Path analysis under generalized assortative mating

I. Theory

By D. C. Rao*, N. E. Morton*, and C. R. Cloninger †

Population Genetics Laboratory, University of Hawaii, Honolulu, Hawaii 96822

(Received 26 January 1979)

SUMMARY

Assortative mating is generalized to include social homogamy and phenotypic homogamy as two special cases. This generalization, called mixed homogamy, enables tests of hypotheses on the nature of assortment. Cultural inheritance is also extended to include two components: transmitted from parental environments, and non-transmitted sibship environment. Familial correlations are derived for a variety of relationships under mixed homogamy.

1. INTRODUCTION

Assortative mating is a system in which either like individuals (positive assortative mating) or unlike individuals (negative assortative mating) preferentially mate with each other. This preference may be based on two broad characteristics: phenotype, or social class including status, tastes, contacts and other aspects of group membership. Assortative mating may accordingly be classified into two major types: *phenotypic homogamy* and *social homogamy*. Under phenotypic homogamy the phenotypic correlation between spouses is primary, leading to secondary correlations between genotypes and environments of spouses. Under social homogamy, on the other hand, mates choose each other on the basis of their group membership which generates primary correlations between the genotypes and environments of spouses, and the phenotypic correlation between spouses becomes secondary. Models for phenotypic homogamy in the presence of cultural inheritance were developed and studied extensively by Sewall Wright (1978), Cloninger, Rice & Reich (1979) and Jencks (1972). We have so far concentrated on social homogamy (Morton & Rao, 1978; Rao & Morton, 1978), with major interest in physiological variables such as lipoprotein concentrations (Rao *et al.* 1979), hypertension (Morton *et al.* 1978), etc. So far, the choice between the two models of assortative mating was arbitrary. We here generalize assortative mating to incorporate phenotypic homogamy and social homogamy as two special cases, revise our model for cultural inheritance, and present a comprehensive

theory for the resolution of biological and cultural inheritance in arbitrary pedigrees. We call the generalized assortative mating *mixed homogamy*, which was briefly introduced in Morton & Rao (1979).

A subsequent paper will report applications of these methods to several bodies of data.

2. GENERALIZED ASSORTATIVE MATING

We assume that environment (C) acts additively with genotype (G) to produce a phenotype (P), all interactions being negligible. Fig. 1 displays the relationships between these variables for a couple, where subscripts F and M denote father and mother respectively. While dealing with phenotypic homogamy which is shown in Fig. 1 by a direct path (p) between P_F and P_M, we follow a recent technique of Wright (1978) and reverse the paths from G and C to P and replace the path coefficients by the correlation coefficients. We follow another technique of Cloninger *et al.* (1979) and avoid repetition of parental phenotypes. In Fig. 1 the solid lines represent phenotypic homogamy. Notice that the compound path coefficients γ and ϕ for the reversed paths are actually correlations of the phenotype with the genotype and indexed environment respectively:

$$\gamma = hz + cya, \quad \phi = cy + hza.$$

All these parameters are defined in Table 1. Total correlation between the genotype (G) and indexed environment (C) of an individual is denoted by a, a function of other parameters. This is split into two components under phenotypic homogamy: one direct correlational path between the two $(a - \gamma\phi)$, and another part via the phenotype (P) with a value of $\gamma\phi$, making the total $a - \gamma\phi + \gamma\phi = a$. Residual paths for the two spouses, not shown in Fig. 1, are assumed to be uncorrelated.

In Fig. 1 let us now superimpose social homogamy (H) onto phenotypic homogamy. Consequences of social homogamy are denoted by broken lines. Social homogamy has a direct path $\sqrt{m}$ to each of the two parental genotypes, and a direct path $\sqrt{u}$ to each of their indexed environments. This induces an indirect correlation $(s = \sqrt{mu})$ between G and C of an individual, and hence the direct correlation between G and C is shown as $a - s$ for either parent (by broken lines). It should be noted that this system of parallel paths in terms of solid and broken lines does not correspond to reciprocal causation (Wright, 1968). If cohabitation induces phenotypic correlation, this will be reflected by u. The extent of premarital resemblance of spouses due to phenotypic homogamy is measured by p.

In Fig. 1, all solid paths correspond to phenotypic homogamy, and all broken paths correspond to social homogamy. Notice that $p = 0$ gives social homogamy, and $m = u = 0$ gives phenotypic homogamy.

At this stage we introduce the following two conventions in reading broken paths for deriving familial correlations under mixed homogamy:

