Vol. 24, Fall 1977, pp. 225–233

Socioeconomic Assortative Mating in Santiago, Chile: A Demonstration Using Stochastic Matrices of Mother-child Relationships Applied to ABO Blood Groups

Carlos Y. Valenzuela and Zuraiya Harb

Departamento de Biología Celular y Genética
Facultad de Medicina
Universidad de Chile
Santiago, Chile

ABSTRACT: The gene and phenotype frequencies for the ABO blood groups system were studied in two socioeconomically different subpopulations of Santiago, Chile. The data were taken from the maternity services and blood banks of two hospitals, one serving mainly low socioeconomic classes and the other serving middle and high socioeconomic groups. Results show a clear difference in gene and phenotype frequencies between the two subpopulations, a difference maintained in two generations (mothers and their children). These results reinforce other studies showing that socioeconomic forces have resulted in assortative mating and different genetic subpopulations in Chile.

The population of Santiago can be segmented into a variety of socioeconomic and cultural classes as can most, if not all, urban populations. In Chile there are studies suggesting that these different urban socioeconomic classes have different genetic compositions (see, for example, the ABO blood group studies of Pinto-Cisternas et al., 1971; Rothhammer, 1973; and Palomino, 1976). If this is also true of Santiago, we would expect to find different gene frequencies in the different socioeconomic groups. Since human matings probably occur in an assortative way with respect to socioeconomic variables, the evolution of these populations will be conditioned by their socioeconomic structures.

The Chilean population stems mainly from the admixture of Spanish people with Chilean aborigines. The ABO blood group frequencies in these two peoples were very different. While the aborigines were presumably all of type O, the Spanish settlers had types A and B in appreciable frequencies. Thus, the occurrence of types A and B may be correlated with the degree of Spanish or European admixture. The papers cited above suggest that the upper socioeconomic classes have higher frequencies of types A and B than do lower socioeconomic classes.

Our aims in this paper are threefold. First, we will show that there are significant differences between two very different socioeconomic subpopulations of Santiago in their ABO blood group phenotypes and in the frequencies of the alleles O, A, and B. Second, we will show that the socioeconomic structure conditions the mating structure in the population of Santiago in such a way that the gene frequencies of mothers and their children are almost the same within each socioeconomically defined subpopulation, that is to say, there is a strong tendency toward as-

sortative matings. Finally, we will show that one needs to be very cautious when interpreting gene frequencies obtained after assuming Hardy-Weinberg equilibrium.

Chilean health service benefits reach Chileans either through the National Health Service or through private medical practice. In general, the latter benefits the middle and upper socioeconomic classes; the former, the lower ones. Among the benefits available to all groups are maternity services and blood banks. Maternity services offer an opportunity to study the gene frequencies of mothers, who represent one generation, and their infants, who represent another. Blood banks permit one to study a sample of donors which presumably represents the general population from which the mothers are drawn. The relation between the gene frequencies of the mothers and the gene frequencies of their husbands can be studied through the phenotypes of children born to O blood group mothers. Helpful in this context is the method of stochastic matrices of conditional probabilities developed by Li and Sacks (1954), Li (1958, 1969) and applied recently by Harb and Valenzuela (1976).

MATERIALS AND METHODS

NOTATION AND FORMULAS

A, B, and O represent the alleles of the ABO system; p, q, and r, their respective population frequencies. A, B, O, and AB are the phenotypes of this blood group system. T is a particular matrix that describes the conditional probabilities between individuals related as mother-child. The elements of this matrix are given in Table 1. It should be noted that in this matrix the Hardy-Weinberg equilibrium has been assumed, but it is possible to construct a matrix which does not make this assumption by adding the appropriate coefficients related to conditions which lead to an equilibrium different from the Hardy-Weinberg.

Children born to mothers belonging to the O blood group represent a direct sample of the fathers' gene pool, as is demonstrated in Table 2. Again, if mating is at random, f_A, f_B, and f_O calculated by this method will be equal to p, q, and r of the general fathers' population. If panmixia does not obtain, then this statement does not hold; but the frequencies f_A, f_B and f_O will then provide insight into the nature of the deviation from simple randomness.

