Studies on Genetic Selection in a Completely Ascertained Caucasian Population. II. Family Analyses of 11 Blood Group Systems

C. F. SING,[1] D. C. SHREFFLER,[1] J. V. NEEL,[1] AND J. A. NAPIER[2]

In a preceding paper [1] we have described the Tecumseh Community Health Study and the opportunity it afforded to investigate the relationship between phenotype and/or genotype and a number of indices of genetic selection. Data were presented on first-order associations between genotypes, and on genotype-related age trends and sex effects, for 12 different genetic systems. In the present paper we will analyze the results of matings with respect to 11 of these 12 systems, with emphasis on tests for departures from Mendelian expectations which might indicate the action of selection. For the twelfth system, transferrin, only 94 individuals were of a type other than C, an inadequate sample for this type of analysis.

The majority of past searches for evidence of selection—including some of our own—have emphasized the few "positive" findings to emerge out of many tests, with no proper consideration of the magnitude of the selection coefficients which might be implied. Here, we shall emphasize a consideration of the selection coefficients which our study either indicates or excludes at various significance levels. It will be shown that this study—the most extensive to date—is unable to detect selection coefficients of the magnitude commonly used in the formulations of population genetics. On the other hand, the selection coefficients implied by the few significant deviations of genetic ratios which were detected in this study are so large that from the biological standpoint they seem unlikely. Some problems in the design of future studies will also be discussed.

SUBJECTS AND METHODS

The Genetic Variables

The 11 genetic systems (32 factors) which were available for analysis are described in table 1. The Lewis system was classified on the basis of both red blood cell type and secretion of ABH substance in saliva but, because the secretor data were deemed more reliable (see [1] for a discussion), only these were included in the analysis. Details regarding the methods of typing are given in a previous paper [1].

Received June 26, 1970; revised September 15, 1970.

This study was supported by U.S. Atomic Energy Commission contract AT(11-1)-1552 to the Department of Human Genetics and by U.S. Public Health Service program project grant H-6378 to the Cardiovascular Research Center, University of Michigan. Author Shreffler is supported by U.S. Public Health Service career development award HE 24,980.

[1] Department of Human Genetics, University of Michigan, Ann Arbor, Michigan 48104.

[2] Department of Epidemiology, University of Michigan, Ann Arbor, Michigan 48104.

TABLE 1

ENUMERATION OF GENETIC SYSTEMS EMPLOYED IN THE STUDY

System	Factor
ABO	A_1, A_2, B
MNSs	M, N, S, s
Rh	C, c, D, E, e
Kell	K, k, Kp^a, Kp^b
Duffy	Fy^a
Kidd	Jk^a or Jk^b
P	P_1
Lewis secretion	Le^a, Le^b
ABH secretion	A, B, H
Haptoglobin	Hp1, Hp2
Gc	Gc1, Gc2

The Population Studied

The sample of families considered here is drawn from the 9,182 Caucasoid individuals typed in connection with the Tecumseh Community Health Study. Blood and saliva specimens were taken for genetic typing as the subjects presented themselves for physical examination; kindreds were constructed, after examinations were concluded, from family information collected as part of the study. Occasionally, for various reasons, one of the three types of specimens (red blood cells, serum, or saliva) was not collected from an individual, or a genetic classification was not completed because of technical difficulty. This accounts for the variation in number of individuals analyzed for the different systems.

Of these 9,182 persons, 13.6% could be assembled into three-generation kindreds. However, for purposes of the analysis reported here, only two-generation kindreds were recognized; hence, an individual who was a child in generation 2 may also appear in tabulations as a parent of generation 3. We will refer to the two-generation kindred as the nuclear family in this report. Furthermore, for the analyses to follow, we will treat all these nuclear families as genetically independent.

In many kindreds, not all members of the nuclear family were typed. When neither parent was in the study, a family record was constructed only if two or more siblings were available. There were 3,482 (37.9%) individuals in the sample of 9,182 who did not have any relatives included in the study and hence could not provide family information. Table 2 presents the distribution of families by number of children tested and availability of the parents of the 2,507 nuclear families which were assembled. These families represented 3,567 parents and 5,700 children. Twenty-one percent of the families had only one parent, and 18% had two or more children but no parents in the study. When only one parent presented for examination, it was usually the mother. Sixty percent of the family records (1,517) included both parents, and these with their 3,592 children provided the data for the analysis of genetic ratios. Although at least one child was available for comparison with the parents in each of these 1,517 families, not all children ever born of these parents were included. An estimate of the level of completeness of the record for a nuclear family was obtained by checking 158 families (approximately 10%) randomly selected from a listing of the 1,517 families. These 158 families were found to have 474 children, with an average of three children per family. The distribution by sibship size is given in table 3. The table also records, for each sibship size, the number of sibships in which all members were in the study. It should be noted that, in order to be included in this study, a "child" had to be living in the Tecumseh community at the time medical examinations were given (1962–1965). Therefore, sibships with only one child provide a bias in evaluation of completeness. Of the 459 children in the 143 selected

TABLE 2

DISTRIBUTION OF FAMILIES BY NUMBER OF CHILDREN TESTED AND STATUS* OF PARENTS FOR TECUMSEH POPULATION

No. Children Tested	Both Parents	Father Only	Mother Only	Neither Parent	Total
1	440	67	223	(3,482)†	730
2	501	23	100	325	949
3	313	10	65	83	471
4	166	4	19	28	217
5	60	3	11	12	86
6	21	1	4	7	33
7	9	.....	1	.....	10
8	5	.....	1	1	7
9	2	.....	.....	1	3
10	.....	.....	1	.....	1
Total	1,517	108	425	457	2,507
No. Children represented	3,592	180	798	1,130	5,700
Average no. children per family	2.37	1.67	1.88	2.47	.....

* Based on the presence of the parents in the sample of 9,182 individuals available for study.

† Families of size 1 with neither parent available provide no information for a family analysis, hence not included in the analysis as a child.

TABLE 3

COMPLETENESS OF DATA FOR 158 RANDOMLY SELECTED NUCLEAR FAMILIES ACCORDING TO SIBSHIP SIZE

	Completed Sibship Size						Total
	1	2	3	4	5	6–9	
No.	15	56	39	27	11	10	158
No. with all members in study	15 (1.00)*	40 (0.71)	19 (0.49)	14 (0.52)	4 (0.36)	1 (0.10)	93
Expected no.†	15	31	16	8	2	.....	.....

