

TESTING A MODEL OF DOMINANT INHERITANCE FOR METRIC TRAITS IN MAN

I. BARRAI, L. L. CAVALLI-SFORZA

*Laboratorio Internazionale di Genetica e Biofisica, Sezione di Pavia,
c/o Istituto di Genetica, Università' di Pavia*

and

M. MAINARDI

Istituto di Genetica, Parma

Received 10.v.64

1. INTRODUCTION

THE present paper contains an analysis of two metrical traits in man (stature and chest girth) based on joint analysis of the inbreeding effects and of sib and twin correlations. The material consists of data collected at conscription in the region of the Parma Valley (Northern Italy) on 20-year-old males born between 1892 and 1911, excluding the town of Parma. Part of these data was already partially analysed (Barrai, Cavalli and Mainardi, 1960) and the analysis showed that an increase in sample size of the inbred fractions would be desirable. The sample size was therefore doubled.

The effect of inbreeding on metrical traits has been considered theoretically by several authors. We shall give below a treatment in such a form that the results can be more easily linked with those of sib-correlations in random mating populations.

2. EFFECT OF INBREEDING ON METRIC TRAITS

Mather (1949) gives a model for the study of metrical traits in diploid populations. He deals with random mating populations, and his methods are limited to situations in which it is possible to gather information on at least two generations.

We use here an extension of his model to populations with known inbreeding levels, to study models of transmission with data pertaining to a single generation (Dickinson and Jinks, 1956).

Consider, in a diploid population, a diallelic locus $\begin{pmatrix} A, a \\ p, q \end{pmatrix}$, and let the metric values of the three possible genotypes be $+d$ for AA, h for Aa, and $-d$ for aa. Let F be the inbreeding coefficient. The effects of dominance are expressed by the value of the parameter h , which varies between $+d$ and $-d$, being zero when the effects of the alleles are additive. The model may be summarised as follows:—

Genotype	Frequency	Metric value
AA	$p^2 + pqF$	$+d$
Aa	$2pq(1-F)$	h
aa	$q^2 + pqF$	$-d$

The mean of the metric trait in this population will be:

$$m_F = (p-q)d + 2pqh(1-F) \quad (1)$$

and the following may be defined:

a—mean of a random mating population

$$m_R = (p-q)d + 2pqh \quad (2)$$

b—mean of a homozygous population

$$m_I = (p-q)d \quad (3)$$

Rearranging (1), (2) and (3), we find

$$m_F = m_R - (m_R - m_I)F \quad (4)$$

equivalent to

$$m_F = m_R - 2pqhF. \quad (5)$$

The form (4) is due to Wright (1921) and was used by Morton (1958). We prefer the form (5) because it shows explicitly the effect of dominance. A test for significance of the linear component of regression of the means of the trait on *F* is a test for dominance effects.

From the model, the variance is derived as:

$$\begin{aligned} \sigma_F^2 = & 2pq[d^2 + h^2 - 2pqh^2 - 2(p-q)dh] \\ & + 2pq[d^2 - h^2 + 4pqh^2 + 2(p-q)dh]F - 4p^2q^2h^2F^2 \end{aligned} \quad (6)$$

which is a quadratic function of *F*.

Using Mather's symbols, and omitting the summation over all genes

$$\begin{aligned} D &= 4pq[d + h(q-p)]^2 \\ H &= 16h^2p^2q^2 \end{aligned}$$

and introducing the symbol:

$$D' = 4pqd^2 \quad (\text{Jinks, 1954})$$

equation (6) simplifies to:

$$\sigma_F^2 = \frac{1}{2}D + \frac{1}{4}H + F(D' - \frac{1}{2}D) - \frac{F^2}{4}H.$$

For a random mating population:

$$\sigma_R^2 = \frac{1}{2}D + \frac{1}{4}H$$

and therefore (6) can be written:

$$\sigma_F^2 = \sigma_R^2 + F(D' - \frac{1}{2}D) - \frac{F^2}{4}H.$$

The relevant quantity D' could thus be estimated if suitable data were available. If there is no dominance, $H = 0$ and $D = D'$, and therefore

$$\sigma_F^2 = \sigma_R^2 + \frac{F}{2} D' = \frac{1}{2} D' (1 + F) = \sigma_R^2 (1 + F)$$

so that the variance is linear with F and the slope of the regression is equal to $\frac{1}{2} D'$.

The test of linearity of the variance on F , however, would be a poor test of the absence of dominance. In fact, even with complete dominance, ($h = d$) and gene frequencies maximising H , ($p = q$) the deviation from linearity would be hardly detectable, being of the order of 3 per cent. at an inbreeding coefficient as high as $1/8$, one of the highest that can be observed in Man.

3. MATERIALS AND METHODS

We have a practically complete and up-to-date list of the consanguineous marriages that took place in the diocese of Parma in northern Italy. Results of the anthropometric examinations for stature and chest girth of conscripts from the same region in the period 1912-32 were made available for this study.

Since military service is compulsory for physically fit individuals and such fitness is ascertained by medical examination carried out at the same time as the anthropometric examination, our sample contains practically all the males of 20 years of age living in the area.