(i) Within an individual, either solid lines or broken lines are used but not both. For example, the correlation between P_F and C_F of Fig. 1 is: $\phi = cy + hza$ through solid lines and $cy + hz(a - \sqrt{(mu)}) + hz\sqrt{(mu)} = \phi$ through broken lines, conserving equality.
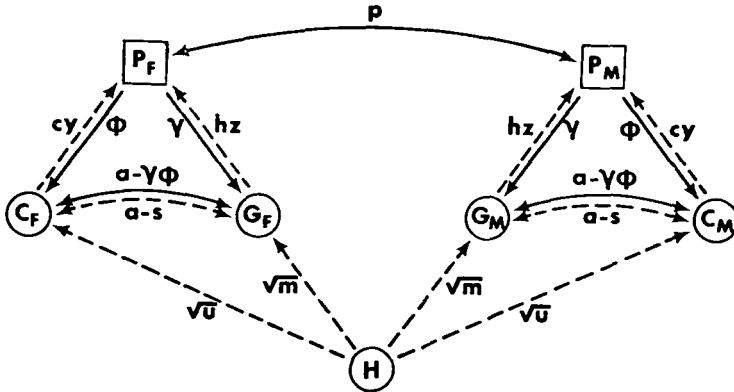
Fig. 1. Path diagram for assortative mating under mixed homogamy. P, G, C, H, denote phenotype, genotype, indexed environment and social homogamy respectively. Subscripts F and M denote father and mother. See Table 1 for definition of the parameters. $\gamma = hz + cya$, $\phi = cy + hza$, $s = \sqrt{(mu)}$.

Table 1. *Parameters of the mixed homogamy model*

Symbol

h effect of genotype on child's phenotype (square-root of 'heritability').

hz effect of genotype on adult's phenotype

c effect of child's indexed environment on the child's phenotype

cy effect of adult's indexed environment on the adult's phenotype

p primary correlation between parental phenotypes, not due to secondary resemblance through social homogamy (H)

m correlation between parental genotypes through social homogamy (H)

u correlation between parental indexed environments through social homogamy (H)

f_F effect of father's indexed environment on child's indexed environment

f_M effect of mother's indexed environment on child's indexed environment

b effect of non-transmitted common sibship environment on child's indexed environment

bx effect of non-transmitted common sibship environment on adult's indexed environment

i effect of child's indexed environment on the child's index.

i_F effect of father's indexed environment on father's index

i_M effect of mother's indexed environment on mother's index

Derived:

$\gamma = hz + cya$, correlation between the genotype and phenotype of an adult

$\phi = cy + hza$, correlation between the indexed environment and phenotype of an adult

$s = \sqrt{mu}$, correlation between an adult's indexed environment and spouse's genotype under social homogamy

$a =$ correlation between genotype and indexed environment of an individual

$\psi = f_F^2 + f_M^2 + 2f_F f_M (u + p\phi^2)$, correlation between indexed environments of sibs derived from parental environments.

(ii) Between individuals, any given chain may contain either the solid paths or the broken paths between mates but not both. For example, the correlation between P_F and P_M involves five chains in all:

direct path between P_F, P_M p
through G_F, G_M $(hz)^2 m$
through C_F, C_M $(cy)^2 u$
through G_F, C_M & C_F, G_M $2hzcys$ $(s = \sqrt{(mu)})$

therefore

correlation between P_F, $P_M = p + (hz)^2 m + (cy)^2 u + 2\,hz\,cy\,s$

These conventions are equivalent to alternative conventions justified on statistical grounds (Cloninger, 1979 *a*, *b*).

3. THE MODEL

Fig. 2 displays the model for biological and cultural inheritance in nuclear families under mixed homogamy. As before (Rao & Morton, 1978) we retain specific maternal effects by distinguishing f_F and f_M where f denotes the effect of a parent's environment on that provided to a child the parent rears. Direct effects of parental phenotypes on the environment of a child are assumed negligible since they were not significant even for IQ (Rao & Morton, 1978). Intergenerational differences in genetic and environmental effects are maintained: whereas the genetic and environmental heritabilities in childhood are h^2 and c^2 respectively, they are $h^2 z^2$ and $c^2 y^2$ in adulthood. Separate indices (I_F, I_M) are retained for the parents.

Apart from mixed homogamy for assortative mating, the most significant deviation of this model from our previous models is in representing the sibship environment. Previously we postulated one sibship environment common for all members of the sibship, and estimated this by one index (average of the index values for all members of the sibship). Realizing that this is only approximate, we now introduce different environments for different members of a sibship which are partly transmitted from the previous generation (C_F, C_M) and partly determined by a non-transmitted sibship environment (B). Accordingly, we now have one index for each member of a sibship. This model is illustrated with two full sibs in Fig. 2. All the 14 functionally independent parameters are defined in Table 1. Note that the parameter x has a different meaning from Rao, Morton & Yee, (1976): Whereas the correlation between sibling environments due to non-transmitted factors is b^2 in childhood, it is $b^2 x^2$ for adult sibs. From Fig. 2, the correlation between a genotype and a child's indexed environment of any child (reared together) is

$$(a + s + p\phi\gamma)\ (f_F + f_M)/2 = a$$

assuming equilibrium. This correlation (a) is therefore not an independent para-

meter. Remembering that γ and ϕ also involve a, solution of the following quadratic equation yields a as a function of other parameters:

where

$$
\left.\begin{array}{c}
a^2 + a\beta + 1 + \dfrac{s}{phzcy} = 0, \\[2ex]
\beta = \dfrac{\lambda p\,(h^2 z^2 + c^2 y^2) - 1}{\lambda phzcy}, \\[2ex]
\lambda = (f_F + f_M)/(2 - f_F - f_M).
\end{array}\right\} \tag{1}
$$