Children born to O-blood mothers represent a population backcross. The phenotypic frequencies obtained from these children are directly the gene frequencies of their fathers. It is thus possible to obtain the variance-covariance matrix of p, q, and r for the fathers, as a simple variance-covariance matrix of a trinomial distribution (Lebart and Fenelon, 1973). However, since $p + q + r = 1$, it is sufficient to work with two frequencies. We

TABLE 1

T MATRIX OF MOTHER-CHILD RELATIONSHIP FOR ABO BLOOD GROUP

Mother Genotype	Child Genotype					
	AA	*AO*	*BO*	*BB*	*AB*	*OO*
AA	p	r	0	0	q	0
AO	$\frac{1}{2}p$	$\frac{1}{2}(p + r)$	$\frac{1}{2}q$	0	$\frac{1}{2}q$	$\frac{1}{2}r$
BO	0	$\frac{1}{2}p$	$\frac{1}{2}(q + r)$	$\frac{1}{2}q$	$\frac{1}{2}p$	$\frac{1}{2}r$
BB	0	0	r	q	p	0
AB	$\frac{1}{2}p$	$\frac{1}{2}r$	$\frac{1}{2}r$	$\frac{1}{2}q$	$\frac{1}{2}(p + q)$	0
OO	0	p	q	0	0	r

TABLE 2

SEGREGATION FREQUENCIES OF ABO GENOTYPES IN O-BLOOD MOTHERS

FATHER GENOTYPE	FREQUENCY*	CHILD GENOTYPE		
		AO	*BO*	*OO*
AA	D_{AA}	D_{AA}	...	...
AO	H_{AO}	$\frac{1}{2}H_{AO}$	...	$\frac{1}{2}H_{AO}$
BO	H_{BO}	...	$\frac{1}{2}H_{BO}$	$\frac{1}{2}H_{BO}$
BB	D_{BB}	...	D_{BB}	...
AB	H_{AB}	$\frac{1}{2}H_{AB}$	$\frac{1}{2}H_{AB}$	...
OO	D_{OO}	...	...	D_{OO}
Total	1	$D_{AA} + \frac{1}{2}H_{AO} + \frac{1}{2}H_{AB} = f_A$	$D_{BB} + \frac{1}{2}H_{BO} + \frac{1}{2}H_{AB} = f_B$	$D_{OO} + \frac{1}{2}H_{BO} + \frac{1}{2}H_{AO} = f_O$

* D = frequency of homozygotes; H = frequency of heterozygotes; f_A, f_B, f_O = respective frequencies p, q, and r in the fathers.

have chosen p and q to obtain the estimators needed in the next stage of our analysis. The variance-covariance matrix under discussion is shown in Table 3. The sum of the elements of this matrix is the variance of r. The gene frequencies obtained in this manner can readily be shown to be a sufficient statistic and therefore maximum likelihood estimators (Mood and Graybill 1963).

TABLE 3

MATRIX OF VARIANCE-COVARIANCE FOR A TRINOMIAL DISTRIBUTION

$p(1-p)$	$-pq$
$-pq$	$q(1-q)$

Gene frequencies from data belonging to mothers and their children and from the blood banks were calculated by the maximum likelihood method as shown in Stevens (1938). This method also gives a good approximation to the variance-covariance of the gene frequencies estimators. We used Bernsteins' coefficient statistics (Cavalli-Sforza and Bodmer, 1971) as trial values to initiate the iterative procedure which culminates in the maximum likelihood estimates. We have also calculated gene frequencies from the matrix of mother-child phenotypes. This matrix is given in Table 4 and can be used to generate a polynomial distribution from which we have obtained the maximum likelihood estimators of p, q, and r (Li, 1958). This method enables one to use all of the information present in the mother-child combinations.

In Table 4, we again assume a Hardy-Weinberg equilibrium. It is also possible to construct the variance-covariance matrix of the gene frequencies directly from the maximum likelihood estimates as has been previously shown in Table 3. The actual number of independent genes can be obtained by dividing the elements of the variance-covariance matrix of the gene frequencies by the corresponding ones in the variance-covariance matrix of the gene frequencies estimators.