* Proportion of families of this sibship size which were included.

† Computed as (number of completed sibships of size s) $\times$ (the proportion of children examined = 0.743)^s.

families with at least two children, 340 (74.3%) were examined, gave an adequate sample of blood for typing, and were included in the analysis to be presented here. There were 26 children who were examined but from whom no blood was drawn (or amount was inadequate). Of the remaining 93 children, seven died before the study started, one was born after the study ended, 13 were living in the area but declined examination, six were away at school or in the service, and 66 were living elsewhere than in Tecumseh.

The proportion of completed sibships included in the study (table 3) decreases as the size of sibship increases. The observed number of families exceeds the expected number, based on percentage of children examined in the sample of 158 families for all sibship sizes. We interpret this to indicate that families who enter the health study tend either to bring all members

into the clinic for examination more often than would be expected by random distribution of examination among children of the sibships, or else tend to show poor participation.

After the 2,507 nuclear families were established, a comparison (by computer) was made to detect any inconsistencies with the expected pattern of Mendelian inheritance between the parental and offspring typings (and among offspring when parents were not typed). When a discrepancy was found, it was flagged for retyping if inspection of the individual records did not reveal the nature of the problem. Six types of discrepancy were identified, of which four did not require retyping. These four were errors in transcribing the identification number of an individual's record, errors in coding a phenotype in the laboratory, errors in transcribing the data from written laboratory records to punched cards, and unrecorded adoption. For those families in which a discrepancy remained unresolved after a records check, the individuals were retyped for the systems which identified the inconsistency. This distinguished between a group of nuclear families with a blood-typing error and a residual class of families in which there was an inconsistency which could not be resolved. The latter group of families were classified as having a biological discrepancy. Table 4 gives the number of nuclear families, by the type of inconsistency, which were discovered for each of the 11 genetic systems. The success of a system in detecting inconsistencies in family records is determined by the number of alleles, the frequencies of the alleles, dominance, family size, and inclusion of parental records in the study. We note that these criteria are reflected by the larger percentage of inconsistencies detected by the MNSs and Rh groups than by the P and Secretor groups.

After correction of records for the four types of clerical error, we are left with 7.82% (196) of the 2,507 nuclear families with a discrepancy in one or more systems. We emphasize that, since the individuals were distributed in these families according to table 2 (families varied in size and inclusion of parents), the probability of detecting an error was not equal for each nuclear family. Thus, our presentation in table 4 is necessarily a conservative estimate of the degree to which various sources of inconsistency occur in a sample of this nature. A measure of the variable ability to detect the 222 independent families with discrepancies is reflected by considering the relationship between size of family and incidence of "errors" given in table 5. The "power" of our search for errors was approximately linear with family size.

In an effort to resolve these discrepancies, 627 individuals were retyped for one or more systems. In 101 cases, an error in typing was demonstrated. The exact level of demonstrated typing error ranged from 0.01% for the ABO and Kell systems to 0.24% for the Gc system (table 4). Overall, 198,234 typings were done, and 0.05% were found to be in error. However, it is obvious that this approach can only give a minimal estimate of error because not all mistypings will result in inconsistencies between parents and offspring. The role of these errors in confounding a search for selection effects will be discussed later.

After correction of records for typing errors, there remained 109 nuclear families in which the genotype of one or more children was inconsistent with that of the parents. The non-random birth order of these children within the families suggests that in the majority of instances the result was due to a discrepancy between stated and biological parentage. The bias is illustrated by examining the record of the firstborn in each of these families. Of the 109 independent nuclear families in this class, 97 were identified by one discrepant child, 11 by two, and one by three. In 60% of the 97 families identified by one discrepant child, the child was the first issue of the recorded parents, whereas in the material as a whole, firstborn children constitute 44% of the total. This figure is in close agreement with the report of Schacht and Gershowitz [2] that, in another Michigan sample of Caucasoids, 57% of the parental exclusions were firstborn.

The nuclear family records were revised to exclude these 122 children. Furthermore, 17 families were found to involve a "sibship exclusion" (i.e., it was impossible to determine which offspring was discordant), hence the entire family had to be dropped from the analysis. A further reduction in the number of nuclear families available for the genetic-ratio analysis occurred for given genetic systems because of incomplete typing data on some children. The family records available for analysis after these adjustments are summarized in table 6.

TABLE 4

ANALYSIS OF FAMILY RECORDS IN WHICH INCONSISTENCIES WERE DETECTED IN TECUMSEH STUDY

SOURCE OF INCONSISTENCY	NO. FAMILIES WITH AN INCONSISTENCY BY SYSTEM										ERROR SUMMARY		
	ABO	MNSs	Rh	Kell	Duffy	Kidd	P	Secretor (Lewis and ABH)	Hapto-globin	Gc	Total	No. Inde-pendent Families*	Percentage of 2,507 Family Records
Mislabeled I.D.	.....	3	4	1	.....	.....	.....	.....	.....	.....	8	8	0.32
Coding error	.....	5	4	.....	.....	.....	.....	.....	.....	.....	9	9	0.36
Transcribing error	2	6	3	.....	.....	7	1	1	1	.....	21	8	0.32
Unrecorded adoption	1	.....	.....	.....	.....	.....	.....	.....	.....	.....	1	1	0.04
Blood-typing errors	3	13	19	3	4	7	10	3	14	15	91	87	3.47
Remaining biological discrepancies	20	39	49	4	4	6	5	7	13	21	168	109	4.35
Total	26	66	79	8	8	20	16	11	28	36	298	222	8.86
Sibship Exclusions	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	17	0.67
Percentage of 2,507 family records	1.04	2.63	3.15	0.32	0.32	0.80	0.63	0.44	1.12	1.44	.....	.....	9.53
Percentage of total typings in error	0.01	0.04	0.04	0.01	0.05	0.15	0.15	0.02	0.11	0.24	0.05	.....	.....

* This entry may not equal the sum of the entries because a family is detected by more than one system.