Only one root is admissible, with absolute value less than unity and the same sign as the product of λ and the correlation between mates (P_F and P_M).
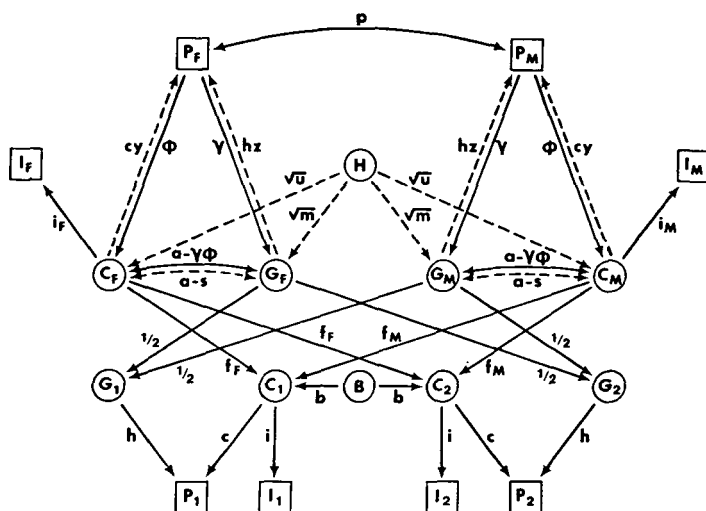

Fig. 2. Treatment of mixed homogamy in nuclear familes. P, G, C, I denote phenotype, genotype, indexed environment and index respectively. Subscripts F and M denote father and mother, 1 and 2 denote two children respectively. B denotes non-transmitted common environment for a sibship. Table 1 defines the parameters of the model: $\gamma = hz + cya$, $\phi = cy + hza$, $s = \sqrt{(mu)}$.

Correlation between the indexed environments of two sibs (C_1 and C_2 in Fig. 2) is easily derived from Fig. 2 as $b^2 + \psi$ where

$$
\psi = f_F^2 + f_M^2 + 2f_F f_M (u + p\phi^2) \tag{2}
$$

measures the correlation between sibling environments derived from parental environments. Correlation between C_1 and C_2 is $b^2 x^2 + \psi$ if the sibs are adults.

4. NUCLEAR FAMILIES

Phenotypes and indices of two parents and one child generate $\dbinom{6}{2} = 15$ correlations. Phenotypes and indices of children generate 3 more correlations: between phenotypes of children, between indices of children, and between child's

phenotype and another child's index. Therefore, the complete model in 14 parameters leaves 4 degrees of freedom in nuclear families for goodness-of-fit test. Additional $d.f.$ are available if certain parameters (like m) are eliminated under a null hypothesis. Also, in the absence of adult sibs the parameter x is eliminated.

Consider the correlation between mother's phenotype P_M and a child's phenotype P_1, say. The correlation between P_M and G_1 is, from Fig. 2,

$$\rho_{P_M G_1} = \tfrac{1}{2}[\gamma(1+p) + hzm + cys] \tag{3}$$

and similarly, the correlation between P_M and C_1 is given by

$$\rho_{P_M C_1} = \phi(f_M + pf_F) + f_F(hzs + cyu). \tag{4}$$

Using (3) and (4) we easily get the mother-child correlation as

$$\rho_{P_M P_1} = h\rho_{P_M G_1} + c\rho_{P_M C_1} \tag{5}$$

Table 2. *Expected correlations in nuclear families under mixed homogamy* (see Fig. 2)

Relation	Variables	Expected correlation (ρ)
FMT (marital)	P_F, I_F	ϕi_F
	P_F, P_M	$p + (hz)^2 m + (cy)^2 u + 2hzcys$
	P_F, I_M	$(p\phi + hzs + cyu)\, i_M$
	I_F, P_M	$(p\phi + hzs + cyu)\, i_F$
	I_F, I_M	$(p\phi^2 + u)\, i_F i_M$
	P_M, I_M	ϕi_M
OPT (parent-offspring)	P_F, P_C	$\tfrac{1}{2}h[\gamma(1+p) + hzm + cys] + c\phi(f_F + pf_M)$ $+ cf_M(hzs + cyu)$
	P_F, I_C	$[\phi(f_F + pf_M) + f_M(hzs + cyu)]i$
	I_F, P_C	$[\tfrac{1}{2}h(a + s + \gamma p\phi) + cf_F + cf_M(u + \phi^2 p)]i_F$
	I_F, I_C	$[f_F + f_M(u + p\phi^2)]ii_F$
	P_M, P_C	$\tfrac{1}{2}h[\gamma(1+p) + hzm + cys] + c\phi(f_M + pf_F)$ $+ cf_F(hzs + cyu)$
	P_M, I_C	$[\phi(f_M + pf_F) + f_F(hzs + cyu)]i$
	I_M, P_C	$[\tfrac{1}{2}h(a + s + \gamma p\phi) + cf_M + cf_F(u + p\phi^2)]i_M$
	I_M, I_C	$[f_M + f_F(u + p\phi^2)]ii_M$
SST (filial)	(P_1, I_1) or (P_2, I_2)	$(c + ha)i$
	(P_1, I_2) or (P_2, I_1)	$[ha + c(b^2 + \psi)]i$
	I_1, I_2	$(b^2 + \psi)i^2$
	P_1, P_2	$\tfrac{1}{2}h^2(1 + m + p\gamma^2) + c^2(b^2 + \psi) + 2hca$