TABLE 4

MOTHER-CHILD PHENOTYPIC MATRIX

MOTHER	CHILD			
	AB	*A*	*B*	*O*
AB	$pq(p+q)$	$pq(p+r)$	$pq(q+r)$	...
A	$pq(p+r)$	$p(p^2+3pr+r^2)$	pqr	pr^2
B	$pq(q+r)$	pqr	$q(q^2+3qr+r^2)$	qr^2
O	...	pr^2	qr^2	r^3

Theoretical numbers can be obtained as has been shown by Li (1958); but since genetic dominance introduces complications into these calculations, we have chosen to work with the actual numbers. The latter are slightly different if one chooses the variance of p, or the covariance of p and q, or the variance of q; these numbers also differ slightly from the limits within which the theoretical numbers must lie. We have used the smallest actual number in all cases to avoid artificially increasing the significances of the statistics. This procedure introduces an element of approximation in the computation of the Hotelling T^2 and F; but since the numbers which are used are large, we believe the true values of T^2 and F to be little different from the calculated ones.

Now, given p and q, and their respective variances and covariances from two samples, it is possible to calculate the statistical significance of a bivariate Hotelling T^2 (Morrison 1967). The computation proceeds as follows:

$$T^2 = \frac{N_1 N_2}{N_1 + N_2} (p_1 - p_2, q_1 - q_2) S_t^{-1} \begin{bmatrix} p_1 - p_2 \\ q_1 - q_2 \end{bmatrix}$$

where the subscripts 1 and 2 indicate samples 1 and 2, N is the sample size or the actual number and S_t is the pooled variance-covariance matrix which corresponds to

$$S_t = (N_1 + N_2 - 2)^{-1} [(N_1 - 1)S_1 + (N_2 - 1)S_2].$$

The significance of the T^2 is obtained by an F test, where

$$F = \frac{N_1 + N_2 - d - 1}{(N_1 + N_2 - 2)d} T^2$$

and d is one of the two degrees of freedom (in this case, 2). The other is $N_1 + N_2 - d - 1$.[1]

The Hardy-Weinberg equilibrium for phenotypes was checked by squaring the gene frequencies $(p + q + r)$ obtained and then applying the χ^2 test for phenotypes as found in Cavalli-Sforza (1971).

SAMPLES

Our samples of observations were drawn from the Hospital San José (HSJ), a National Health Service Hospital in the northern part of Santiago, and from the Clínica Alemana (CAL), a private clinic located in the eastern part of Santiago whose patients are predominantly from the middle and upper socioeconomic classes. In these hospitals we obtained data from the maternity services on mothers and their children and from donors to the blood bank at HSJ and receptors at CAL. We choose blood bank receptors at CAL rather than donors on the assumption that the donors may not belong to the same socioeconomic levels as the mothers. Blood typing of infants was done at birth from cord blood.

RESULTS

The phenotypic distributions of the mothers, children, donors (HSJ), and receptors (CAL) are shown in Tables 5 and 6. The significance of the difference in the phenotypic distribution between the hospitals for these various classes of individuals has been tested by chi-square. Significance is assumed if the observed deviation exceeds that predicted by chance one in twenty times, which corresponds to a chi-squre of 7.81 for three degrees of freedom and 5.99 for two degrees of free-

[1] It is also possible to treat the product $(p_1 - p_2, q_1 - q_2) S^{-1} \begin{bmatrix} p_1 - p_2 \\ q_1 - q_2 \end{bmatrix}$ as a chi-square value and then to obtain significance level with one degree of freedom. This can be done because N_1 and N_2 are large numbers.

TABLE 5

PHENOTYPIC DISTRIBUTION OF *T* MATRIX AND BLOOD DONORS IN HSJ FOR ABO BLOOD GROUP

	CHILD				
MOTHER	*O*	*A*	*B*	*AB*	Total
O	458	93	38	0	589
A	100	164	11	4	279
B	34	11	31	7	83
AB	0	7	8	2	17
Total	592	275	88	13	968
Donors	1,511	696	215	50	2,472

TABLE 6

PHENOTYPIC DISTRIBUTION OF *T* MATRIX AND BLOOD RECEPTORS IN CAL FOR ABO BLOOD GROUP

	CHILD				
MOTHER	*O*	*A*	*B*	*AB*	Total
O	850	280	80	0	1,210
A	266	547	23	20	856
B	71	27	113	24	235
AB	0	30	24	6	60
Total	1,187	884	240	50	2,361
Receptors	472	347	94	20	933

dom. Table 7 gives the results of these tests. The frequencies and the method used to estimate them are shown in Table 8. We tested the Hardy-Weinberg equilibrium with p, q, and r calculated from one source of data and phenotypic numbers from the same or another source (Table 9). A comparison between the HSJ and CAL samples was not made, since the phenotypic distribution revealed these samples to be drawn from different populations (see Table 7). The chi-square test where p, q, and r and the phenotypic frequencies belong to the same sample has one degree of freedom. Two degrees of freedom are lost in estimating p and q from the observed distribution of phenotypes (Cavalli-Sforza, 1971; Li, 1958). If p, q, and r belong to a different sample, then the chi-square test has three degrees of freedom.[2]