TABLE 5

RELATIONSHIP BETWEEN SIZE OF NUCLEAR FAMILIES AND FREQUENCY OF DISCREPANCY BETWEEN PARENTS AND OFFSPRING

No. Children	No. Nuclear Families		
	Total	With "Error"	Percentage
1.....	730	36	4.93
2.....	949	68	7.16
3.....	471	56	11.89
4 or more.....	357	62	17.37

The Statistical Methods

The primary aspect of the Tecumseh study which is reported here is an analysis to determine whether the distribution of genetic factors among children within sibships deviated from Mendelian expectations. Here each genetic system was analyzed independently in a search for first-order distortions of genetic ratios which might indicate the operation of selection. An analysis to detect the deviation of the joint distribution of factors of two or more systems will be presented in a subsequent paper.

Four analyses of each genetic system were conducted. The first, of mating-type frequencies, was done to determine if mating was at random with respect to the genetic factors of each system. While assortative mating with respect to these systems is unlikely, it must be considered before an evaluation of the evidence for selection can be undertaken. The second, third, and fourth analyses of each genetic system were conducted to determine whether evidence for selection was present in this population. These analyses were, respectively: number of children included in the study per mating; parental and offspring phenotype distributions; and genetic ratios for matings, classified according to segregating phenotype, mating type, reciprocal (i.e., sex which is the segregating phenotype), and size of family. To reduce the possible bias in evaluation of selection which might be introduced by a nonrepresentative sample of incomplete family records, only families with both parents tested were used in the first, second, and fourth analyses. All available parents and children were utilized for the third analysis (phenotype distributions).

The standard χ^2 analysis for goodness-of-fit was employed to establish whether the mating-type frequencies, number of children typed per mating type, and phenotype frequencies of children (or parents) deviated significantly from expectations. Expected mating-type frequencies and the frequency distribution of parental phenotypes were derived in the usual manner, assuming Hardy-Weinberg equilibrium conditions using gene frequencies estimated from the parents by the maximum-likelihood method discussed in [1]. The expected Hardy-Weinberg distribution of children's phenotypes was computed in two ways: from gene frequencies estimated from the children's phenotypes, and from parental gene frequencies. The expected proportion of children for a mating type was taken to be the observed frequency of the mating.

In an analysis of genetic ratios, we have two objectives. First, we wish an estimate of the parameter, say θ, which defines the proportion of segregant phenotypes in the children of either backcross (one parental phenotype segregating) or intercross (both parents segregating) matings. (The phenotypes of parents are presented throughout with the allele defining the nonsegregant offspring given first; e.g., A_1 from $A_1 \times O$ and M from $MN \times M$ are the implied nonsegregant offspring.) Second, we wish to determine whether the observed estimate,

TABLE 6

SUMMARY OF NUMBER OF TESTS PERFORMED CLASSIFIED BY AVAILABILITY OF PARENT AND GENETIC SYSTEM

Genetic System	Both Parents	Father Only	Mother Only	Neither Parent	Totals	Percentage of Children Untested
ABO	1,498 (3,313)* (226)†	109 (176) (9)	420 (734) (49)	463 (1,130) (7)	2,490 (5,353) (291)	5.16
MN	1,499 (3,213) (225)	109 (176) (9)	421 (734) (50)	461 (1,228) (9)	2,490 (5,351) (293)	5.19
MNS	1,368 (3,069) (224)	99 (163) (9)	404 (713) (52)	380 (920) (9)	2,251 (4,865) (294)	5.70
Rh (total)	1,370 (2,994) (267)	169 (257) (82)	464 (757) (132)	487 (1,075) (80)	2,490 (5,083) (561)	9.94
Kell-Kk	1,497 (3,229) (267)	110 (180) (9)	421 (729) (55)	462 (1,165) (10)	2,490 (5,303) (341)	6.04
Kell-Kp^a Kp^b	1,381 (3,086) (247)	101 (170) (9)	405 (706) (54)	383 (932) (7)	2,270 (4,894) (317)	6.08
Duffy	1,497 (3,264) (272)	109 (175) (10)	421 (730) (54)	463 (1,068) (71)	2,490 (5,237) (407)	7.21
Kidd-Jk^a	982 (2,212) (165)	73 (123) (4)	307 (515) (46)	329 (653) (92)	1,691 (3,503) (307)	8.06
Kidd-Jk^b	267 (549) (63)	13 (21) (5)	73 (112) (13)	90 (125) (48)	443 (807) (129)	13.78
P	1,497 (3,298) (240)	109 (176) (9)	422 (732) (53)	462 (1,129) (7)	2,490 (5,335) (309)	5.47
Lewis secretion	1,491 (3,079) (435)	108 (171) (11)	417 (726) (55)	473 (1,142) (22)	2,489 (5,118) (523)	9.27
ABH secretion	1,304 (2,718) (430)	117 (175) (41)	434 (737) (83)	634 (1,129) (328)	2,489 (4,759) (882)	15.64
Haptoglobin	1,347 (2,436) (810)	122 (182) (39)	462 (722) (154)	559 (1,167) (134)	2,490 (4,507) (1,137)	20.15
Gc	1,356 (2,473) (802)	122 (181) (39)	461 (728) (149)	551 (1,142) (130)	2,497 (4,524) (1,120)	19.84

* No. tested children for each parental classification.

† No. untested children for each parental classification.

$\hat{\theta}$, derived from a sample of sibships, deviates from some null value, θ_0, which is expected in the absence of any distorting effects. And further, it is of interest to use the set of estimates $\hat{\theta}_1, \hat{\theta}_2, \ldots \hat{\theta}_C$, estimated from C subsets of all families, to test whether subsets are derived from some common population of families with the same parameter θ, *not necessarily the null value* θ_0. The estimation is accomplished here by the standard maximum-likelihood procedures suggested by Fisher [3] utilizing iteration techniques employing the scoring method summarized by Bailey [4, pp. 276–282], with certain exceptions cited below. The testing of hypotheses is carried out in a different manner than has been previously suggested by Morton [5, 6] and discussed by Mi [7] and Elandt-Johnson [8]. Because we chose a somewhat different approach, we propose, for clarity, to present the argument for estimation and hypothesis testing in detail.