The conscription records also give information on the surnames and christian names of the conscripts' parents. By matching them with the lists of consanguineous marriages of the region, it was possible to ascertain 1556 inbreds out of a population of about 35,000 conscripts (4.4 per cent.). This corresponds reasonably well to the proportion of consanguineous marriages observed in the region in a comparable period (5.4 per cent.).

The inbreeding coefficient of each individual was punched on cards along with other information available in addition to stature and chest girth, forming for each individual a vector with the following nine variates:

- y_1 = Stature as recorded at the examination. It is expressed in centimetres, to the nearest unit.
- y_2 = Chest girth as recorded at the examination. It is measured immediately under the nipples: expressed in centimetres to the nearest unit.
- F = Inbreeding coefficient of the individual.
- x_1 = Age in months at the examination.
- x_2 = Altitude on sea level of the birthplace measured in metres (to be considered as an arbitrary environmental parameter).
- x_3 = Year of birth. The oldest individuals were born in 1892 and the youngest in 1912.
- x_4 = Migration, defined as a binary variate having value 1 if the place of residence was different from the place of birth, zero otherwise.
- x_5 = Rurality. A binary variate having value zero if the profession declared at examination was "farmer"; 1 if other.
- x_6 = Literacy. A binary variate having value 1 if the individual examined was able to read, zero otherwise.

The six concomitant variables, x_i , take into account some environmental factors variously associated with consanguinity and the dependent variables, y_i . We have but little knowledge of the action of physical factors on the phenotypic traits we are

considering, and little also is known on the effect of social factors. Factors like x_4 , x_5 , x_6 , even if coded by binary variates should provide some information on these points.

It is known that consanguinity varies in time and space and with social conditions in this area (Cavalli-Sforza, 1959). Stature and chest girth also show geographical variation (Conterio and Cavalli-Sforza, 1959) and therefore a proper choice of non-consanguineous controls was necessary.

1646 non-consanguineous controls were chosen at random from the same population for comparison with the consanguineous group. "Non-consanguineous" is taken here to indicate a relationship between parents more remote than third cousins. The F value for controls was taken to be zero. Actually, a small fraction of controls might arise from undetected consanguineous marriages, if the marriage of the parents took place outside the Parma Diocese. However, considering the low mobility of this population, such a fraction must be exceedingly small.

Controls were chosen with suitable stratification by year and residential unit ("comuni") so as to match the consanguineous group. Even so, one cannot hope to have equalised all possible differences in social factors between controls and the various consanguinity groups. For this reason adjustment by covariance analysis was carried out, using all the information given before on both consanguineous and control groups.

In addition, from the first half of the time interval considered (1892-1901), paternity and maternity data of all 16,997 conscripts of this period were used for the recognition of sibs and twin pairs among conscripts. This was done by punching on cards for each conscript the first two letters of the surname and christian name of the conscript's father and mother, then sorting and collating these cards, finally checking on the original records if the full paternity and maternity were the same. Sibs born on the same day were considered twins; 25 twin pairs were found. Sib data were not corrected for concomitant variables.

(i) *Analysis of variance and regression (uncorrected variates)*

The distribution of the means of stature and chest girth on the value of F is given in table 1; it is apparent that the means are not a

TABLE 1

Distribution of mean stature and chest girth on degrees of consanguinity

F	Number of individuals	Mean stature	Variance	Mean chest girth	Variance
0	1646	165.46 ± 0.16	40.03	84.93 ± 0.11	21.16
1/256	264	167.14 ± 0.42	46.56	85.62 ± 0.31	25.39
1/128	240	165.91 ± 0.42	42.37	85.03 ± 0.31	22.12
1/64	620	165.20 ± 0.25	37.95	85.03 ± 0.19	22.08
1/32	120	164.89 ± 0.53	34.03	84.99 ± 0.47	26.43
1/16	297	165.29 ± 0.42	52.32	84.28 ± 0.27	21.76
1/8	15	164.87 ± 1.90	54.12	83.07 ± 0.91	11.92
All	3202	165.54 ± 0.11	41.51	84.95 ± 0.08	22.02

linear function of the inbreeding coefficient. The most puzzling fact is that the means for the group of individuals born to third cousins' marriages are higher than the other values, and we do not see any immediate reason for this, except from social factors that we were

unable to identify with certainty. The means of the groups, both for stature and chest girth are significantly different as it is shown by the analysis of variance in tables 2 and 3.

TABLE 2
Analysis of variance of stature, y_1

Source of variation	J_1^2	d.f.	S^2	F	P
Total	132874.292	3201			
Between groups with different consanguinity	860.710	6	143.452	3.742	< 1 per cent.
Regression on consanguinity	69.484	1	69.484	1.682	< 30 per cent.
Deviation from linearity	791.226	5	158.245	3.380	< 30 per cent.
Residual	132013.582	3195	41.319		

For stature the sum of squares removed by a simple regression on F is not significant: while highly significant is indeed the deviation of the means from a straight line. The heterogeneity between means of chest girth on the other hand has a component linear with F which,

TABLE 3
Analysis of variance of chest girth, y_2

Source of variation	Σy_2^2	d.f.	S^2	F	P
Total	70507.600	3201			
Between groups with different consanguinity	310.980	6	51.830	2.359	< 5 per cent.
Regression on consanguinity	157.217	1	157.217	7.156	< 1 per cent.
Deviation from linearity	153.763	5	30.753	1.400	< 30 per cent.
Residual	70196.620	3195	21.971		

tested on error variance, removes a highly significant sum of squares. The regression coefficient ($b = -0.04365$) is negative, thus indicating that the factors responsible for "large" chest girth are dominant. There is no significant deviation from linearity.