Note: $s = \sqrt{mu}$, $\gamma = hz + cya$, $\phi = cy + hza$, $\psi = f_F^2 + f_M^2 + 2f_F f_M(u + p\phi^2)$.

which is given in Table 2. The technique represented by equations (3) to (5) is useful, especially in deriving correlations for more remote vertical and collateral relatives. Note that in equation (5) h and c would be replaced by hz and cy respectively if P_1 is the phenotype of an adult child. Correlations involving the indices are derived similarly. For example,

$$\rho_{C_M C_1} = f_M + f_F(u + p\phi^2), \tag{6}$$

which gives

$$\rho_{I_M I_1} = ii_M \rho_{C_M C_1}. \tag{7}$$

Also,

$$\rho_{C_M G_1} = (\gamma p\phi + s + a)/2 \tag{8}$$

and hence the correlation between P_1 and I_M is given by

$$\rho_{P_1 I_M} = (h\rho_{C_M G_1} + c\,\rho_{C_M C_1})i_M \tag{9}$$

and finally,

$$\rho_{P_M I_1} = i\rho_{P_M C_1}, \tag{10}$$

which may be expanded using (4). In all, 18 such correlations for nuclear families are presented in Table 2.

5. VERTICAL RELATIVES

Consider the genotypes and indexed environments of two spouses under social homogamy as shown in Fig. 3(a), where the common cause H of Fig 1 is replaced by direct correlational paths. In Fig. 3(a), m denotes the total correlation between the genotypes of spouses, u denotes the total correlation between their indexed environments and so on. Consider the standardized partial regression coefficients m^*, u^*, s^* and t^* in the regression of each variable of the mother (G_M, C_M) separately on each of the father's variables (G_F, C_F), represented in Fig. 3(b) by unidirectional paths. The partial regression coefficients m^*, u^*, s^* and t^* are easily expressed in terms of the ambiguous correlations m, u, s and a as follows. The total correlation between G_F and G_M is m from Fig. 3(a), and $m^* + at^*$ from Fig. 3(b). Deriving all such correlations and equating them yields the following equations

$$\left.\begin{aligned}
m &= m^* + at^*, \\
u &= u^* + as^*, \\
s &= s^* + au^* = t^* + am^*.
\end{aligned}\right\} \tag{11}$$

Solution of these equations gives

$$\left.\begin{aligned}
m^* &= (m - as)/(1 - a^2), \\
u^* &= (u - as)/(1 - a^2), \\
s^* &= (s - au)/(1 - a^2), \\
t^* &= (s - am)/(1 - a^2).
\end{aligned}\right\} \tag{12}$$

Notice that in the absence of genotype-environment correlation ($a = 0$), all the partial regression coefficients will be equal to the corresponding correlations ($m^* = m$, $u^* = u$ and $s^* = t^* = s = \sqrt{(mu)}$). In the absence of genotypic correlation ($m = 0$, and hence $s = \sqrt{(mu)} = 0$), $m^* = t^* = 0$, $u^* > 0$, but $s^* < 0$. In general, when m and u are both positive, m^* and s^* will have the same sign, and u^*, t^* will both have the opposite sign. Finally, it may be mentioned that the correlation between G_M and C_M is $a^* + m^*s^* + t^*u^* + a(m^*u^* + s^*t^*) \equiv a$ by definition. Therefore,

$$a^* = a - s(m + u - 2as).$$

Conversion of two-headed correlations (m, u, s) into unidirectional paths (m^*, u^*, s^*, t^*) avoids much confusion and uncertainty in tracing correlations beyond nuclear families. It should be remembered, however, that both methods are equivalent ways of expressing marital correlations. More details of this approach are given elsewhere (Cloninger, 1979a). In the ancestral generation we shall always

use marital correlations m, u and s, and in subsequent generations we shall use the marital partial regressions m^*, u^*, s^* and t^* appropriately. Similarly, in dealing with phenotypic homogamy it should be noted that the partial regression p^* of either parent's phenotype on that of the other is the same as p. This justifies drawing this path either as a two-headed arrow or as a unidirectional path as Cloninger $et\ al$ (1979) did. This is not causation, but merely represents functional dependence due to selection as discussed in detail elsewhere (Cloninger, 1979a, b). We shall denote this path by a two-headed arrow in the ancestral generation, and by a unidirectional path in subsequent generations.
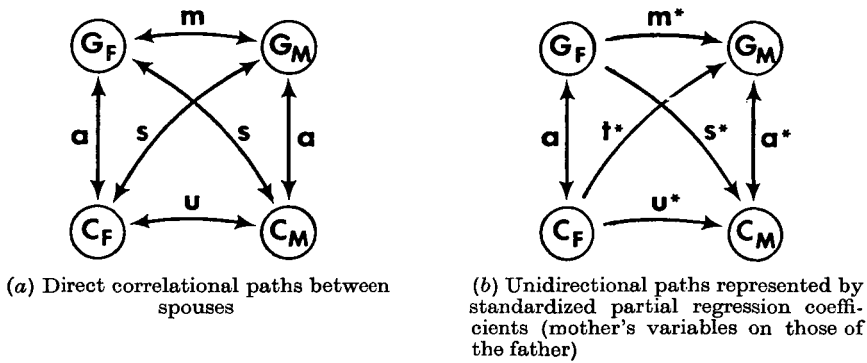

(a) Direct correlational paths between spouses

(b) Unidirectional paths represented by standardized partial regression coefficients (mother's variables on those of the father)

Fig. 3. Representing ambiguous correlations m, u and s (a) in terms of standardized partial regression coefficients m^*, u^*, s^* and t^* (b). G and C denote genotype and indexed environment, subscripts F, M denote father and mother respectively.