Estimation

We begin by defining the likelihood function, L_{sr}, for a given array of children in a family. In this study, the sample of families was derived in a random manner without consideration of the children's phenotypes. This sampling method has been called "ascertainment through the parents" (see [9] for discussion). For this case, the likelihood function (segregation distribution) which defines the probability of observing r segregants in a sibship of size s is simply the binomial

$$L_{sr} = \binom{s}{r} \theta^r (1 - \theta)^{s-r} ; \quad r = 0, 1, \ldots, s ; \quad 0 < \theta < 1 ; \tag{1}$$

where θ is the true proportion of segregants,

$$\sum_{r=0}^{s} L_{sr} = 1 ,$$

and the genotypes of the parents are known. For those matings which have parents whose genotypes are not defined by their phenotypes, the probability that a family cannot segregate because of undetectable homozygosity of one or both parents, designated by the parameter h, must be considered. When h is not zero, the segregation distribution becomes

$$L_{sr} = h + (1 - h)(1 - \theta)^s , \quad \text{when } r = 0 ; \quad 0 < \theta . \quad h < 1 , \tag{2a}$$

and

$$L_{sr} = (1 - h) \binom{s}{r} \theta^r (1 - \theta)^{s-r} , \quad \text{for cases } r > 0 ; \quad 0 < \theta , \quad h < 1 . \tag{2b}$$

Equations (2a) and (2b) are used to obtain simultaneous estimates of θ and h. Or, h may be considered to be a function of expected homozygosity when mating is at random and there is no selection, and may be computed from the gene frequencies of the parents. Both analyses were carried out. Because multiple alleles may be involved, a more general definition of h is the probability of either homozygosity or heterozygosity not detectable because of the nature of the mating (the $A_2 \times A_1$ mating is an example). The general derivation of h values in terms of gene frequencies is given by Morton [5] and presented in table 7 for backcrosses. In the backcross, we are concerned about the probability, say h_b, that the parental type which is expected to segregate cannot do so for the reasons mentioned above. For intercrosses, both parents are involved in producing segregants. The probability, say h_I, of no segregation in intercross progeny becomes $h_b^2 + 2h_b (1 - h_b)$.

The total number of families which have the same Mendelian expectation for a genetic system were partitioned into subsets classified according to the phenotype segregating, the mating type, the sex of the segregating phenotype (reciprocal for backcrosses only), and the size of the family. Each subset was described by a likelihood equation which was used to obtain maximum-likelihood estimates (by numerical differentiation) of the parameters of interest. Hence, a likelihood equation is a function of the number of families, a_{ijksr}, of size s,

TABLE 7

PROBABILITY (h_b) THAT DESIGNATED PARENTAL PHENOTYPE WILL NOT PRODUCE SEGREGANT PHENOTYPES IN BACKCROSS PROGENY FOR SYSTEMS STUDIED

Segregating Parent	h_b
A_1	$f_{A_1}/[f_{A_1}+2(f_{A_2}+f_0)]$
A_2	$f_{A_2}/(f_{A_2}+2f_0)$
A_2*	$1-[2f_0/(f_{A_2}+2f_0)][f_0/(f_{A_1}+f_0)]$
B	$f_B/(f_B+2f_0)$
Rh-D, Fy(a+), Jk(a+), Jk(b+), P(+), Secr., LeS	$f_+/(f_++2f_-)$

NOTE.—f=population gene frequency.

* Where A_2 is mated to A_1, and A_1 offspring are excluded from the analysis.

with r segregants ($r = 0, \ldots, s$; $s = 1, 2, \ldots, S$, in the kth reciprocal; $k = 1, K$ [$K = 2$ for backcrosses, $K = 1$ for intercrosses] of the jth mating type; $j = 1, \ldots, n$, which involves the ith segregating phenotype; $i = 1, \ldots, m$). We let θ_{ijks} and h_{ijks} (if applicable) be the parameters associated with the a_{ijks}. observation. We note that

$$\sum_i \sum_j \sum_k \sum_s \sum_r a_{ijksr} = a \ldots\ldots,$$

letting the dot subscript denote the sum over all i, j, k, s, or r.

Each subset of the total number of families, $a\ldots\ldots$, can be expressed, then, as a likelihood equation compounded in terms of the likelihood, L_{sr}, for a single family, given by either equation (1) or (2). It follows that the likelihood equation corresponding to each subset is used to obtain estimates of the parameters which will have the specifications which identify the subset. Let

$$L_{ijks} = \left(\frac{a_{ijks.}!}{\prod_{r=0}^{s} a_{ijksr}!} \right) \prod_{r=0}^{s} (L_{sr})^{a_{ijks}} \tag{3}$$

be the likelihood for a sample of $a_{ijks.}$ families of size s with the ijk classification, where L_{sr} is either equation (1) or (2) depending on the nature of the ith phenotype which is being studied. The construction of the likelihoods for the higher-order samples is simplified by using the fact that the values of the parameters which will maximize L, the maximum-likelihood (ML) estimates, also maximize log L. Therefore, letting L equal log L and ignoring constant terms, equation (3) becomes

$$L_{ijks} = \sum_{r=0}^{s} a_{ijksr} L_{sr} ; \tag{4}$$

for all families in ijk, sizes pooled,

$$L_{ijk} = \sum_{s=1}^{S} \sum_{r=0}^{s} a_{ijksr} L_{sr} ; \tag{5}$$

for all families in ijs, reciprocals pooled,

$$L_{ijs} = \sum_{k=1}^{K} \sum_{r=0}^{s} a_{ijksr} L_{sr} \tag{6}$$

for all families in ij, reciprocals and sizes pooled,

$$L_{ij} = \sum_{k=1}^{K} \sum_{s=1}^{S} \sum_{r=0}^{s} a_{ijksr} L_{sr} ; \tag{7}$$

for all families with the ith segregating phenotype, matings, reciprocals, and sizes pooled,

$$L_i = \sum_{j=1}^{n} \sum_{k=1}^{K} \sum_{s=1}^{S} \sum_{r=0}^{s} a_{ijksr} L_{sr} ; \tag{8}$$

and, finally, for all families with the same hypothesis, the pooled likelihood is

$$L = \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{k=1}^{K} \sum_{s=0}^{S} \sum_{r=0}^{s} a_{ijksr} L_{sr} . \tag{9}$$

The log likelihoods, equations (4) through (9), were analyzed for each system to obtain the set of ML estimates [when L_{sr} = eq. (1), and all families are characterized by the same segregation hypothesis] given in table 8.

The ML estimates were obtained for each subset of families by using the general maximum-likelihood program, MAXLIK [10], as a subroutine in a mainline program designed to construct the likelihoods and manipulate the corresponding subsets of data available for each genetic system.