Stature and girth being highly correlated, one may look for the linear combination of them which maximises the sum of squares removed by a term linear in F. Such a function was found to be

$$L = y_1 + 18.622y_2.$$

But the gain obtained in comparison to chest girth alone is trivial.

TABLE 4
Correlations among stature, chest girth, consanguinity and the concomitant variables

	Chest girth	Consanguinity	Age	Altitude of birth place	Year of birth	Migration	Non-rurality	Literacy
Stature	0.424	-0.023	0.146	0.042	0.052	-0.004	0.114	0.074
Chest girth		-0.047	0.194	0.120	0.078	-0.005	0.009	0.052
Consanguinity			0.028	0.109	0.043	-0.113	-0.048	0.017
Age				0.094	0.180	-0.066	0.019	0.055
Altitude of birth place					0.220	-3.35	-0.056	0.017
Year of birth						-0.069	0.018	0.185
Migration							-0.033	-0.076
Non-rurality								0.147

With 3200 degrees of freedom, a value of $r = .035$ is significant at the 5 per cent. level.

(ii) *Adjustment of means for environmental variables*

The correlations among the variables considered are shown in table 4. Consanguinity is not significantly correlated with stature, age and literacy: the correlation with chest girth, migration and rurality is negative. There is a strong increase in consanguinity with altitude on sea level, and the year of birth is significantly associated

TABLE 5

Distribution of mean stature and chest girth on social condition and consanguinity

F	n	Stature	Chest girth
Farmers			
0	787	164.78 ± 0.24	85.10 ± 0.17
1/256	135	165.53 ± 0.55	84.57 ± 0.42
1/128	120	166.00 ± 0.68	84.85 ± 0.44
1/64	333	164.56 ± 0.34	85.21 ± 0.25
1/32	59	163.75 ± 0.72	84.37 ± 0.67
1/16	168	164.64 ± 0.58	84.01 ± 0.37
1/8	8	161.00 ± 2.08	82.50 ± 1.10
Total	1610	164.82 ± 0.17	84.91 ± 0.12
Non-farmers			
0	859	166.09 ± 0.21	84.77 ± 0.15
1/256	129	168.82 ± 0.60	86.72 ± 0.43
1/128	120	165.82 ± 0.50	85.22 ± 0.42
1/64	287	165.95 ± 0.36	84.83 ± 0.29
1/32	61	166.00 ± 0.76	85.59 ± 0.66
1/16	129	166.15 ± 0.60	84.64 ± 0.39
1/8	7	169.29 ± 2.49	83.71 ± 1.49
Total	1592	166.28 ± 0.15	84.99 ± 0.12

with inbreeding. The other variables are variously intertwined, so that equalisation of stature and girth for their effects within consanguinity groups is necessary. In table 5 we have a striking example of the effects of social conditions on the dependent variables. The data are split into two groups, one formed by the individuals whose profession was declared as "farmer" at the examination, the other by all non-farmers. The non-farmers group has an average stature significantly higher than the farmers group ($F = 42.165$ from an analysis of variance); chest girth does not differ significantly between the two groups. It is apparent that stature is extremely sensitive to variation in ecological conditions, whereas chest girth looks rather insensitive to such variation.

We need the value of the means of stature and chest girth within groups of different consanguinity, adjusted for the six concomitant variables. The appropriate correction is given by the formula

$$\bar{y}'_i = \bar{y}_i - \sum_j b_j(\bar{x}_{ij} - \bar{\bar{x}}_j)$$

where *i* indicates the consanguinity group, and *j* the *j*th variate. The *b_j*s are partial regression coefficient obtained using the variance-covariance matrix within groups. The regression coefficients, their

TABLE 6
Partial regressions within groups

	Stature <i>b_i</i> + <i>S_{b_i}</i>	<i>t</i>	Chest girth <i>b_i</i> + <i>S_{b_i}</i>	<i>t</i>		
Age	0.120484 ± 0.014285	8.434**	0.110158 ± 0.010357	10.636**		
Altitude	0.007557 ± 0.004234	1.785	0.017078 ± 0.003069	5.565*		
Year	0.013244 ± 0.020108	0.659	0.011131 ± 0.014579	0.763		
Migration	0.542990 ± 0.334166	1.625	0.180213 ± 0.242281	0.744		
Non-rurality	1.409815 ± 0.226130	6.235**	0.073796 ± 0.163956	0.450		
Ability to read	1.213520 ± 0.331444	3.661*	0.773050 ± 0.240308	3.217*		
Analysis of variance						
S. of V.	Σ <i>y</i> ₁ ²	d.f.	S ²	Σ <i>y</i> ₂ ²	d.f.	S ²
Within groups	132013.582	1195		70196.620	3195	
Regression	5444.490	6		3682.260	6	
Residual within	126569.092	3189	39.6893	66514.360	3189	20.8574

errors and *t*s, together with the analysis of variance of the regression are given in table 6.