Consider a simple three-generation family shown in Fig. 4, consisting of paternal grandparents (generation 3), parents (generation 2) and a child (generation 1). Let K_2 denote the total value of all backward paths between G_{2F} and P_1, and Λ_2 denote the same between C_{2F} and P_1. It is clear that equation (3) also gives $\rho_{P_{3M}G_{2F}}$ in Fig. 3, and further, equation (4) yields $\rho_{P_{3M}C_{2F}}$. Therefore, the correlation between the phenotypes of paternal grandmother and a grandchild is simply

$$\rho_{P_{3M}P_1} = K_2\,\rho_{P_{3M}G_{2F}} + \Lambda_2\,\rho_{P_{3M}C_{2F}}. \tag{13}$$

Note that K_2 and Λ_2 are special kinds of compound paths (not involving ambiguous correlations, and never changing direction) and therefore equation (13) does not correspond to multiplication of correlations which is not permissible in path analysis. It is easy to derive the following quantities from Fig. 4:

$$\left.\begin{aligned} K_2 &= \tfrac{1}{2}h(1+m^*+\gamma phz) + cf_M(s^*+\phi phz) \\ \Lambda_2 &= c[f_F + f_M(u^*+\phi pcy)] + \tfrac{1}{2}h(t^*+\gamma pcy) \end{aligned}\right\} \tag{14}$$

and

and therefore the paternal grandmother-grandchild correlation is obtained by inserting equations (14), (3) and (4) into (13):

$$\begin{aligned} \rho_{P_{3M}P_1} = \ &\tfrac{1}{2}[\gamma(1+p) + hzm + cys][\tfrac{1}{2}h(1+m^*+\gamma phz) + cf_M(s^{*\prime}+\phi phz)] \\ &+ [\phi(f_M + pf_F) + f_F(hzs + cyu)][cf_F + cf_M(u^*+\phi pcy) + \tfrac{1}{2}h(t^*+\gamma pcy)] \end{aligned} \tag{15}$$

Sex-labelling of every parent is important in the presence of maternal effects. Bearing this in mind, the correlation between vertical relatives separated by n generations may be written down as

$$\rho_{Vn} = \tfrac{1}{2}[\gamma(1+p)+hzm+cys]K_n + [\phi(f_i+pf_j)+f_j(hzs+cyu)]\Lambda_n, \qquad (16)$$

where $i \neq j = M$ or F, $n = 1$ for parent-offspring, $n = 2$ for grandparent-grandchild, etc., and, $K_1 = h$ or hz, $\Lambda_1 = c$ or cy depending on whether the descendent is a child or an adult, and

$$\left.\begin{aligned} K_n &= \tfrac{1}{2}(1+m^*+\gamma phz)K_{n-1}+f_l(s^*+\phi phz)\Lambda_{n-1}, \\ \Lambda_n &= [f_k+f_l(u^*+\phi pcy)]\Lambda_{n-1}+\tfrac{1}{2}(t^*+\gamma pcy)K_{n-1}, \end{aligned}\right\} \qquad (17)$$

where $k \neq l = F$ or M depending on the sexes of the intermediate parents.
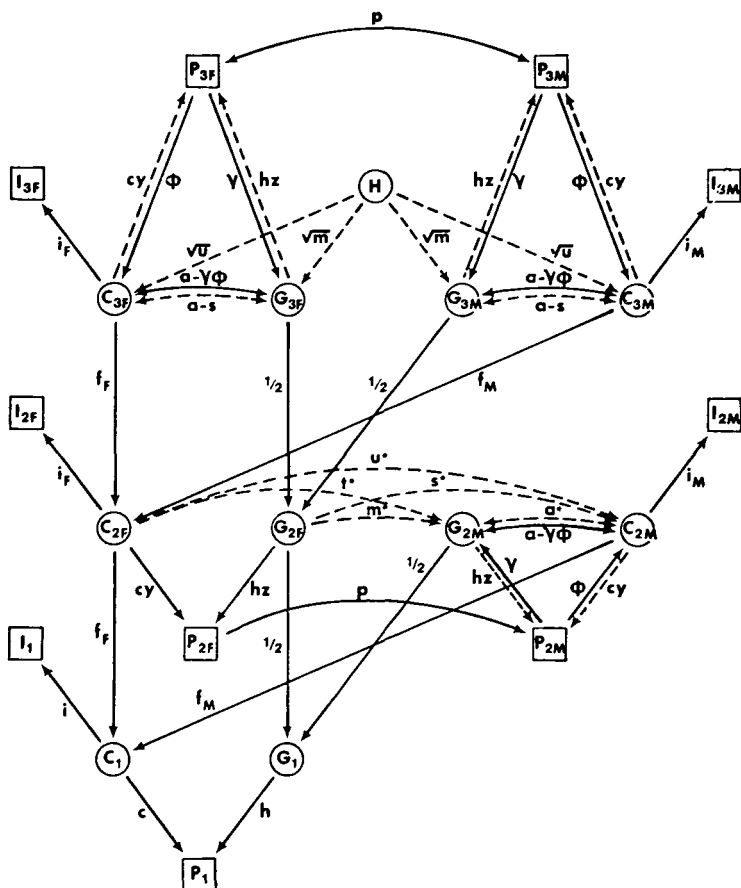