The observed variance of the ML estimates derived from a specified subset of families was taken to be the inverse of the expected information computed from the data evaluated at the estimates of $\hat{\theta}$ and $\hat{h}$, if appropriate. Let DF $(\hat{\theta})_r$ and DF $(\hat{h})_r$ be the first derivatives of L_{sr} with respect to θ and h, respectively, computed by numerical differentiation and evaluated in the region of $(\hat{\theta}, \hat{h})$. Alternately stated, DF_r is the rate of change of the likelihood,

TABLE 8

SUMMARY OF LIKELIHOODS ASSOCIATED WITH THE GENETIC-RATIO ANALYSIS

Maximum-Likelihood Estimate	Log Likelihood	No. Estimates for Each System*
$\hat{\theta}$	L	1
$\hat{\theta}_i$	L_i	m
$\hat{\theta}_{ij}$	L_{ij}	$\sum_{i=1}^{m} \sum_{j=1}^{n} \delta_{ij}$
$\hat{\theta}_{ijk}$	L_{ijk}	$\sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{k=1}^{K} \delta_{ijk}$
$\hat{\theta}_{ijks}$	L_{ijks}	$\sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{k=1}^{K} \sum_{s=1}^{S} \delta_{ijks}$

* δ equals 1 if there are data for the subscripted combination, and equals 0 otherwise.

L_{sr}, in the region of the ML estimates. To illustrate the formulation, consider the likelihood L_{ijks}. The elements of the information matrix are

$$I_{\hat{\theta}_{ijks}} = a_{ijks.} \sum_{r=0}^{s} \frac{[\mathrm{DF}(\hat{\theta})_r]^2}{L_{sr}},$$

$$I_{\hat{h}_{ijks}} = a_{ijks.} \sum_{r=0}^{s} \frac{[\mathrm{DF}(\hat{h})_r]^2}{L_{sr}},$$

and

$$I_{\hat{\theta}_{ijks}\hat{h}_{ijks}} = a_{ijks.} \sum_{r=0}^{s} \frac{\mathrm{DF}(\hat{\theta})_r \mathrm{DF}(\hat{h})_r}{L_{sr}},$$

with L_{sr} evaluated at $(\hat{\theta}, \hat{h})$ in each case. Therefore, in the two-parameter case, the variance-covariance matrix becomes (for any likelihood)

$$I^{-1} = \begin{bmatrix} I_{\hat{\theta}} & I_{\hat{\theta},\hat{h}} \\ I_{\hat{\theta},\hat{h}} & I_{\hat{h}} \end{bmatrix}^{-1}.$$

When L_{sr} is equation (1), I^{-1} is simply $I_{\hat{\theta}}^{-1}$.

In the treatment developed by Morton [5, 6], the first derivatives (designated as the U scores) and the information (designated as K) for each family are taken to be the appropriate *expectations* evaluated at the null values, θ_0 and h_0. As a basis for comparison of the tests of significance obtained with the present method (presented below), which are based on I^{-1} evaluated at $\hat{\theta}$ (and $\hat{h}$ when appropriate), with tests based on Morton's U scores and K values, U and K were computed simultaneously for each subset of data. Presently, there is no mathematical basis for an argument that either approach is more correct than the other. The difference resides in the inferences one makes from the test results. For lack of space, these values will not be presented for each analysis, but the comparisons will be summarized in the discussion.

Hypothesis Testing

Hypothesis testing was of two sorts. The first class of tests was made to determine if estimates of parameters deviated significantly from expected values. In the two-parameter case,

$$\underset{(1 \times 2)}{X^2} = \underset{(1 \times 2)}{(\hat{\theta} - \theta_0, \hat{h} - h_0)} \underset{(2 \times 2)}{I} \underset{(2 \times 1)}{(\hat{\theta} - \theta_0, \hat{h} - h_0)'} \tag{10}$$

was taken to be distributed approximately as a χ^2 with 2 df. When $\hat{h}$ is taken to be the null value, h_0, computed from parental gene frequencies, or when h is zero because of the mating type, then

$$X^2 = \frac{(\hat{\theta} - \theta_0)^2}{I_{\hat{\theta}}^{-1}} \tag{11}$$

was taken as approximately χ^2 with 1 df.

The objective of a second class of tests was to detect heterogeneity among estimates of the parameters made from C subsets of families, classified according to the parental sex which was segregating, size of the family, and mating type. (We chose to ignore heterogeneity among families on the assumption that the genetic hypothesis was homogeneous for the

systems studied.) Essentially, this was done by a weighted χ^2 analysis of variance. For clarity, we present the univariate case to illustrate the logic of the test. First, let

$$\sum_{c=1}^{C} W_c \hat{\theta}_c = \hat{\bar{\theta}}$$

be the weighted average of the C estimates of the genetic ratio, where

$$W_c = \frac{I_{\hat{\theta}_c}}{I_{\hat{\theta}}},$$

$I_{\hat{\theta}_c}$ being the observed information for $\hat{\theta}_c$, and

$$I_{\hat{\theta}} = \sum_{c=1}^{C} I_{\hat{\theta}_c}.$$

Then

$$X^2 = \sum_{c=1}^{C} W_c[(\hat{\theta}_c - \hat{\theta})^2] I_{\hat{\theta}} \tag{12}$$

is distributed approximately as a χ^2 statistic with $C - 1$ df. Equation (12) may be rewritten in a more convenient form:

$$X^2 = \sum_{c=1}^{C} I_{\hat{\theta}_c} (\hat{\theta}_c - \hat{\bar{\theta}})^2 .$$

It follows that for the two-parameter case,

$$X^2 = \sum_{c=1}^{C} [I_{\hat{\theta}_c} (\hat{\theta}_c - \hat{\bar{\theta}})^2 + I_{\hat{h}_c} (\hat{h}_c - \hat{\bar{h}})^2 + 2I_{\hat{\theta}_c \hat{h}_c} (\hat{\theta}_c - \hat{\bar{\theta}})(\hat{h}_c - \hat{\bar{h}})] , \tag{13}$$

which is approximately the χ^2 statistic with $2C - 2$ df. Because these formulations for hypothesis testing differ from those of Morton, we computed routinely his analogy for testing θ [our eqs. (10) and (11)], namely,

$$X^2 = \frac{U^2}{K}, \tag{14}$$

and his tests of heterogeneity [our eqs. (12) and (13)], namely,

$$X^2 = \sum_{c=1}^{C} (U_c^2 / K_c) - \frac{\left(\sum_{c=1}^{C} U_c\right)^2}{\left(\sum_{c=1}^{C} K_c\right)}. \tag{15}$$

Because of limited space, we present only a summary of this comparison in the discussion.

