The Effect of Reproductive Compensation on the Long Term Maintenance of the Rh Polymorphism: The Rh Crossroad Revisited

BRUCE R. LEVIN

Department of Human Genetics,
The University of Michigan Medical School,
Ann Arbor, Michigan 48104.

To DATE, the digital computer has been used in genetics primarily as a device for the solution of analytically intractable equations and the analysis of large volumes of data. Occasionally, however, it has been the means for the development of genetically testable population models, particularly those of a multilocus nature. The use of the computer in this latter, more fundamental manner has been referred to as "computer genetics" and briefly involves the simulation of a system of interest using Monte Carlo techniques and a large digital machine. The investigator, through the use of input parameters, asks the system appropriate questions to determine how it will behave under particular conditions. The approach is in many respects analogous to that used by the laboratory experimentalist, and, like the latter, the computer geneticist is also subject to surprises by his results, especially if the model is sufficiently complex.

It is the purpose of this presentation to describe both a deterministic and a stochastic model of a particular genetic system, namely, the Rh locus in man, and more particularly the maintenance of the Rh polymorphism. Here, unlike most of the other polymorphic systems, there is some indication that selection may be operating through the heterozygous offspring of Rh-negative mothers who die due to hemolytic disease of the newborn. This leads to a system of selection against the heterozygote, which, according to the generally accepted theory, presents an unstable situation wherein the allele in lowest frequency must be ultimately eliminated, barring of course recurrent mutation or counter selection. The implication of this latter point is that the Rh polymorphism is at best transitory. It is, however, with us and does not appear to be heading toward fixation.

A number of mechanisms have been advanced to account for the maintenance of the polymorphism at this locus. Among these is reproductive compensation. That is to say, it has been suggested that familial replacement of offspring who have died due to hemolytic disease of the newborn may stabilize the gene frequency. This effect has been considered at the theoretical level by Spencer (1947) and Li (1953), and Glass (1950) has presented evidence which he believed supported this point of view. Reproductive compensation,

This work was supported by U. S. Public Health Service Training Grant GM 00071.

as indicated by this work, can lead to a situation wherein the Rh negative alleles have a selective advantage, even when these alleles are in lower frequency than their Rh-positive counterparts.

In this discussion, an alternate form of the compensation equation will be presented and employed in both the deterministic and stochastic models. Input parameters will then be chosen to facilitate a test of the effect of compensation on the long term maintenance of the Rh polymorphism. The deterministic model has been programed in 1130 FORTRAN and the stochastic model in 7090 MAD. "Hard copy" of both programs is available upon request from the author.

THE DETERMINISTIC MODEL

Consider only the major Rh alleles, D and d. Selection through maternal-fetal incompatibility will operate only on the Dd offspring of dd mothers; the value of the selection coefficient will be a function of the probability of prior immunization and of the probability of the death of an offspring given that immunization has occurred. Since the probability of immunization of a dd mother is dependent upon the genotype of the father, the selection coefficients associated with the two matings in which selection occurs are not equivalent. Let us suppose that s_1 is the probability of the death of a heterozygous offspring born to the mating of a DD father and a dd mother, and that s_2 is the probability of the death of a Dd offspring in a $Dd \times dd$ mating. Since the probability of immunization of the mother is somewhat greater in the $DD \times dd$ mating than in the $Dd \times dd$ mating, because all offspring in the former mating are incompatible, s_1 will be greater than s_2. A more complete discussion of the relation between s_1 and s_2 is presented in the appendix. If t, the compensation coefficient, is defined as the probability of the replacement of a dead offspring, then the fitness of the Dd offspring of the $DD \times dd$ mating is given by

$$W_{Dd_1} = 1 - s_1 + s_1t - s_1^2t + s_1^2t^2 - \ldots + s_1^nt^n - s_1^{n+1}t^n + \ldots$$

where $s_1^nt^n$ is the replacement term and $s_1^{n+1}t^n$ the loss term for the replacement offspring. The above fitness can be expressed as the geometric progression $a + ar + ar^2 + \ldots + ar^n + \ldots$, where a is $1 - s_1$ and r is s_1t. As n approaches infinity, this progression approaches $a/(1 - r)$; hence the fitness of the Dd offspring in this mating tends to $W_{Dd_1} = (1 - s_1)/(1 - s_1t)$. In the $Dd \times dd$ mating, the effect of compensation is somewhat more complex, for the replacement offspring may be of either the Dd or dd genotype with equal probability. This implies that the replacement term for this mating, $s_2^nt^n$, must be continually divided among the two possible offspring. Thus the fitness of the Dd offspring of this mating is