As seen from the tests of significance, age, year of birth, rurality and literacy have an effect on stature. The increase due to age is easily understood; however, considering the age of the individuals, this effect is small, because at about twenty years the growth curve approaches an asymptote. The effect of the year of birth is expected, because of the well-known recent trend to the increase in stature (Conterio and Cavalli-Sforza, 1959). The effect of non-rurality is approximately of the same order as that of the ability to read: both these factors are diagnostic of a more favourable environment. Age, year of birth, rurality and literacy are significantly associated with girth, and the considerations to be made are similar to those given for stature; however, altitude also has an effect in increasing chest girth; even if this effect is small, it is definitely significant. This increase might reflect an adaptive adjustment to diminished oxygen pressure.

(iii) *Regression of the adjusted means on F*

The variance of the adjusted means is

$$\text{Var}(\bar{y}_i) = \{1/n_i + \Sigma(\bar{x}_{ij} - \bar{x}_j)^2/SS_{x_j}\}s_w^2$$

s_w^2 is the error variance, obtained from the residual SSD after fitting the multiple regression using the dispersion matrix within groups.

The inverse of the variance of the adjusted means was used to weight their regression on consanguinity degree. The resulting sums

TABLE 7
Regression of adjusted means on consanguinity

F	Stature		Chest girth	
	Adjusted mean	Weight, $w = \frac{1}{s^2}$	Adjusted mean	Weight, $w = \frac{1}{s^2}$
0	165.51 ± 0.16	38.9051	85.09 ± 0.12	74.0099
1/256	167.10 ± 0.39	6.5052	85.41 ± 0.28	13.3750
1/128	165.90 ± 0.41	5.9714	84.90 ± 0.30	11.3595
1/64	165.16 ± 0.26	15.0100	84.82 ± 0.19	28.5538
1/32	164.79 ± 0.58	3.0069	84.78 ± 0.42	5.7201
1/16	165.37 ± 0.37	7.4458	84.22 ± 0.27	14.1671
1/32	165.00 ± 1.63	0.3776	83.04 ± 1.16	0.7183

Analysis of χ^2 s from the regression						
S. of V.	d.f.	χ^2	P	d.f.	χ^2	P
Total	6	20.787	< 1 per cent.	6	14.910	< 5 per cent.
Regression	1	1.662	< 30 per cent.	1	12.955	< 1 per cent.
Deviation from linearity	5	19.128	< 1 per cent.	5	1.955	< 90 per cent.

of squares $\Sigma w_i(\bar{y}'_i - \bar{y}_i)^2$ are distributed approximately like χ^2 , because the quantity w_i is the inverse of the variance of \bar{y}'_i .

The results of the adjustments of the means, their weights and the analysis of variance of the regressions are given in table 7, and the adjusted means are plotted against F in figs. 1 and 2.

The analysis is given, in this case, by a partition of chi-squares. The deviations from the grand mean are significant both for stature and chest girth; however, the regression of the adjusted means of stature on consanguinity is not significant. Thus, no factor that can account for the peculiar behaviour of this trait has been isolated even after removing environmental variables. Deviation from linearity is very erratic, and no indication on the mode of transmission was found in the data under analysis.

