards. Otherwise we take the values  $A$ ,  $B$  to be, nearly enough, the final estimates. From these we can find the ratio  $B/A$ . But it may be convenient to deal with  $\log(B/A)$  (to base 10) since in small samples any inaccuracy in the denominator  $A$  may very considerably affect the ratio  $B/A$ , and one may reasonably expect the standard error of the estimated  $\log(B/A)$  to be better than the standard error of  $B/A$  itself as a guide to the probable range of values of the true ratio  $B/A$ . The error variances are then found as follows:

$$\text{var}(B) = 1/\Delta_{FF},$$

$$\text{var}(A) = 1/W + B^2/\Delta_{FF},$$

$$\text{var}[\log(B/A)] = 0.1886 [1/W A^2 + (\bar{F}/A - 1/B)^2/\Delta_{FF}].$$

The standard errors are the square roots of the error variances. In addition we may note that

$$\chi^2 = \sum w_i (Z_i - A - BF_i)^2$$

is distributed on the null hypothesis approximately as  $\chi^2$  with degrees of freedom equal to (the number of different classes - 2). This gives a simple and useful check on the correctness of the hypothesis expressed by equation (1).

The calculation can be done fairly conveniently manually, or it can be straightforwardly programmed on an electronic computer. The University College London computer was used to analyse the data of Table 3 of the paper of Kumar *et al.*, and the following values were found (Table 5).

Table 5.

Type of mortality	$A$	$B$	$\log. (B/A)$	$B/A$	$\chi^2 (1.D.F.)$
Still births and miscarriages	0.032 ± 0.0065	1.07 ± 0.18	1.52 ± 0.095	33	4.0
Infant and juvenile deaths	0.091 ± 0.011	0.90 ± 0.21	1.00 ± 0.21	10	0.1
Total mortality	0.13 ± 0.013	2.27 ± 0.31	1.25 ± 0.06	18	0.9

These differ appreciably from the values given by Kumar *et al.* and suggest an estimate of around 2.5 lethal equivalents per gamete, or 5 per zygote, rather than the 7 found by them.