$$W_{Dd_2} = 1 - s_2 + s_2t/2 - s_2^2t/2 + s_2^2t^2/4 - \ldots$$

whereas the fitness of the dd offspring is

$$W_{dd} = 1 + s_2t/2 + s_2^2t^2/4 + \ldots$$

TABLE 1. FREQUENCIES AND FITNESS VALUES FOR OFFSPRING IN THE RH SYSTEM

Mating		Frequency and fitness of offspring		
Female Male	Frequency	*DD*	*Dd*	*dd*
$DD \times DD$	p^4	p^4		
$DD \times Dd$	$2p^3q$	p^3q	p^3q	
$DD \times dd$	p^2q^2		p^2q^2	
$Dd \times DD$	$2p^3q$	p^3q	p^3q	
$Dd \times Dd$	$4p^2q^2$	p^2q^2	$2p^2q^2$	p^2q^2
$Dd \times dd$	$2pq^3$		pq^3	pq^3
$dd \times DD$	p^2q^2		$p^2q^2\left(\frac{1-s_1}{1-s_1t}\right)$	
$dd \times Dd$	$2pq^3$		$pq^3\left(\frac{1-s_2}{1-s_2t/2}\right)$	$pq^3\left(\frac{1}{1-s_2t/2}\right)$
$dd \times dd$	q^4			q^4

From the geometric progression, the fitnesses of these individuals will approach

$$W_{Dd_2} = \frac{1-s_2}{1-s_2t/2}$$

and

$$W_{dd_2} = \frac{1}{1-s_2t/2}$$

It should be pointed out that we have assumed that, although s_1 and s_2 are different, the probability of replacement of a dead offspring, t, is equal for both matings. This assumption seems justified on the basis of the relatively low values of s_1 and s_2 and the general nature of reproductive compensation in human populations, but one can, needless to say, consider two compensation coefficients, t_1 and t_2, corresponding to the two selection coefficients.

The frequencies and fitnesses of the offspring from the matings in this system are presented in Table 1. One is able to obtain from the sum of products of the fitnesses and frequencies the distribution of offspring in the succeeding generation. From the latter, the new frequencies of the alleles can be computed. Since we have chosen to use a computer to solve successively the equations for the gene frequency, it is unnecessary to make the Hardy-Weinberg assumption (see Edwards, 1961). The offspring frequencies of one generation serve, of course, as the parental frequencies of the next. A diagrammatic representation of the steps used in the construction of this computer program is presented as a logical flow chart in Fig. 1.

THE STOCHASTIC MODEL

The stochastic model has been designed to obtain some insight into the stability of the system in finite populations and involves the Monte

Input p, s_1, s_2, t, gen

Initialize genotypic frequencies

$q = 1 - p$

$D = p^2$

$H = 2pq$

$R = q^2$

$G = 0$

Start iterations

$G = G + 1$

α

$Dx = D^2 + DH + H^2/4$

$Hx = DH + H^2/2 + HR/2 + DR(1 - s_1)/(1 - s_1t) + HR(1 - s_2)/2(1 - s_2t/2) + DR$

$Rx = H^2/4 + HR/2 + HR/2(1 - s_2t/2) + R^2$

$\text{Tot} = Dx + Hx + Rx$

$$q = \frac{Hx + 2Rx}{2\text{Tot}}$$

$D = Dx/\text{Tot}$

$H = Hx/\text{Tot}$

$R = Rx/\text{Tot}$

Print out G, q

Go to α until $G >$ gen

FIG. 1. Logical flow diagram of a deterministic model for selection and reproductive compensation in the Rh system.

Carlo simulation of the mating and selection processes that comprise the model previously outlined. The simulation used, although it is specific in the sense of being primarily for a system of selection due to maternal-fetal incompatibility, is also representative of a simulation of a monogamous mating system with discrete generations. The program under discussion here is similar in many respects to that used earlier by Schull and Levin (1964).

Since this presentation is primarily concerned with methodology, a reiteration of the basic technique seems appropriate. In order to make a decision with respect to a particular event in a Monte Carlo simulation, a pseudo random number x is generated ($0 \leq x \leq 1$) and compared to the probabilities of the various outcomes of the event. In the simplest case, that of a binomial decision, only one comparison has to be made. As an illustration, consider the determination of the sex of an offspring. If the random number is less than or equal to the probability of the offspring being a male, the offspring will be considered a male; if the random number is greater than this probability, the offspring will be considered to be a female. When decision-making involves a choice among three possibilities, only two comparisons need be made. As an illustration of this, consider the assignment of a genotype to an individual randomly drawn from a parental pool. If $P(DD)$, $P(Dd)$ and $P(dd)$ are the frequencies of the DD, Dd, and dd males, respectively, and $P(DD) + P(Dd) + P(dd) = 1$, then the male parent is chosen in the following manner: If the random number x is less than or equal

to $P(DD)$, the male parent will be DD. If x is greater than $P(DD)$ but less than or equal to the sum $P(DD) + P(Dd)$, then the male parent will be Dd. If x is greater than $P(DD) + P(Dd)$, the male parent will be dd. If sampling without replacement is desired, at this point the total number of males of the genotype just chosen is diminished by one and the male parental frequencies recomputed.