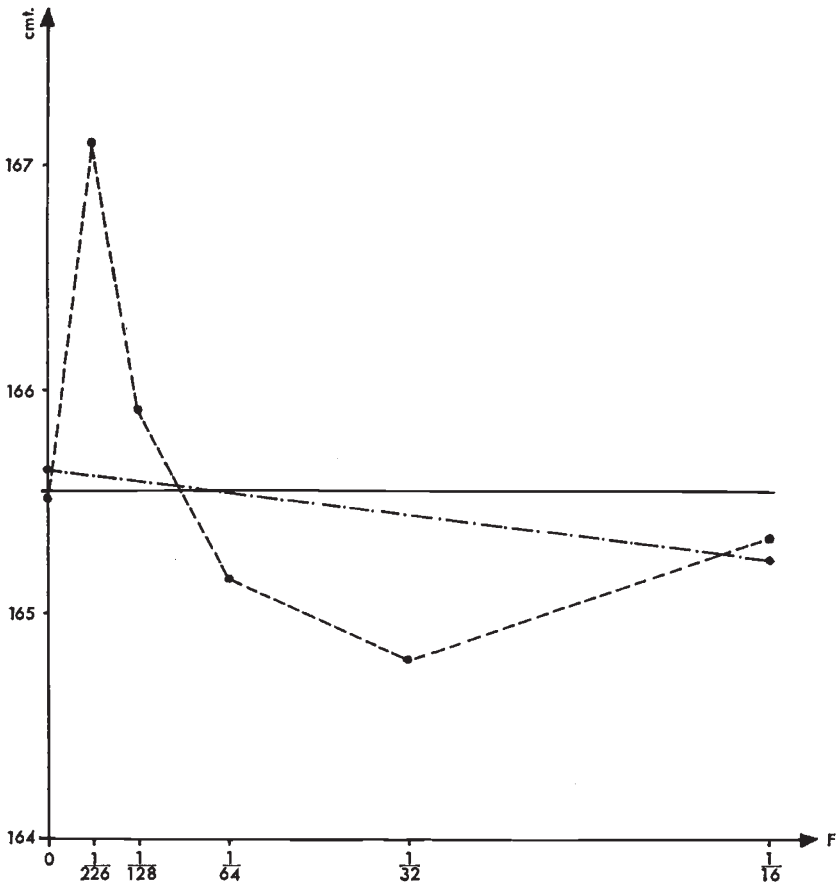


FIG. 1.—Regression of corrected means of stature on F.

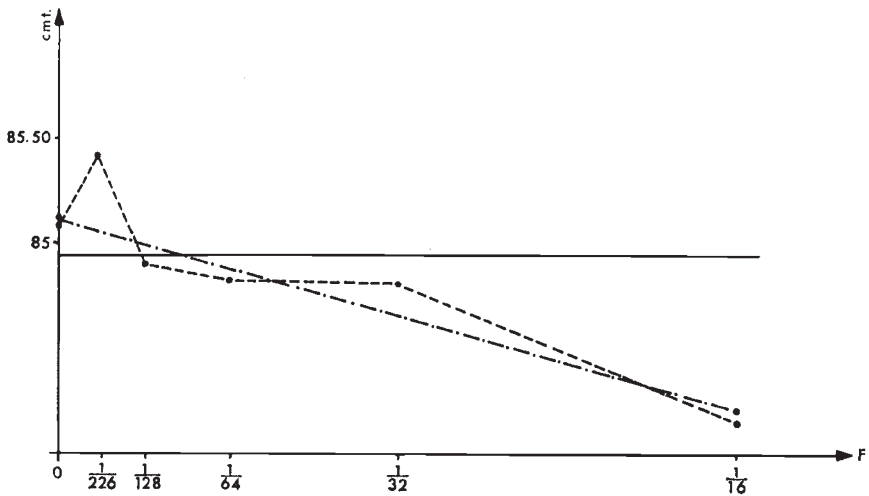


FIG. 2.—Regression of corrected means of chest girth on F.

The situation is quite different for chest girth. Most of the heterogeneity between the adjusted means of the groups is removed by the regression on F, leaving a residual chi-square for deviations from linearity which is not significant; the regression itself is significant at the 1 per cent. level. There is a strong indication of dominance of "plus" alleles in the mode of transmission of the factors determining the width of chest girth. The regression equations are:

$$\bar{y}'_i = 165.65 - (0.02848 \pm 0.02209)F \text{ for stature}$$

$$\bar{y}'_i = 85.12 - (0.05766 \pm 0.01602)F \text{ for chest girth.}$$

We tested the heterogeneity of variance in groups of different consanguinity; the results of Bartlett's test are as follows:

	d.f.	χ^2	P
Stature	6	16.672	1 per cent.
Chest girth	6	8.274	< 20 per cent.

The variances for stature are heterogeneous, and do not follow a pattern indicative of either a first or a second degree function. No information would be obtained for our purposes by fitting a polynomial of higher order. The variances of girth do not differ significantly in the various groups, so the fitting of any function is arbitrary.

(iv) *Sibship data*

We also had available the whole group of conscripts born from 1892 to 1901. By grouping together individuals who had the same

TABLE 8

Analysis of variance of stature between sibships of different size

S. of V.	SSD	d.f.	S ²	F	P
Total	300525.61	7615	39.46		
Between sizes	153.68	2	76.84	3.210	< 5 per cent.
Between sibships within sizes	198422.82	3356	59.12	2.470	< 1 per cent.
Within sibships within sizes .	101949.11	4259	23.94		

father and mother we obtained 2607 sibships of size two, 668 of size three and 123 of size four; 3 sibships of larger size were neglected. Analysis of variance was then performed for chest girth and stature between and within sibships.

The results of the analysis are given in tables 8 and 9 and analysis of twins (sibs born on the same day) on table 10. The degrees of freedom are different for stature and chest because in the original

TABLE 9

Analysis of variance of chest girth between sibships of different size

S. of V.	SSD	d.f.	S ²	F	P
Total	170593·02	7709	22·13		
Between sizes	369·38	2	164·69	9·241	<0·5 per cent.
Between sibships within sizes	93440·04	3398	27·52	1·544	
Within sibships	76823·60	4312	17·82		

TABLE 10

Analysis of variance of stature and chest girth in twins

S. of V.	Stature				Chest girth			
	SSD	d.f.	S ²	F	SSD	d.f.	S ²	F
Total	1884·80	49	38·47		617·38	49	12·60	
Between pairs	1648·79	24	69·70	7·27	380·88	24	15·87	1·678
Within pairs	236·01	25	9·44	P < 1 per cent.	236·50	25	9·46	P < 10 per cent.