These last tests, those in Table 9, must be interpreted with care. There are striking differences when comparing reciprocal situations, that is to say, if one compares the chi-square values calculated applying p, q, and r of sample A with phenotypes of sample B with chi-square values calculated applying gene frequencies of B with phenotypes from A, substantial differences are seen. This happens because the p, q, and r of any one sample are not estimated without error; their variance or confidence intervals need to be considered. Thus, the only clear significance is that found within CAL's children. The significance between matrix-donors from HSJ and matrix-children from CAL needs to be discussed further.

To solve the problem of interpretation mentioned above, it is necessary to include

[2] In this case we are assuming that p, q, and r are the parameters of the population.

TABLE 7

CHI-SQUARE VALUES FOR PHENOTYPES' DISTRIBUTIONS

	HSJ		CAL			
Sample	Children	Donors	Mothers	Children	Receptors	HSJ *T* Matrix
HSJ						
Mothers	0.716	0.38	26.03*	31.22*	20.58*	...
Children	...	1.89	29.91*	34.28*	22.91*	...
Donors	...	...	49.07*	60.13*	32.43*	...
CAL						
Mothers	...	...	...	1.63	0.66	...
Children	...	...	...	...	0.03	...
T matrix	...	...	...	...	...	13.41*†

* $p < 0.05$.
† With 2 df; the remainder with 3 df.

TABLE 8

ABO GENE FREQUENCIES IN THE DIFFERENT SAMPLES

SAMPLE	METHOD*	GENE FREQUENCIES		
		p	q	r
HSJ				
Mothers	M.L.	0.16681	0.05315	0.78004
Children	M.L.	0.16206	0.05368	0.78426
Donors	M.L.	0.16429	0.05508	0.78063
T matrix	T matrix	0.15789	0.06452	0.77759
Ph. matrix	Ph. matrix (M.L.)	0.16579	0.05508	0.77913
CAL				
Mothers	M.L.	0.21788	0.06462	0.71750
Children	M.L.	0.22312	0.06359	0.71329
Receptors	M.L.	0.22163	0.06322	0.71515
T matrix	T matrix	0.23141	0.06116	0.70743
Ph. matrix	Ph. matrix (M.L.)	0.22106	0.06327	0.71567

* M.L. = maximum likelihood; Ph. = phenotypic matrix.

TABLE 9

CHI-SQUARE VALUES FOR HARDY-WEINBERG EQUILIBRIUM TEST

SOURCE OF p, q, AND r	SOURCE OF PHENOTYPES		
	Mothers	Children	Donors (HSJ)-Receptors (CAL)
	Within HSJ		
HSJ			
Mothers	0.00*	1.38	1.31
Children	0.29	1.10*	1.20
Donors	0.21	1.21	0.78*
T matrix	4.79	4.74	8.57†
Ph. matrix	0.18	1.33	0.85
	Within CAL		
CAL			
Mothers	0.86*	6.64	2.14
Children	1.56	5.92*†	1.99
Receptors	1.31	6.05	1.98*
T matrix	5.55	7.84†	3.08
Ph. matrix	1.21	6.11	1.99

* With 1 df; the remainder with 3 df.
† $p < 0.05$.

gene frequencies and their variances when comparing two samples, which was done by the method previously described (T^2 and F test). The results are presented in Table 10. For 2 and more than 120 degrees of freedom, the 0.05 level of significance corresponds to an F value of 3.00. The 0.01 level of significance is reached with F values over 4.61. Table 10 demonstrates clearly the existence of two samples with different gene frequencies and supports the need for caution in evaluating the chi-square values of Table 9.

We have calculated p, q, and r for all individuals either from CAL or HSJ, and with these frequencies we have tested the Hardy-Weinberg equilibrium. The values for p, q, and r were 0.19547, 0.05972, and 0.74481 respectively, and the chi-square value was 3.49 with one degree of freedom ($p > 0.05$).