In general, the choice of outcome of a series of mutually exclusive and exhaustive events can be viewed as determining the position of a random point on a line of unit dimension which is divided into segments such that the size of the ith segment is relative to the probability of the ith outcome for every i. As an example of this general case, consider the determination of offspring number for each mating pair. The possible outcomes of this event are 0, 1, 2, . . . max offspring, which occur with probability $P(0)$, $P(1)$, $P(2)$, . . ., $P(\max)$, respectively. If x is less than or equal to $P(0)$, the mating pair will have no children. If $P(0) < x \leq P(0) + P(1)$, they will have one child. If $P(0) + P(1) < x \leq P(0) + P(1) + P(2)$, they will have 2 children. In general, they will have n children if

$$\sum_{i=0}^{n-1} P(i) < x \leq \sum_{i=0}^{n} P(i)$$

We have assumed the distribution of offspring number to be Poisson with parameter λ. To maintain a stable population, λ was set equal to the desired population size divided by the number of mating pairs, which in a monogamous system is the minimum number of parents available in either the male or female parental pool. The maximum number of offspring possible, *max*, was arbitrarily assumed to be $2\lambda + 8$; this truncation of the distribution obviously introduces an element of approximation albeit a negligible one.

Once a decision is made, it can be implemented in a number of ways. If it is a matter of storage, as in the case of a surviving offspring, a storage indicator, initially set equal to zero at the start of the generation, is incremented by 1. This storage indicator must, of course, be specific for each sex and genotype. If the decision involves the determination of a control value, e.g., the offspring number, a variable will be set equal to a previously specified number. If the outcome of one event is to be used in a decision with respect to another, for example, in the determination of the occurrence of prior immunization, an indicator will be set equal to 0 or 1 depending upon whether immunization has occurred or not.

A logical flow diagram of the complete stochastic program is presented in Fig. 2, and two flow charts of specific segments of the program are presented in Figs. 3 and 4. These latter two represent the actual programing steps used.

The population size, probability of immunization, probability of death after prior immunization, probability of compensation, and the initial gene frequencies are specified and entered into the computer. In addition, control parameters such as the desired number of generations, the number of replicas, and an initializing number for the random number generator are entered. The output of the program is the gene frequency and population size of each

Input p, n, s, t, gen, rpl
Initialize all values and indicators

I. Parental pool
  (1) Choose parents
  (2) Determine number of offspring

II. Generate offspring
  (1) Gamete determination
  (2) Fertilization
  (3) (Sex determination)

III. Selection
  (1) Store
  (2) Do not store
    (a) compensate
  (3) (Sex determination)

IV. Continuation decisions
  (1) Return to I
  (2) Return to II
  (3) Start next generation
    G = 1, 2, ... Gen
  (4) Start next run
    R = 1, 2, ... Rpl
  (5) Finish—Print desired output

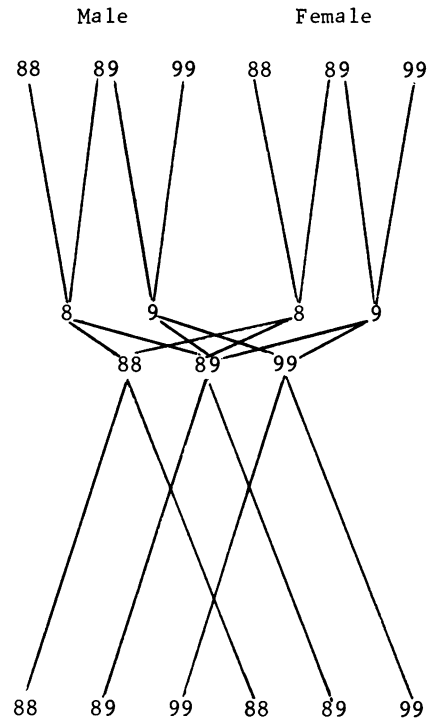

FIG. 2. Logical flow diagram for a Monte Carlo simulation of a single human breeding unit.

generation and each run. After the desired number of runs is completed, the mean and variance of each gene frequency and population size is printed for each generation.

DISCUSSION OF THE RESULTS OF RUNS MADE WITH THESE PROGRAMS

For this presentation two levels of selection were considered, one relatively high and the other in the general range of selection intensities considered by Li (1953). The probability that an Rh-negative mother will be immunized by her Rh-positive offspring, m, was set equal to 0.50 and 0.25, and the probability of death of the offspring after prior immunization, S, was set equal to 0.40 and 0.20 for the high and low selection systems, respectively. Following the scheme presented in the appendix, and assuming a Poisson distribution for offspring number with $\lambda = 2$, the probabilities of death of offspring in this system are as follows: In the high selection system $s_1 = 0.1598$ and $s_2 = 0.0946$. In the low selection system $s_1 = 0.0473$ and $s_2 = 0.0259$. As indicated earlier, m and s serve as the input parameters for the stochastic model.