RESULTS

The results of the four analyses reported in this paper are given in tables 9, 10, and 11. The analysis of mating-type frequencies to detect nonrandom (assortative) mating effects is presented in table 9. This analysis partitions the total χ^2 for each system into a component due to deviations of mating types (reciprocals pooled) from ex-

TABLE 9—ANALYSIS OF MATING TYPE FREQUENCIES AND DISTRIBUTION OF OFFSPRING AMONG MATING TYPES

Source of Variability	df	Mating Freq. (χ^2)	Offspring Freq. Given Mating-Type Freq. (χ^2)
ABO:			
Total	35	36.38	18.19
Among mating types	20	13.11	7.64
Among reciprocal mating types	15	23.26	10.54
MN:			
Total	8	8.01	5.07
Among mating types	5	1.77	2.73
Among reciprocal mating types	3	6.24	2.34
Ss:			
Total	8	8.66	6.23
Among mating types	5	4.29	5.29
Among reciprocal mating types	3	4.37	0.94
MNSs:			
Total	80	75.33	47.20
Among mating types	44	39.61	24.67
Among reciprocal mating types	36	35.72	22.53
Rh-C:			
Total	8	3.63	4.20
Among mating types	5	2.14	3.13
Among reciprocal mating types	3	1.49	1.07
Rh-D			
Total	3	3.31	0.35
Among mating types	2	0.97	0.28
Among reciprocal mating types	1	2.34	0.07
Rh-E:			
Total	8	7.53	4.87
Among mating types	5	3.56	3.39
Among reciprocal mating types	3	3.96	1.48
Kell-Kk:			
Total	8	3.67	2.77
Among mating types	5	3.59	1.11
Among reciprocal mating types	3	0.08	1.66
Kell-Kp^aKp^b:			
Total	8	2.66	0.82
Among mating types	5	1.73	0.80
Among reciprocal mating types	3	0.93	0.02
Duffy:			
Total	3	3.37	2.45
Among mating types	2	2.78	1.19
Among reciprocal mating types	1	0.59	1.26
Kidd-Jk^a:			
Total	3	4.80	0.24
Among mating types	2	0.42	0.13
Among reciprocal mating types	1	4.38*	0.11
Kidd-Jk^b:			
Total	3	2.55	1.57
Among mating types	2	2.55	1.11
Among reciprocal mating types	1	0.00	0.46
P:			
Total	3	3.88	0.43
Among mating types	2	3.81	0.42
Among reciprocal mating types	1	0.07	0.01
Lewis secretion:			
Total	3	3.21	1.13
Among mating types	2	0.39	1.10
Among reciprocal mating types	1	2.82	0.03
ABH secretion:			
Total	3	0.16	1.48
Among mating types	2	0.10	0.26
Among reciprocal mating types	1	0.06	1.22
Haptoglobin:			
Total	8	9.78	7.41
Among mating types	5	3.49	6.99
Among reciprocal mating types	3	6.29	0.42
Gc:			
Total	8	3.59	10.14
Among mating types	5	2.11	3.97
Among reciprocal mating types	3	1.48	6.17

* Significant at .05 level of probability.

TABLE 10

ANALYSIS OF PARENTAL AND OFFSPRING PHENOTYPE AND GENE FREQUENCIES

Phenotype	Parents' Phenotypes			Children's Phenotypes					Gene	Parents' Gene Frequency	Children's Gene Frequency
	Observed	Expected	χ^2	Observed	Expectation 1 from Children's Gene Frequency	χ^2	Expectation 2 from Parental Gene Frequency	χ^2			
ABO:											
A_1	1,152	1,147.21	.....	1,116	1,114.59	.....	1,085.49	.....	A_1	.1985	.2039
A_2	329	334.36	.....	331	333.43	.....	316.33	.....	A_2	.0689	.0732
A_1B	88	92.80	.....	86	87.48	.....	86.79	.....	B	.0674	.0653
A_2B	38	32.10	.....	34	31.40	.....	30.47	.....	O	.6652	.6574
B	326	326.73	.....	295	295.95	.....	309.11	.....	.....	.....	.....
O	1,535	1,534.56	.....	1,419	1,418.16	.....	1,451.81	.....	.....	.....	.....
	(3,468)	.....	1.44 (2)	(3,281)	.....	0.26 (2)	.....	3.37 (5)	.....	.....	.....
MN:											
MM	1,018	1,025.68	.....	1,013	1,021.95	.....	969.54	.....	M	.5436	.5581
MN	1,738	1,722.32	.....	1,636	1,618.35	.....	1,628.03	.....	N	.4564	.4419
NN	715	723.00	.....	632	640.70	.....	683.44	.....	.....	.....	.....
	(3,471)	.....	0.29 (1)	(3,281)	.....	0.39 (1)	.....	5.86 (2)	.....	.....	.....
Ss:											
SS	330	327.67	.....	325	322.23	.....	312.74	.....	S	.3209	.3258
Ss	1,382	1,386.86	.....	1,329	1,334.45	.....	1,323.66	.....	s	.6791	.6742
ss	1,470	1,467.47	.....	1,383	1,380.32	.....	1,400.59	.....	.....	.....	.....
	(3,182)	.....	0.04 (1)	(3,037)	.....	0.05 (1)	.....	0.72 (2)	.....	.....	.....