Finally, we examined the phenotypic matrices to see whether they were distributed according to the Hardy-Weinberg equilibrium. The chi-square values for the CAL matrix and for the HSJ matrix were 10.59 and 9.86 respectively, which are not significant with eleven degrees of freedom.

DISCUSSION

Since different socioeconomic levels in the populations of admission are assured by the different admission requirements of the hospitals, our data support the view that the genetic compositions of these socioeconomically different populations are different. Previous studies have also reached this same conclusion. Here, how-

TABLE 10

MATRIX OF F VALUES FOR PAIRED SAMPLES

	HSJ				CAL				
	Children	Donors	T Matrix	Ph. Matrix	Mothers	Children	Receptors	T Matrix	Ph. Matrix
HSJ									
Mothers	0.14	0.68	0.62	0.04	12.61*	14.42*	9.77*	10.74*	14.58*
Children	...	0.05	0.50	0.08	14.61*	16.60*	11.34*	12.33*	16.90*
Donors	...	...	0.48	0.01	23.83*	27.41*	15.14*	15.76*	29.80*
T matrix	...	...	...	0.46	5.71*	6.60*	5.48*	6.57*	6.39*
Ph. matrix ...	...	...	...	...	16.05*	18.48*	11.59*	12.55*	19.15*
CAL									
Mothers	...	...	...	...	...	0.17	0.06	0.54	0.09
Children	...	...	...	...	...	...	0.01	0.21	0.03
Receptors	...	...	...	...	...	...	...	0.20	0.00
T matrix	...	...	...	...	...	...	...	...	0.33

*$p < 0.01$.

ever, we add evidence that not only do the phenotypic distributions differ but so too do the gene frequencies. We also show that these differences are maintained at least for two consecutive generations.

It is necessary for further discussion to emphasize the different degrees of socioeconomic heterogeneity within each group. While the HSJ sample represents an almost homogeneous population from a socioeconomic viewpoint, the CAL sample involves a range of socioeconomic groups, one extending from the middle to the highest socioeconomic levels. We believe this difference is why there is no significance in the chi-square test for the Hardy-Weinberg equilibrium in the HSJ samples (see Table 9). The significance seen when p, q, and r from T matrix are used to calculate expected values of donors' phenotypes can be explained as a result of not considering the variance of p, q, and r and from the further possibility that the donors are not drawn from the same population as the mothers. Donors need not be beneficiaries of the National Health Services.

The chi-square for the Hardy-Weinberg equilibrium in the CAL sample of children is significant at the 0.02 level. It is clear then that the children are not from a generation in Hardy-Weinberg equilibrium, or alternatively that the mothers are not from the generation of their children. The significance seen when the p, q, and r calculated from the T-matrix on the phenotypes of children is used can also be explained; it is a consequence of the failure to take cognizance of the gene frequency variances. It is also possible, of course, that this latter significance stems from a deviation from panmixia, but the number of observations needed to establish this fact unequivocally would be very large.

The use of T matrices to estimate gene frequencies represents an alternative approach to the methods of Bernstein or the classic maximum likelihood technique. It requires, however, that the recessive homozygote genotypes be in sufficient number, since the variances of estimators will be small. The T matrix approach yields maximum likelihood estimates directly and does not necessarily assume that the population sampled is in panmictic equilibrium. We have shown that children seen at the CAL do not fit the Hardy-Weinberg equilibrium, but their gene frequencies do not differ significantly from those of their mothers. These two statements may be interpreted by assuming that the fathers of these children do not differ from the mothers in their gene frequencies, but that some sort of assortative mating is occurring. This interpretation

would be in agreement with the known heterogeneous socioeconomic composition of the CAL sample. Of course, it is possible also that a small difference in gene frequencies between fathers and mothers may exist, but the method is not sufficient to demonstrate it. Evidence for assortative mating is necessarily indirect because paternal blood typing was not possible. In the HSJ observations where there is more socioeconomic homogeneity, no discrepancy with the Hardy-Weinberg assumptions is found.

The most conspicuous of our results is the sharp difference in phenotypes and in gene frequencies between the CAL and HSJ samples. This finding strongly suggests that there are at least two subpopulations in the population of Santiago. Since these differences are maintained over two generations (mother and children), it is possible to conclude that the socioeconomic factor is a very important one in the origin and maintenance of these subpopulational differences. It remains for further studies to demonstrate if socioeconomic or cultural factors predominate in the maintenance of these subpopulations.