In the deterministic model, at both levels of selection (see Figs. 5 and 6) an increase in the level of compensation increases the fitness of the Rh-negative allele, as indicated by the direction of gene frequency change. There is, however, an exception to this situation. In the portion of the graphs above an Rh-negative frequency of 0.80, the increase in the rate of com-

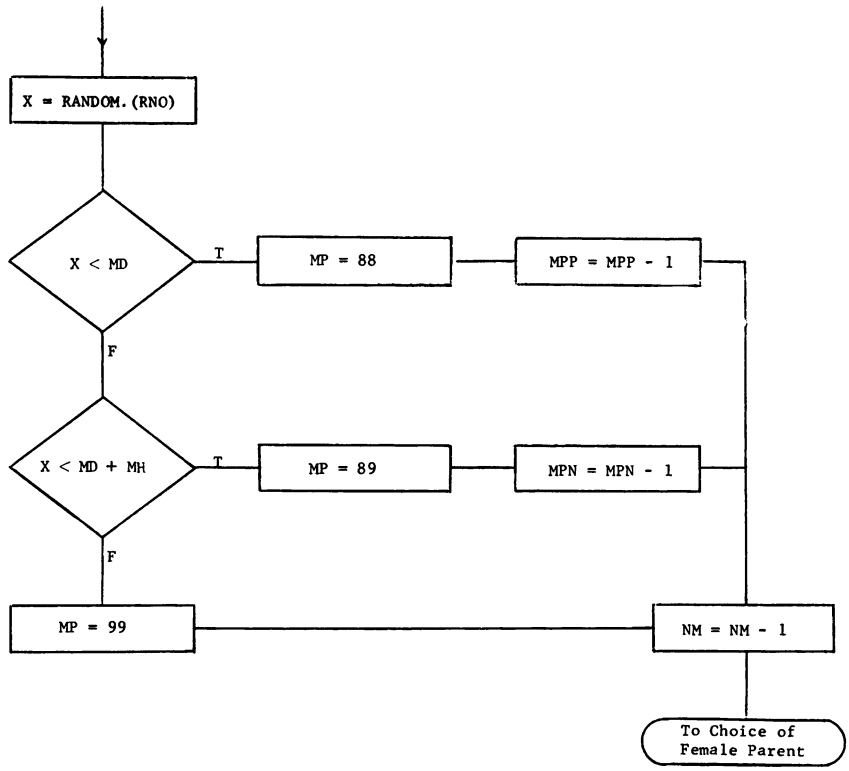

FIG. 3. Computer flow diagram for the choice of a male parent in a Monte Carlo simulation of a single human breeding unit. Definitions: NM = total number of males; MPP = number of *DD* males; MPN = number of *Dd* males; MD = MPP/NM = *f(DD)*; MH = MPN/NM = *f(Dd)*.

pensation reduces the fitness of this allele. In order to see what is occurring in this almost paradoxical situation, the contribution to the fitness of each offspring must be considered for each mating in which selection occurs. The fitness of each genotype is the sum of the fitness values of each offspring of that genotype weighted by the frequency of that offspring and divided by the frequency of that genotype in the absence of selection. Using the Hardy-Weinberg values for this latter frequency, we have

$$W_{Dd} = \alpha + (pq/2)\left(\frac{1 - s_1}{1 - s_1 t}\right) + (q^2/2)\left(\frac{1 - s_2}{1 - s_2 t/2}\right)$$

and

$$W_{dd} = \beta + pq\left(\frac{1}{1 - s_2 t/2}\right)$$

where α and β are the contributions to the fitness of nonselected offspring. As can be observed from the above, an increase in the value of q increases the contribution of the $q^2/2$ term. The contribution to the fitness of the