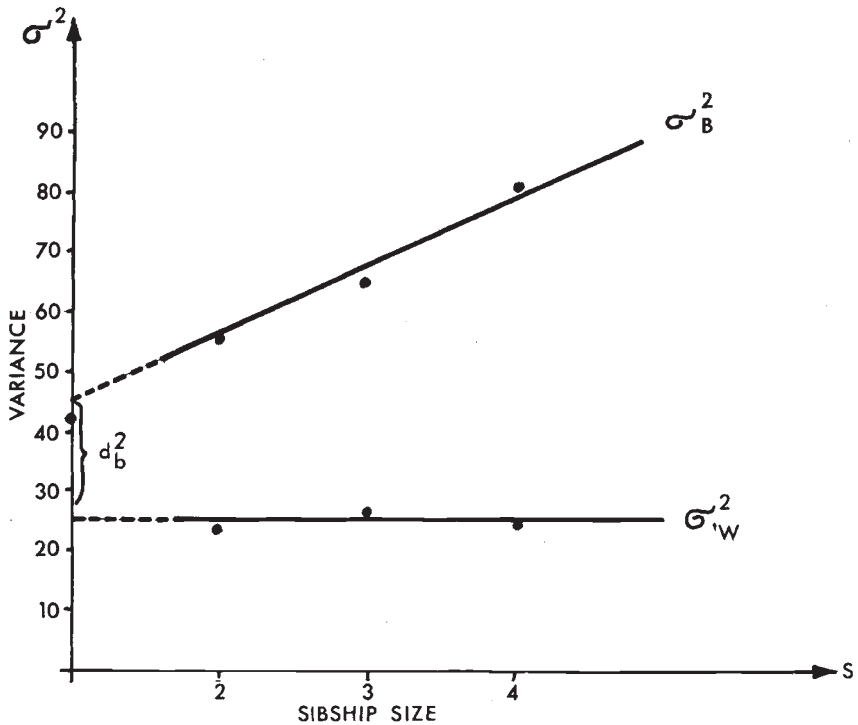


FIG. 3.—Stature variances and sibship size.

source of data one of the two items was occasionally missing (or incomprehensible). Between and within sibship variances are plotted against sibship size in figs. 3 and 4. Within sibship variances are independent of family size, and between sibship variances increase linearly with it,

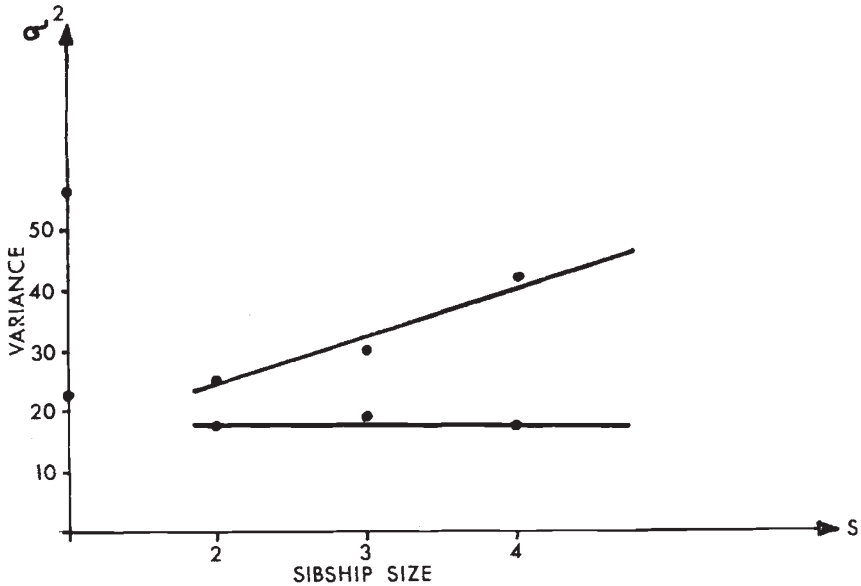


FIG. 4.—Chest girth variances and sibship size.

as expected, the difference between the intercepts of the two regression lines at sibship size = 1 being the estimate for σ_B^2 , the variance between sibships corrected for sampling errors. The corrected variances between sibs and within sibs (and their errors) were estimated by maximum likelihood as given in the appendix.

TABLE 11

Variation of means with sibship size

Sibship size	Stature	Chest girth
1	164.85 ± 0.07	84.32 ± 0.05
2	164.82 ± 0.09	84.36 ± 0.06
3	165.03 ± 0.14	84.73 ± 0.11
4	165.33 ± 0.28	85.04 ± 0.22

The estimates of the corrected variance between sibships, and the average variances within sibships, as well as the “heritability” estimates derived from them by the usual formula

$$\frac{1}{2}h^2 = \frac{\sigma_B^2}{(\sigma_B^2 + \sigma_w^2)}$$

are given in table 12.

Using Mather's equations to analyse in terms of D and H (see introduction) we obtain for chest girth the value:

$$\sigma_w^2 = 1/4D + 3/16H + E_1 = 18.35 \pm 0.34 \quad (9)$$

for the variance within sibships, and

$$\sigma_B^2 = 1/4D + 1/16H + E_2 = 4.73 \pm 0.32 \quad (10)$$

for the variance between sibships, where E_1 and E_2 are the within and between environmental variances.