TABLE 10—*Continued*

Phenotype	Parents' Phenotypes			Children's Phenotypes					Gene	Parents' Gene Frequency	Children's Gene Frequency
	Observed	Expected	χ^2	Observed	Expectation 1 from Children's Gene Frequency	χ^2	Expectation 2 from Parental Gene Frequency	χ^2			
MNSs:											
MSMS	213	207.59	.....	216	211.95	.....	198.10	.....	*MS*	.2554	.2641
MSMs	462	468.79	.....	469	470.30	.....	447.39	.....	*Ms*	.2884	.2930
MSNS	106	106.35	.....	100	98.87	.....	101.45	.....	*NS*	.0654	.0616
MNSs	751	755.24	.....	712	721.23	.....	720.66	.....	*Ns*	.3907	.3811
MsMs	263	264.67	.....	246	260.89	.....	252.60	.....	.....	.....	.....
MsNs	728	717.19	.....	711	678.48	.....	684.41	.....	.....	.....	.....
NSNS	11	13.62	.....	9	11.53	.....	12.99	.....	.....	.....	.....
NSNs	169	162.70	.....	148	142.64	.....	155.20	.....	.....	.....	.....
NsNs	479	485.86	.....	426	441.12	.....	462.59	.....	.....	.....	.....
	(3,182)	.....	1.28 (5)	(3,037)	.....	3.90 (5)	.....	10.38 (8)	.....	.....	.....
Rh:											
CC	555	565.90	.....	508	518.77	.....	496.44	.....	*C*	.4096	.4185
Cc	1,653	1,631.37	.....	1,463	1,441.65	.....	1,431.14	.....	*c*	.5904	.5815
cc	1,165	1,175.73	.....	991	1,001.58	.....	1,031.43	.....	.....	.....	.....
	(3,373)	.....	0.59 (1)	(2,962)	.....	0.65 (1)	.....	2.81 (2)	.....	.....	.....
D	2,804	.....	.....	2,486	2,485.95	.....	2,462.39	.....	*D*	.5893	.5991
dd	569	.....	.....	476	476.05	.....	499.61	.....	*d*	.4107	.4008
	(3,373)	.....	.....	(2,962)	.....	.....	.....	1.34 (1)	.....	.....	.....
EE	87	82.30	.....	50	63.74	.....	72.36	.....	*E*	.1563	.1467
Ee	880	889.70	.....	769	741.57	.....	781.20	.....	*e*	.8437	.8533
ee	2,406	2,401.00	.....	2,143	2,156.69	.....	2,108.44	.....	.....	.....	.....
	(3,373)	.....	0.38 (1)	(2,962)	.....	4.05* (1)	.....	7.67* (2)	.....	.....	.....

* Significant at .05 level of probability.

TABLE 10—*Continued*

Phenotype	Parents' Phenotypes			Children's Phenotypes					Gene	Parents' Gene Frequency	Children's Gene Frequency
	Observed	Expected	χ^2	Observed	Expectation 1 from Children's Gene Frequency	χ^2	Expectation 2 from Parental Gene Frequency	χ^2			
Kell-Kk:											
KK	2	4.20	.....	2	3.69	.....	3.94	.....	*K*	.0348	.0337
Kk	238	233.78	.....	215	211.80	.....	218.44	.....	*k*	.9651	.9663
kk	3,230	3,232.02	.....	3,035	3,036.51	.....	3,028.97	.....	.....	.....	.....
	(3,470)	.....	1.14 (1)	(3,252)	.....	0.82 (1)	.....	1.02 (2)	.....	.....	.....
Kell-Kp^{a}Kp^b:											
Kp(a+b−)	1	0.22	.....	0	0.13	.....	0.21	.....	*Kp^a*	.0082	.0065
Kp(a+b+)	51	52.26	.....	40	39.45	.....	49.67	.....	*Kp^b*	.9918	.9935
Kp(a−b+)	3,161	3,160.52	.....	3,014	3,014.42	.....	3,003.51	.....	.....	.....	.....
	(3,213)	.....	0.64 (1)	(3,054)	.....	0.13 (1)	.....	2.13 (2)	.....	.....	.....
Duffy:											
Fy(a+)	2,338	.....	.....	2,236	.....	.....	2,165.23	.....	*Fy^a*	.4198	.4388
Fy(a−)	1,186	.....	.....	1,028	.....	.....	1,098.77	.....	*Fy^b*	.5802	.5612
	(3,524)	.....	.....	(3,264)	.....	.....	.....	6.87** (1)	.....	.....	.....
Kidd-Jk^a:											
Jk(a+)	1,761	.....	.....	1,656	.....	.....	1,677.74	.....	*Jk^a*	.5142	.5041
Jk(a−)	544	.....	.....	540	.....	.....	518.26	.....	*Jk^b*	.4858	.4959
	(2,305)	.....	.....	(2,196)	.....	.....	.....	1.19 (1)	.....	.....	.....
Kidd-Jk^b:											
Jk(b+)	441	.....	.....	422	.....	.....	395.64	.....	*Jk^b*	.4816	.5309
Jk(b−)	162	.....	.....	119	.....	.....	145.33	.....	*Jk^a*	.5184	.4691
	(603)	.....	.....	(541)	.....	.....	.....	6.52* (1)	.....	.....	.....

* Significant at .05 level of probability.

** Significant at .01 level of probability.

TABLE 10—*Continued*

Phenotype	Parents' Phenotypes			Children's Phenotypes					Gene	Parents' Gene Frequency	Children's Gene Frequency
	Observed	Expected	χ^2	Observed	Expectation 1 from Children's Gene Frequency	χ^2	Expectation 2 from Parental Gene Frequency	χ^2			
P:											
P(+)	2,710			2,528			2,552.20		P^1	.5324	.5246
P(−)	758			738			713.80		P^2	.4675	.4754
	(3,468)			(3,266)				1.05 (1)			
Lewis secretion:											
LeS	2,919			2,503			2,528.47		*Le*	.7570	.7384
nL	183			184			158.53		*le*	.2429	.2616
	(3,102)			(2,687)				4.35* (1)			
ABH secretion-Se:											
Secr	2,619			2,261			2,316.77		*Se*	.5101	.4919
Nonsecr	827			787			731.23		*se*	.4898	.5081
	(3,446)			(3,048)				5.60* (1)			
Haptoglobin:											
Hp 1-1	564	548.15		424	443.17		409.74	0.50	Hp^1	.4125	.4290
Hp 1-2	1,531	1,562.41		1,218	1,179.72		1,167.12	2.22	Hp^2	.5875	.5710
Hp 2-2	1,128	1,112.44		766	785.11		831.14	5.11			
	(3,223)		1.31 (1)	(2,408)		2.54 (1)		7.82* (2)			
Gc:											
Gc 1-1	1,674	1,678.22		1,219	1,220.40		1,266.43		Gc^1	.7197	.7065
Gc 2-1	1,316	1,307.22		1,017	1,013.98		986.47		Gc^2	.2803	.2935
Gc 2-2	250	254.56		209	210.62		192.10				
	(3,240)		0.14 (1)	(2,445)		0.01 (1)		4.21 (2)			

* Significant at .05 level of probability.