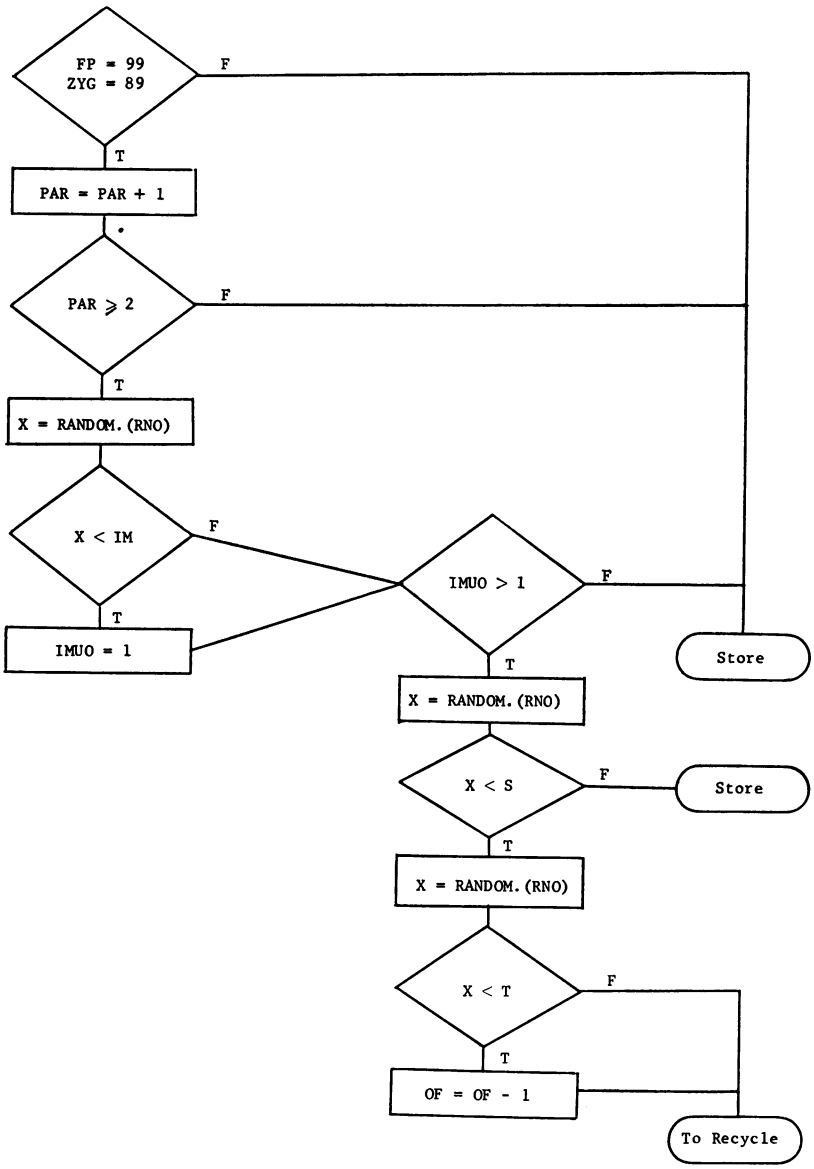

FIG. 4. Computer flow diagram for selection and compensation in the Rh system in a Monte Carlo simulation of a single human breeding unit. Definitions: IM = prob. (of immunization); S = prob. (death, prior immunization); T = prob. (compensation for a dead offspring).

$pq/2$ and the pq terms is maximal when $p = q$. Thus at high values of q, the heterozygote has a proportionally higher fitness. This fitness of the heterozygote is enhanced by increasing t, and at a particular value of q, the gain in fitness of the d allele due to the increased fitness of the dd individual is offset by the gain in fitness of the Dd individual. This reverses the effect of t.

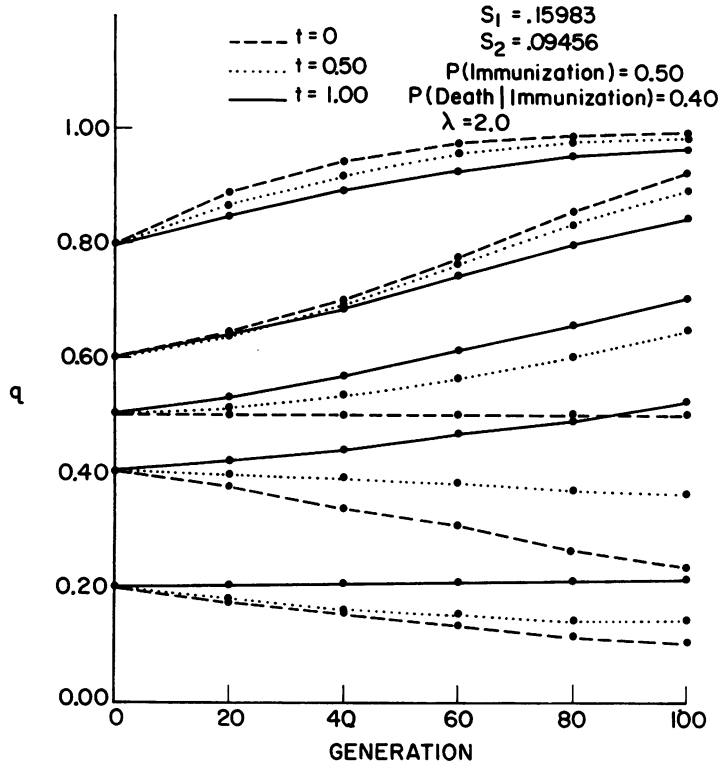

Fig. 5. High level of selection in the deterministic model.

As earlier work indicated, through reproductive compensation, the Rh-negative allele can acquire a selective advantage at frequencies other than $q > p$. As q increases, a smaller value of t is required to give the Rh-negative allele a selective advantage. At the intermediate value of q, initially equal to 0.4 and $t = 0.5$, there is virtually no change in the frequency of the Rh-negative allele. At lower frequencies of this allele, the reduced selective pressure resulting from the lower frequencies of the appropriate matings reduces the rate of gene frequency change.

In addition to the two sets of input parameters described for the stochastic model, a system of no selection was also used. This latter system is intended to facilitate a comparison of the selection systems with a genetic drift situation. The drift system is presented in Fig. 7; the high and low selection systems are presented in Figs. 8 and 9, respectively. A desired population size of $N = 100$ was used for all runs. In these figures, the mean gene frequency of the indicated number of runs is plotted as a function of generation number.