TABLE 12

	Stature	Chest girth
Variance between sibs, corrected (σ_B^2) . . .	15.88 ± 0.57	4.73 ± 0.32
Variance within sibs, corrected (σ_w^2) . . .	24.15 ± 0.59	18.35 ± 0.34
Heritability	0.397 × 2	0.205 × 2

Assuming that in our small sample of twins there is an equal number of mono- and dizygotic pairs, as expected (approximately) for like-sexed twin pairs in Italy, we can obtain another equation from them. In monozygotic pairs there will be only environmental variance, which one can assume as a first approximation to be the same as in full sibs, and in dizygotic pairs the components will be as in full sibs. The variance within twin pairs σ_T^2 shall be:

$$\begin{aligned} 1/2E_1 + 1/2(1/4D + 3/16H + E_1) &= \sigma_T^2 = 9.46 \\ \text{or } 1/8 + D/32 + H/32 + E_1 &= 9.46. \end{aligned} \quad (11)$$

Associating (9) and (11) we find

$$E_1 = 2\sigma_T^2 - \sigma_w^2 = 18.92 - 18.35 = 0.57 \pm 2.73$$

a surprisingly small value for E_1 . However, this is subject to a large error because of the small number of twin pairs. We have no way to estimate E_2 , so that we will solve for D and H the system:

$$\begin{aligned} 1/4D + 1/16H + E_2 &= 4.73 \pm 0.32 \\ 1/4D + 3/16H &= 18.35 - 0.57 = 17.78 \pm 2.75 \end{aligned}$$

obtaining D and H as functions of E_2 :

$$\begin{aligned} D &= -7.18 - 6E_2 \\ H &= 104.40 + 8E_2. \end{aligned}$$

Any positive value of E_2 , the environmental variance between families, would inflate H and decrease D. On the assumption that E_2 is zero, we have:

$$\begin{aligned} D &= -7.18 \pm 6.00 \\ H &= 104.40 \pm 22.40. \end{aligned}$$

In practice, D is not significantly different from zero while H is significantly different from zero and from D .

For stature, we obtain by the same procedure:

$$E_1 = -5.27 \pm 2.75$$

$$D = 36.44 - 6E_2 \pm 23.28$$

$$H = 108.32 - 8E_2 \pm 6.88.$$

It is reasonable to assume that E_2 is greater than zero: however, we are interested in the considerations that may be made on the ratio H/D .

A positive value of E_2 would inflate H and decrease D . If we take E_2 equal zero, the ratio H/D will be estimated by defect. From the formulæ given in the introduction and considering for simplicity a single locus,

$$\frac{H}{D} = \frac{16h^2p^2q^2}{4pq[d+h(q-p)]^2}.$$

Excluding overdominance, the upper limit of h is d , and in this case the ratio $H/D = p/q$ measures the relative frequency of the dominant allele with respect to the recessive.

We can conclude that in the case of chest girth, D must be small, because its estimate is in our analysis smaller than zero; and H is much larger. Therefore, alleles responsible for greater size of chest girth should be more frequent than those for small size.

The same could be said for stature, where there also seems to be an important dominance component. However, the consanguinity data seem to show that the average dominance ("potency" according to Wigan, *i.e.* Σh) must be close to zero.

4. SUMMARY

1. A combination of the two approaches, correlation between sibs (or twins) and inbreeding effect has been tried on two metrical data in Man, by taking stature and chest girth as measured at conscription on appropriate samples drawn from some 35,000 conscripts born between 1892 and 1911 in a rural fraction of the Parma region.

2. From these, a total of 1556 inbred individuals could be obtained. A depression due to inbreeding was found for chest girth only; the character varied linearly with F , as expected if "plus" alleles determining large chest girth were dominant over their "minus" alleles.

3. In good agreement with this, the H values for chest girth measuring the sum of squares of dominance effects, determined from correlations between sibs and twins, was very large. The D values (according to Mather's definition) was close to zero. This might be explained assuming that "plus" alleles are on average more frequent than "minus" alleles in the population.

4. With stature the inbreeding effect is not observed. There is instead an unexplained peak of stature in the progeny of third cousins. The explanation of the latter fact is probably that the progeny of third cousins must be made, in part, of a socially highly selected fraction of the population, even though we had no direct evidence of this.

5. Also with stature the H value from sib data is relatively high. If there is dominance in the factors determining stature, therefore, it must be well balanced, namely there must be an approximately equal frequency of "plus" and of "minus" dominant factors.

6. Various environmental factors were found to affect the measurements considered. There is a trend towards the increase with time for both stature and chest girth in our data; but there are also effects of farming or non-farming profession and other indicators of socio-economic conditions. Chest girth shows also a clear cut increase with altitude. The observed effect of sibship is probably indirect.

7. It has been shown that, theoretically, the analysis of variance as a function of F could prove a very useful tool, used in conjunction with sib analysis.

5. APPENDIX

Estimation of variances between and within sibship

A variance between sibs for sibships of size k as estimated by the analysis of variance as a "between sibs" variance has an expected value:

$$k\theta_1 + \theta_2$$

where θ_1 is the variance between sibships, and θ_2 is the variance within sibships.