Fig. 4. Path diagram for grandchild (subscript 1) and paternal grandparents (subscript 3) under mixed homogamy. Paths irrelevant to grandparent-grandchild correlation are omitted.

6. COLLATERAL RELATIVES

For collateral relatives not involving multiple spouses, we shall denote the marital relationships in terms of m^*, u^*, s^*, t^* even in the ancestral generation. Phenotypic correlations are also shown as unidirectional paths. Fig. 5 displays the path diagram for two full sibs (brother and sister, subscripted as *1F* and *2M* respectively), their spouses and one child for each couple. Apart from other more remote relatives such as like in-laws and unlike in-laws (Rao *et al.* 1976), correlations for uncle–niece and first cousins can be derived from Fig 5, taking care to label the sex properly.

The direct correlational paths between the genotypes and environments of the sibs (G_{1F}, G_{2M}, C_{1F}, C_{2M}) are derived from Fig. 2, where b is replaced by bx since these are adult sibs. The phenotypic correlation between paternal aunt and niece (nephew) is easily seen to be

$$\rho_{P_{2M}P_{12}} = [\tfrac{1}{2}hz(1+m+p\gamma^2)+cya]K_2 + [cy(b^2x^2+\psi)+hza]\Lambda_2, \qquad (18)$$

where K_2 and Λ_2 are given by equation (14), and ψ is given by equation (2). Similarly, equation (18) also yields maternal uncle–niece (nephew) correlation with
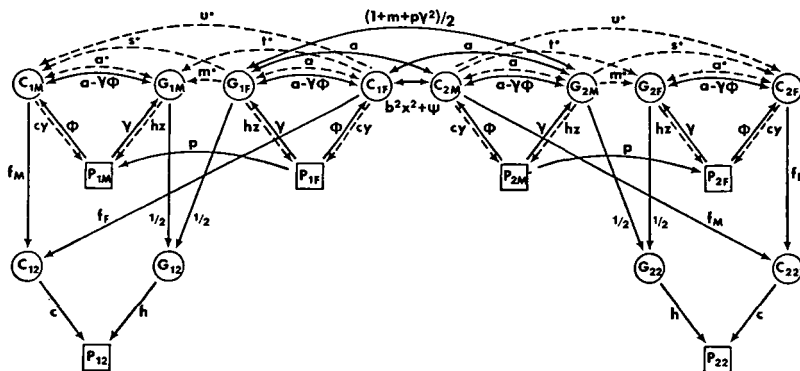

Fig. 5. Path diagram for cross first cousins under mixed homogamy. $\psi = f_M^2 + f_F^2 + 2f_F f_M(u+\phi^2 p)$ denotes the correlation between indexed environments of sibs derived from those of the parents. Subscripts 1 and 2 denote the two sibs in the original generation. In subsequent generations, the first digit of the subscript identifies the original generation, and the second digit for the generation number within the vertical line.

K_2 and Λ_2 replaced by K_2' and Λ_2' where sexes are reversed. By a similar argument we get the phenotypic correlation for cross first cousins as

$$\rho_{P_{12}P_{22}} = \tfrac{1}{2}(1+m+p\gamma^2)K_2K_2' + (b^2x^2+\psi)\Lambda_2\Lambda_2' + a(K_2\Lambda_2'+\Lambda_2K_2'), \qquad (19)$$

where K_2 and Λ_2 are as given in (14), and sexes are reversed in K_2' and Λ_2'.

By a slight generalization, one could easily obtain the phenotypic correlation for an individual (say P_{2M}) and nth descendent of her sib. This is actually given by equation (18) where K_2 and Λ_2 are replaced by K_n and Λ_n respectively. Similarly, the phenotypic correlation between mth and nth descendents of an ancestral pair

is given by a slight rearrangement of equation (19):

$$\rho_{mn} = \tfrac{1}{2}(1 + m + p\gamma^2)K_m K'_n + (b^2 x^2 + \psi)\Lambda_m \Lambda'_n + a(K_m \Lambda'_n + K'_n \Lambda_m), \qquad (20)$$

where sexes are properly labelled for all intermediate parents.

7. OTHER RELATIVES

By a variation of Figs. 2 and 5 it is possible to derive the expected correlations for a variety of other useful relationships such as twins, sibs reared apart, foster sibs, half-sibs, children of MZ twins and so on. Some of these are presented in Table 3.