Of primary importance is the similarity of behavior of the genetic drift system in Fig. 7 and the low selection system in Fig. 9. The higher level system, however, presents little similarity to the drift situation. The rate of change in the Rh-negative frequency in this system is quite pronounced. Further support for the minimal effect of the low level of selection on gene

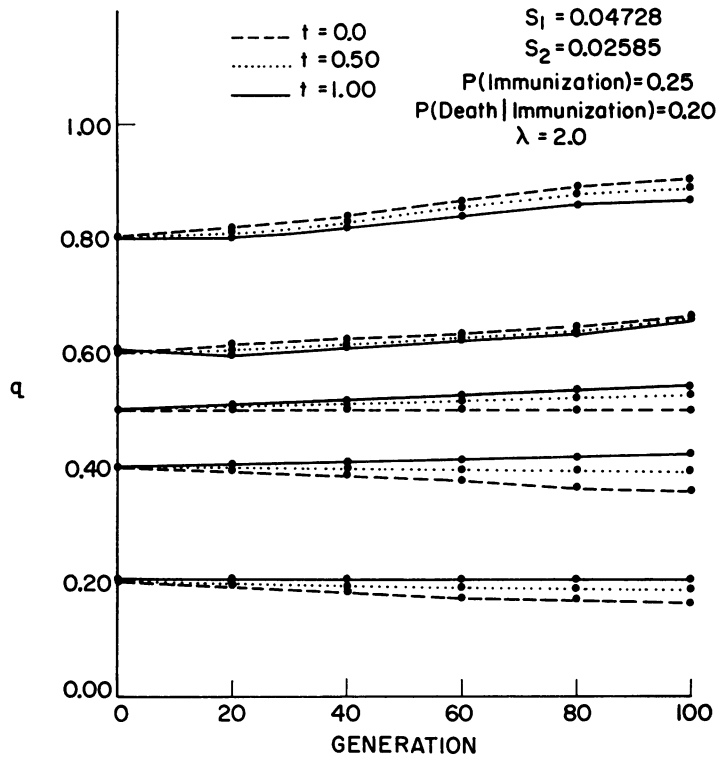

FIG. 6. Low level of selection in the deterministic model.

frequency change is indicated by a comparison with the deterministic model. If one considers the mean Rh-negative level at the 100th generation, the relation between the gene frequency and the level of compensation is not the same in the stochastic model as in its deterministic counterpart. The directional force of increasing compensation seems to be overridden by the stochastic element, and the compensation effect appears to be less patterned. The compensation relationship is, however, more generally met in the high selection system. At low rates of selection, the deterministic drive through selection is apparently offset by stochastic elements.

No doubt an increase in the population size or total number of runs would make the outcome of selection in these systems more deterministic. In this case, the stochastic model should approach the deterministic, but there is, aside from the asymptotic approach of the former to the latter, a possible qualitative difference between the models. As indicated earlier, the rate and direction of selection in these systems is a function of the gene frequency. The latter is, however, subject to random fluctuations in the stochastic model; each change in gene frequency alters the intensity and even the direction of the selective forces. Otherwise stated, there exists an asymmetry in the directional forces, upon which is overlayed a symmetric stochastic element; this leaves a net directional element. The general direction and magnitude

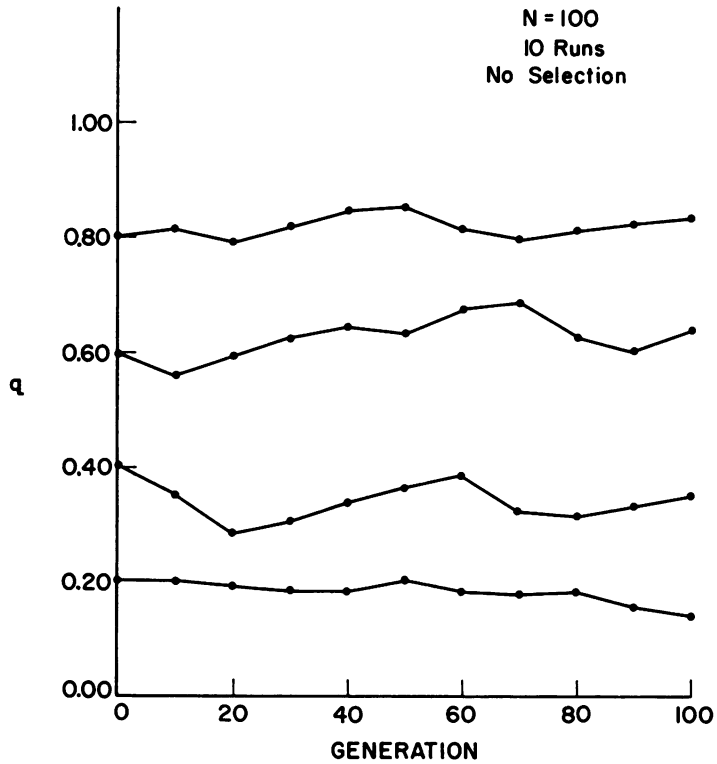

FIG. 7. No selection in the stochastic model.

of the asymmetric deterministic element is indicated by the deterministic model. The importance of this effect is dependent upon the intensity of the selective forces and magnitude of the stochastic element. A further understanding of this phenomenon would require much more extensive use of the stochastic model than has been possible thus far.