When the data from both samples are mixed, as if they were a larger sample of the population of Santiago, the results are striking. There is no significant deviation from Hardy-Weinberg equilibrium, although we know that the samples belong to at least two very different populations. This result is explicable. First, the equations of estimation from which the gene frequencies are derived assume an equilibrium (except for the T matrix method); and second, the test of equilibrium has a very large Type II statistical error (β error). This last result emphasizes the need to be extremely cautious in interpreting the presence or absence of Hardy-Weinberg equilibrium when the classic methods of gene frequency estimation and analysis are used.

There are two technical considerations to be mentioned. First, it is possible that the ABO blood system cannot be typed so clearly from infant cord blood as it can from the blood samples from the mothers. The other possible source of technical error is the misclassification of the weak subgroups of A and B groups. They can be taken as O group (Race and Sanger 1968). If this is happening in our samples, as we suppose, then the actual significance of our results might be larger than those found.

Finally, we want to point out that maternity data are easily available. For purposes of comparison, it would be very interesting to have a study of another population where socioeconomic, cultural, and genetic conditions are different from those found in Santiago.

SUMMARY

The gene and the phenotypic frequencies for the ABO blood groups system were studied in two socioeconomically different subpopulations of Santiago, Chile. The data were taken from the maternity services and blood banks of two hospitals, namely: (1) Hospital San José (HSJ), belonging to the National Health Service and serving mainly low socioeconomic classes; (2) Clínica Alemana (CAL), a private hospital serving middle and high socioeconomic classes.

Gene frequencies were obtained from mothers, children (newborn infants), donors to HSJ blood banks and receptors of CAL blood banks by the maximum likelihood method applied to phenotypes. The stochastic matrix method of mother-child relationship was also used in order to find the fathers' gene frequencies.

Results show a clear difference in phenotypic and gene frequencies between the

two subpopulations. This difference was maintained in two generations (mothers and their children). The fathers' gene frequencies were not different from the mothers gene frequencies in each subpopulation.

These results reinforce the statement that there are socioeconomic forces producing assortative mating in Santiago.

ACKNOWLEDGMENTS

We thank Dr. William J. Schull for his valuable comments and all the personnel in the Hospital San José and Clínica Alemana who made our work possible. We also thank Dr. Arnoldo de Hoyos and Dr. Evelio Fabronni for helping us to solve some statistical problems. Financial support for this study was provided by UNESCO and by the Oficina de Desarrollo Científico y Creación Artística de la Universidad de Chile.

REFERENCES

CAVALLI-SFORZA, L. L., and W. F. BODMER. 1971. The genetics of human populations. Freeman, San Francisco.

HARB, Z., and C. Y. VALENZUELA. 1976. Aplicación de la matriz estocástica madre-hijo en el calculo de frecuencias génicas. Rev. Med. Chile **104**:139–142.

LEBART, L., and J. FENELON. 1973. Statisque et informatique appliquées. 2nd ed. Dunod, Paris.

LI, C. C. 1958. Population genetics. Univ. of Chicago Press, Chicago.

———. 1969. Genética humana. Principios y métodos. Editorial Omega, Barcelona.

LI, C. C., and L. SACKS. 1954. The derivation of joint distribution and correlation between relatives by the use of stochastic matrices. Biometrics **10**:347–360.

MOOD, A., and F. GRAYBILL. 1963. Introduction to the theory of statistics. 2nd ed. McGraw Hill, New York.

MORRISON, D. F. 1967. Multivariate statistical method. McGraw Hill, New York.

PALOMINO, H. 1976. Distribución geográfica de las características genéticas en Chile. Rev. Med. Chile **104**:371–374.

PINTO-CISTERNAS, J., H. FIGUEROA, B. LAZO, C. SALINAS, and C. CAMPUSANO. 1971. Genetic structure of the population of Valparaíso. ABO blood group, color vision deficiency and their relationship to other variables. Hum. Hered. **21**:431–439.

RACE, R. R., and R. SANGER. 1968. Blood groups in man. 5th ed. Blackwell Scientific Publications, London.

ROTHHAMMER, F. 1973. La política de los genes frente a los cambios de estructura. Rev. Med. Chile **101**:247–251.

STEVENS, W. L. 1938. Estimation of blood group gene frequencies. Ann of Eugenics VIII–IX:362–375.