For example, for foster sibs reared together, the correlation is easily derived from Fig. 2 by eliminating the four genetic paths from parental genotypes to those

Table 3. *Expected correlations for other types of relatives*

Relation	Expected correlation
MZ twins reared together by parents	$h^2 + c^2(b^2 + \psi) + 2hca$
MZ twins reared apart, one by true parents	$(h^2 + hac)\theta$
Full sibs reared apart, one by true parents	$[\tfrac{1}{2}h^2(1 + m + p\gamma^2) + hac]\theta$
Full sibs, reared apart by foster parents	$[h^2(1 + m + p\gamma^2)/2]\theta^2$
Foster sibs reared together	$c^2\theta^2(b^2 + \psi)$
Foster sibs reared together by parents of one	$[c^2(b^2 + \psi) + hac]\theta$
Offspring-parent living apart	$\tfrac{1}{2}h\theta[\gamma(1 + p) + hzm + cys]$
Offspring-foster father*	$c\theta[\phi(f_F + pf_M) + f_M(hzs + cyu)]$
Paternal half-sibs reared together by parents of one†	$[\tfrac{1}{4}h^2(1 + 3m + 2\gamma^2 p + \gamma^2 p^2) + c^2(b^2 + \psi)$ $+ 2hca + \tfrac{1}{2}hcf_M(a - s - \nu p^2\phi)]/$ $\sqrt{[1 - hcf_M(a - s - \nu p^2\phi)]}$
Paternal half-sibs reared separately by own parents†	$\tfrac{1}{4}h^2(1 + 3m + 2\gamma^2 p + \gamma^2 p^2)$ $+ c^2[b^2 + \psi - f_M^2(1 - u - \phi^2 p^2)]$ $+ 2hca - hcf_M(a - s - \nu p^2\phi)$
Children of male MZ twins‡	$K_2^2 + 2aK_2\Lambda_2 + (b^2 x^2 + \psi)\Lambda_2^2$

* Reversing subscripts F and M gives offspring-foster mother.
† Reversing subscripts F and M gives maternal half-sibs.
‡ K_2 and Λ_2 are given by equation 12; reversing the roles of sex in K_2 and Λ_2 gives the correlation for children of female MZ twins.
Note: $\psi = f_F^2 + f_M^2 + 2f_F f_M(u + p\phi^2)$.

of the children (foster) and remembering to multiply by the ratio of standard deviations $\theta = 1/\sqrt{(1 - 2hca)}$ (Rao, Morton & Yee, 1974): thus we get the foster-sib correlation as $c^2\theta^2(b^2 + \psi)$ where ψ is given by (2). Random placement of the foster children is assumed.

For half-sibs we assume, as before (Rao *et al.* 1976) that multiple spouses are as similar to each other with respect to social homogamy as either one is to their common spouse. This is facilitated by showing H (social homogamy) as a direct cause of the genotypes and indexed environments of the three spouses, with a

path $\sqrt{m}$ to each genotype and $\sqrt{u}$ to each indexed environment. Two cases are of special interest: half-sibs reared separately by own parents, and half-sibs reared together by parents of one. These two cases are easily distinguished by substituting the contribution of one uncommon parent to one half-sib's indexed environment by that of the other uncommon parent.

For the MZ twin design of Nance, Corey & Boughman (1978) we can derive some interesting correlations from Fig. 5 by writing only one genotype for the twin pair (say, G instead of G_{IF} and G_{2M} in Fig. 5). For example, labelling sexes properly, the correlation between an MZ twin and co-twin's child is

$$\rho = \gamma K_2 + \Lambda_2 [hza + cy(b^2 x^2 + \psi)] \tag{21}$$

and the correlation between children of MZ twins is

$$\rho = K_2^2 + 2aK_2\Lambda_2 + \Lambda_2^2(b^2 x^2 + \psi). \tag{22}$$

Equations (21) and (22) hold for male MZ twins. Substitution of K_2 and Λ_2 by K_2' and Λ_2' yields the same for female MZ twins.

8. STATISTICAL ANALYSIS

Given a set of sample correlation coefficients (r_i, $i = 1, 2, \ldots, m$) and their sample sizes (n_i, $i = 1, 2, \ldots, m$), we propose to take the joint log-likelihood as

$$\ln L = -\chi^2/2 + \text{constant},$$

$$\chi^2 = \sum_{i=1}^{m} [f(r_i) - f(\rho_i)]^2/\sigma_i^2, \tag{23}$$

where $f(x)$ is a suitable function of x, and σ^2 is the variance of $f(x)$. The two appropriate forms are

$$f(x) = \begin{cases} x \text{ with } \sigma^2 = (1-\rho^2)^2/n, \\ z = \frac{1}{2}\ln\dfrac{1+x}{1-x} \text{ with } \sigma^2 = 1/n^*, \end{cases}$$

where $n^* = n - 3$ for interclass correlations, and $n^* = n - 1\cdot5$ for intraclass correlations. Whenever correlations are based on pairs of observations we use the z transforms, and $f(x) = x$ otherwise.