CONCLUSIONS

Although the general compensation formulation has been altered, there seems to be little or no qualitative difference in the general effect of reproductive compensation from that presented by Li (1953). That is, it is possible to obtain an equilibrium level for the Rh polymorphism at frequencies other than $p = q = 0.5$, but this equilibrium is unstable. This statement itself may not be sufficient to formulate a rigid hypothesis for the maintenance of the Rh polymorphism or forecast its future, but if reproductive compensation is incorporated into whatever schema is advanced, intensity of selection can become a population controlled variable. Thus, if the rate of reproductive compensation is varied, a population, even without the aid of modern medicine, is able to alter the intensity of selection at this locus.

This now brings us to the question of the rate of reproductive compensation in human populations. A more extensive discussion of this question is presented

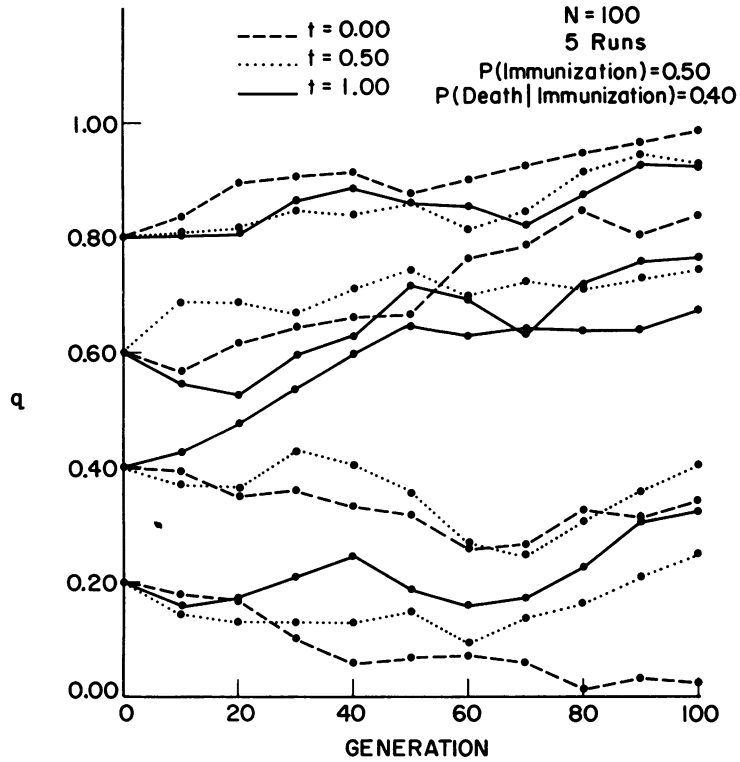

FIG. 8. High level of selection in the stochastic model.

in Levin (1967), but in summary: Reproductive compensation in human populations may be either volitional or nonvolitional. The former involves a conscious decision on the part of parents to replace a dead offspring, for whatever reason. The latter, the nonvolitional form, stems from the inability of the family to support as many children as are born. In a situation of limited resources, if a child dies for genetic reasons, its death enhances the survival probability of its sibs, younger as well as older, and may thus ultimately lead to ensuring that a more or less constant number of the sibs will reach reproductive age. It is only in those situations where the death of one offspring does not enhance the survival probability of its sibs that reproductive compensation would not occur. The latter would hold true if there is an abundance of resources and a need to produce as many children as possible.

Both the volitional and nonvolitional forms of reproductive compensation may reflect particular phases of human existence. The volitional form will more often obtain, it would seem, in a modern society, where family size is relatively small and primarily socially determined. The nonvolitional form would appear more characteristic of a primitive hunting and gathering society, where the resources are limited and the period of nursing is long. Reproductive compensation would have been minimal, presumably, in an early

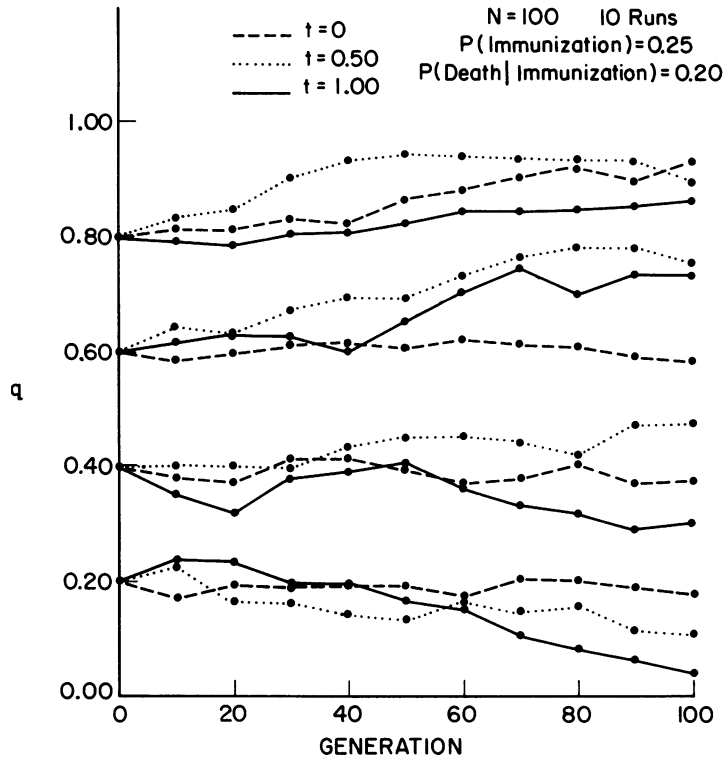

FIG. 9. Low level of selection in the stochastic model.