Application of the method of maximum likelihood permits the estimation of θ_1 and θ_2 and their errors including also the data of sibships of size one.

SS_k will be defined as the sum of squared deviations "between sibs" in the standard analysis of variance for sibships of size k , and will have n_k degrees of freedom. SS_1 is the sum of squares for individuals from sibships of size one. SS_0 is the sum of squares "within sibs" from the analysis of variance of all sibship sizes ≥ 2 .

Assuming that variances have been estimated from a normal population, the quantity

$$\frac{SS_k}{k\theta_1 + \theta_2} = u$$

is distributed like a Gamma variate, *i.e.* with probability proportional to:

$$u^{\frac{n}{2}-1} e^{-u/2}$$

the log likelihood for each SS_k will be

$$L = \binom{n}{2} - 1 \log u - \frac{u}{2}.$$

Trial values for θ_1 and θ_2 can be derived by some approximate method, e.g. averaging the values obtained from

$$V_{Btw} = k\theta_1 + \theta_2$$

for each k value, θ_2 being obtained from the pooled within sibs variance.

Using the trial values θ_{10} , θ_{20} , scores will be computed from the formulæ:

$$S(\theta_1) = -\frac{1}{2} \left(\sum_{k=1}^{\infty} \frac{k(n-2)}{k\theta_1 + \theta_2} - \sum_{k=1}^{\infty} \frac{kSS_k}{(k\theta_1 + \theta_2)^2} \right)$$

$$S(\theta_2) = -\frac{1}{2} \left(\sum_{k=1}^{\infty} \frac{n-2}{k\theta_1 + \theta_2} - \sum_{k=1}^{\infty} \frac{SS_k}{(k\theta_1 + \theta_2)^2} \right)$$

and the information matrix from

$$I(\theta_{11}) = \sum_{k=1}^{\infty} \frac{(n+2)}{(k\theta_1 + \theta_2)^2} / \Delta$$

$$I(\theta_{12}) = \sum_{k=1}^{\infty} \frac{k(n+2)}{(k\theta_1 + \theta_2)^2} / \Delta$$

$$I(\theta_{22}) = \sum_{k=1}^{\infty} \frac{k^2(n+2)}{(k\theta_1 + \theta_2)^2} / \Delta$$

where

$$\Delta = \left(\sum \frac{k^2(n+2)}{(k\theta_1 + \theta_2)^2} \right) \left(\sum \frac{n+2}{(k\theta_1 + \theta_2)^2} \right) - \left[\sum \frac{k(n+2)}{(k\theta_1 + \theta_2)^2} \right]^2$$

The corrections of the trial values will be obtained from

$$\delta\theta_{10} = S(\theta_1)I(\theta_{11}) + S(\theta_2)I(\theta_{12})$$

$$\delta\theta_{20} = S(\theta_1)I(\theta_{12}) + S(\theta_2)I(\theta_{22})$$

recycling if the scores obtained with the corrected values:

$$\theta_{11} = \theta_{10} + \delta\theta_{10} \pm \sqrt{I(\theta_{11})}$$

$$\theta_{21} = \theta_{20} + \delta\theta_{20} \pm \sqrt{I(\theta_{22})}$$

are not sufficiently close to zero.

Acknowledgments.—We thank Dr. M. Kimura for reading this manuscript. The support of this research by Euratom—C.N.R.—C.N.E.N. contract no. 012-61-12 BIAI, the U.S. Atomic Energy Commission, and the Rockefeller Foundation is gratefully acknowledged.

6. REFERENCES

- BARRAI, I., CAVALLI-SFORZA, L. L., AND MAINARDI, M. 1960. Studio pilota per la determinazione degli effetti della consanguineità su caratteri esaminati alla visita di leva. *Atti Ass. Genet. Ital., Pavia*, 5, 313-331.
- CAVALLI-SFORZA, L. L. 1959. Some data on the genetic structure of human populations. *Proc. X Int. Cong. Genetics*, 1, 388.
- CONTERIO, F., AND CAVALLI-SFORZA, L. L. 1959. Evolution of the human constitutional phenotype: an analysis of mortality effects. *Supplemento alla Ricerca Scientifica*, anno 29, 1959.
- DICKINSON, A. G., AND JINKS, J. L. 1956. A generalised analysis of diallel crosses. *Genetics*, 41, 65-78.
- JINKS, J. L. 1954. The analysis of continuous variation in a diallel cross of nicotiana rustica varieties. *Genetics*, 39, 767-788.
- KEMPTHORNE, O. 1957. *An Introduction to genetic statistics*. J. Wiley and Sons, N.Y.
- MATHER, K. 1949. *Biometrical Genetics*. Methuen and Co., London, 1-162.
- MORTON, N. E. 1958. Empirical risks in consanguineous marriages: birth weight, gestation time, and measurements of infants. *Am. J. Hum. Genet.*, 10, 344-349.
- WRIGHT, S. 1921. Systems of matings. II. The effect of inbreeding on the genetic composition of a population. *Genetics*, 6, 124-143.