We have suggested earlier that the z-transforms be bias-corrected prior to data analysis. Such refinements have negligible effect on the results (Goldberger, 1978), and will be neglected hereafter. We have also proposed quadratic forms involving correlations between correlations whenever correlation coefficients are estimated from the same sample (Rao et al. 1979). However, methods with and without correlations between correlations gave substantially similar results for a number of traits (Rao et al. 1979; Morton et al. 1978; Gulbrandsen et al. 1979; Krieger et al. 1978). This justifies equation (23) which ignores correlations between correlations.

9. DISCUSSION

Assortative mating as treated here enables for the first time tests of hypotheses on the nature of assortment. The two null hypotheses of interest are $m = u = 0$ which corresponds to phenotypic assortative mating, and $p = 0$ which represents social homogamy. Experience so far has been limited to these two extremes. We have been using social homogamy with primary interest in physiological traits such as lipoprotein concentrations (Rao *et al.* 1979), blood pressure (Morton *et al.* 1978; Krieger *et al.* 1979) and uric acid (Gulbrandsen *et al.* 1979). Others have been using phenotypic assortative mating with primary interest in behavioural traits such as IQ (Wright, 1978; Cloninger *et al.* 1979; Jencks, 1972). Analyses of IQ data under these two extremes gave different results, but it could not be asserted whether this was due to differences in the modeling of assortative mating, cultural inheritance, or intergenerational differences. The model of mixed homogamy presented here is capable of resolving such uncertainties.

Treatment of indices may be generalized by adding a path from genotype to index. This path coefficient measures the product of two quantities: square-root of genetic heritability for the index, and correlation between the genotypes of phenotype and index. If the path coefficient is found to be significant by a likelihood ratio test, data analyses should incorporate it.

A FORTRAN computer program incorporating these methods is being developed, which will be used to analyse several data sets including IQ. These results will be presented in a subsequent paper.

REFERENCES

CLONINGER, C. R., RICE, J. & REICH, T. (1979). Multifactorial inheritance with cultural transmission and assortative mating. II. A general model of combined polygenic and cultural inheritance. *American Journal of Human Genetics*. (In the Press.)

CLONINGER, C. R. (1979*a*). Resolution of polygenic and cultural inheritance from reciprocal phenotype-environmental interaction. (In preparation.)

CLONINGER, C. R. (1979*b*). Absolute and relative structural dependence: The interpretation of causal relations by path analysis. (In preparation.)

GOLDBERGER, A. S. (1978). Pitfalls in the resolution of IQ inheritance. In *Genetic Epidemiology* (ed. N. E. Morton and C. S. Chung), pp. 195–222. New York: Academic Press.

GULBRANDSEN, C. L., MORTON, N. E., RAO, D. C., RHOADS, G. G. & KAGAN, A. (1979). Determinants of plasma uric acid. *Human Genetics*. (In the Press.)

JENCKS, C. (1972). *Inequality*. New York: Basic Books.

KRIEGER, H., MORTON, N. E., RAO, D. C. & AZEVEDO, E. (1979). Race, height and blood pressure in Northeastern Brazil. (In preparation.)

MORTON, N. E. & RAO, D. C. (1978). Quantitative inheritance in man. *Yearbook of Physical Anthropology*. **21**, 12–41.

MORTON, N. E., GULBRANDSEN, C. L., RAO, D. C., RHOADS, G. G. & KAGAN, A. (1978). Determinants of blood pressure in Japanese-American families. *Human Genetics* (pending).

Morton, N. E. & Rao, D. C. (1979). Causal analysis of family resemblance. In the proceedings on *The Genetic Analysis of Common Diseases: Applications to Predictive Factors in Coronary Heart Disease*. (ed. C. F. Sing and M. Skolnick). New York: Alan R. Liss, Inc. (In the Press.)

Nance, W. E., Corey, L. A. & Boughman, J. A. (1978). Monozygotic twin kinships: a new design for genetic and epidemiologic research. In *Genetic Epidemiology* (ed. N. E. Morton and C. S. Chung). pp. 87–132. New York: Academic Press.

Rao, D. C., Morton, N. E. & Yee, S. (1974). Analysis of family resemblance. II. A linear model for familial correlation. *American Journal of Human Genetics* **26**, 331–359.

Rao, D. C., Morton, N. E. & Yee, S. (1976). Resolution of cultural and biological inheritance by path analysis. *American Journal of Human Genetics* **28**, 228–242.

Rao, D. C. & Morton, N. E. (1978). IQ as a paradigm in genetic epidemiology. In *Genetic Epidemiology* (ed. N. E. Morton and C. S. Chung), pp. 145–194. New York: Academic Press.

Rao, D. C., Morton, N. E., Gulbrandsen, C. L., Rhoads, G. G., Kagan, A. & Yee, S. (1979). Cultural and biological determinants of lipoprotein concentrations. *Annals of Human Genetics*. **42**, 467–477.

Wright, S. (1968). *Evolution and the Genetics of Populations*, Volume 1. *Genetic and Biometric Foundations*. The University of Chicago Press.

Wright, S. (1978). *Evolution and the Genetics of Populations*, Volume 4. *Variability Within and Among Natural Populations*. The University of Chicago Press.