agrarian society, where food was relatively abundant and large numbers of children were an asset as a form of labor.

The general implication of the above is that in the course of human history, as man went from hunter and gatherer to early farmer and to his present state, with his social change went a change in the rate of reproductive compensation. This change in the rate of reproductive compensation could result in changes in both the level and direction of selection at the Rh locus. In addition to this temporal variation in selective pressures, undoubtedly a spatial one has also existed, for not all populations would have the same frequencies of the Rh-negative alleles nor rates of reproductive compensation. These factors, when taken in conjunction with migration between human populations and the relatively low rates at which selection operates at this locus, could account for both the large variation in Rh frequencies and the long term maintenance of the polymorphism.

SUMMARY

A deterministic and stochastic model for selection and reproductive compensation in the Rh system are presented. A detailed description of the computer techniques required for the development and construction of these models is given. The effects of compensation are considered for both a high

and low level of selection. The occurrence of reproductive compensation and its consequence in the maintenance of the Rh polymorphism are discussed.

ACKNOWLEDGMENT

I would like to thank my mentor, Professor William J. Schull, for the critical advice, encouragement, and patience which enabled me to prepare this manuscript. I would also like to thank Dr. Charles F. Sing for reading this manuscript.

APPENDIX

Let S be the probability that a *Dd* offspring born to a *dd* mother who has been previously Rh immunized will die. Assume that S is a mean value for a given family size and accounts for the intensity of the immune response. Let m be the probability of Rh immunization at a given birth and let λ be the mean number of offspring produced by a mating pair.

In the mating of a *DD* male with a *dd* female, the probability of immunization of the mother at any birth is m. In the mating of a *Dd* father with a *dd* mother, the probability of Rh immunization at any birth is $m/2$, due to the segregation in this mating. The probability of any given child dying is the probability of prior immunization multiplied by the probability of death. For the nth child of a $DD \times dd$ mating, the probability of death is

$$[1 - (1 - m)^{n-1}]S$$

where the $(1 - m)^{n-1}$ represents the probability of no prior immunization. For the nth child of the $Dd \times dd$ mating the probability of death is

$$[1 - (1 - m/2)^{n-1}]S$$

Given a Poisson distribution for the number of offspring produced by a mating pair, the mean probabilities of an offspring dying in the $DD \times dd$ and $Dd \times dd$ matings, say s_1 and s_2 respectively, are determined in the following manner:

$$s_1 = \frac{e^{-\lambda}\lambda^2}{2}[1 - (1 - m)]S + \frac{e^{-\lambda}\lambda^3}{3!}[1 - (1 - m)^2]S + \ldots$$

$$s_2 = \frac{e^{-\lambda}\lambda^2}{2}[1 - (1 - m/2)]S + \frac{e^{-\lambda}\lambda^3}{3!}[1 - (1 - m/2)^2]S + \ldots$$

and the ratio of s_1/s_2 is

$$\frac{s_1}{s_2} = \frac{Se^{-\lambda}\sum_{n=2}^{\infty}\frac{\lambda^n}{n!}[1 - (1 - m)^{n-1}]}{Se^{-\lambda}\sum_{n=2}^{\infty}\frac{\lambda^n}{n!}[1 - (1 - m/2)^{n-1}]}$$

$$= \sum_{n=2}^{\infty}\frac{[1 - (1 - m)^{n-1}]}{[1 - (1 - m/2)^{n-1}]}$$

REFERENCES

Edwards, A. W. F. 1961. The limitations of population models. In *Proceedings of the IInd International Congress of Human Genetics*, pp. 222–223.

Glass, B. 1950. The action of selection on the principle Rh alleles. *Amer. J. Hum. Genet.* 2: 269–278.

Levin, B. R. 1967. Reproductive compensation: Its occurrence and importance in the determination of Darwinian fitness. (Manuscript.)

Li, C. C. 1953. Is the Rh facing a crossroad? A critique of the compensation effect. *Amer. Naturalist* 87: 257–261.

Schull, W. J., and Levin, B. R. 1964. Monte Carlo simulation: Some uses in the genetic study of primitive man. In *Stochastic Models in Medicine and Biology*. J. Gurland, ed. Madison: University of Wisconsin Press, pp. 179–196.

Spencer, W. P. 1947. On Rh gene frequencies. *Amer. Naturalist* 81: 237–240.