TABLE 11—RESULTS OF ANALYSIS OF GENETIC RATIOS

Mating Type	No. Fam.	No. Child.	Nonsegregants $1-\theta$	Segregants θ	All Sizes and Reciprocals Pooled				No. Families Required	Heterogeneity among			All Sizes and Reciprocals Pooled				
					θ_0	$\hat{\theta}$	$I_{\hat{\theta}}^{-1}\times 10^{+3}$	χ^2		Reciprocals (χ^2 1 df)	Sizes (χ^2 df)	Mating Types (χ^2 df)	h_0	$\hat{h}$	θ_0	$\ddot{\theta}$	χ^2 2 df
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
ABO:																	
Backcrosses																	
$A_1\times A_1B$	19	26	A_1B	A_2B, B	.50	.3889	11.56	1.07	69	2.22	1.96(2)						
$A_1\times A_2$	87	205	A_1	A_2, OO	.50	.4638	1.65	0.80	421	0.16	4.78(4)		.1190	.0650	.50	.4433	1.27
$A_1\times A_2B$	9	22	A_1, A_1B	A_2, A_2B, B	.50	.5147	15.76	0.01	2,521	0.24	0.01(1)		.1190	.1452	.50	.5253	0.03
$A_1\times B$	98	233	A_1, A_1B	A_2B, OO, B	.50	.4730	1.43	0.51	739	0	4.66(6)		.1190	.1197	.50	.4733	0.51
$A_1\times OO$	397	931	A_1	A_2, OO	.50	.4852	0.36	0.60	2,506	0.78	5.28(7)		.1190	.1378	.50	.4924	0.85
$A_1\times$Pool	610	1,417			.50	.4785	0.24	1.93	1,217	1.13	2.43(7)	1.05(4)	.1190	.1241	.50	.4806	1.93
$A_1B\times A_1$	23	54	A_1	A_1B, A_2B, B	.50	.4814	4.62	0.07	1,180	0.04	0.68(4)						
$A_1B\times A_2$	11	28	A_1	A_2B, B	.50	.4642	8.88	0.14	293	1.09	0.57(3)						
$A_1B\times B$	7	19	A_1, A_1B	B	.50	.4736	13.12	0.05	507	0.59	0.50(2)						
$A_1B\times OO$	35	71	A_1	B	.50	.3943	3.36	3.32	41	0.08	2.15(4)						
$A_1B\times$Pool	76	172			.50	.4418	1.43	2.36	124	0.47	2.82(5)	1.14(3)					
$A_2\times A_1$	57	85	A_2	OO	.50	.5292	5.50	0.16	1,412	3.33	3.61(2)						
$A_2\times A_1B$	8	13	A_2B	B	.50	.6210	21.08	0.69	45	2.25	4.25(2)						
$A_2\times A_2B$	2	3	A_2B	B	.50	.6667	88.93	0.31	25		1.28(1)						
$A_2\times B$	29	68	A_2B, A_2	B, O	.50	.4703	4.19	0.21	529	0.29	6.80(4)						
$A_2\times OO$	121	264	A_2	OO	.50	.4687	1.09	0.90	517	0.26	6.69(5)		.0492	.0378	.50	.4641	0.94
$A_2\times$Pool	160	348	A_2		.50	.5346	1.30	0.93	668	0	5.49(6)	1.90(5)	.0492	.0233	.50	.4564	1.94
$A_2B\times A_1$	9	22	A_1, A_2	A_1B, A_2B, B	.50	.5908	10.99	0.75	47	1.57	1.35(2)						
$A_2B\times A_2$	3	6	A_2	A_2B, B	.50	.4999	41.67	0			0.14(1)						
$A_2B\times OO$	19	40	A_2	B	.50	.4249	6.11	0.92	80	0	2.27(3)						
$A_2B\times$Pool	32	69	A_2	B	.50	.4781	3.62	0.13	928	0.31	2.78(4)	1.62(2)					
$B\times A_1$	98	233	A_1B, A_2B, B	A_1, A_2, OO	.50	.5612	1.22	3.07	123	0.90	2.58(6)		.0482	.0131	.50	.5473	4.39
$B\times A_1B$	5	10	A_1B	A_1	.50	.5145	28.94	0.01	2,643		0.05(1)		.0482	.3331	.50	.5999	0.82
$B\times A_2$	29	68	A_2B, B	A_2, OO	.50	.5669	4.19	1.07	105	0.03	2.76(4)		.0482	.0315	.50	.5597	1.15
$B\times OO$	118	261	B	OO	.50	.4769	1.09	0.49	926	0.81	2.70(6)						
$B\times$Pool	251	573	B		.50	.5217	0.50	0.94	1,024	0.96	3.03(7)	3.60(3)	.0482	.0106	.50	.5072	2.94
Intercrosses:																	
$A_1\times A_1$	169	349	A_1	A_2, OO	.25	.2278	0.75	0.66	988		1.63(5)						
$A_1B\times A_1B$	2	4	A_1, A_1B	B	.25	.2499	46.87	0									
$A_1B\times A_1B$	2	4	A_1B	A_1, B	.50	.7499	46.88	1.33	6								
$A_1B\times A_1B$	2	3	A_1	B	.50	.3333	74.07	0.38	21								
$A_2\times A_2$	11	26	A_2	OO	.25	.2430	7.84	0.01	6,759		0.26(2)		.0492	.3512	.25	.3396	0.92
$B\times B$	11	25	B	OO	.25	.2551	8.31	0			0.24(1)		.0482	.3906	.25	.4258	1.51

TABLE 11—*Continued*

Mating Type	No. Fam.	No. Child.	Nonsegregants $1-\theta$	Segregants θ	All Sizes and Reciprocals Pooled				No. Families Required	Heterogeneity among			All Sizes and Reciprocals Pooled				
					θ_0	$\hat{\theta}$	$I_{\hat{\theta}}^{-1}\times10^{+3}$	χ^2		Reciprocals (χ^2 1 df)	Sizes (χ^2 df)	Mating Types (χ^2 df)	h_0	$\hat{h}$	θ_0	$\hat{\theta}$	χ^2 2 df
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
MN:																	
Backcrosses:																	
MN×M	428	980	M	MN	.50	.4591	0.25	6.69**	246	0.08	9.58(7)						
MN×N	282	634	MN	N	.50	.4826	0.39	0.77	1,395	1.24	8.79(7)						
MN×Pool	710	1,614	*M*	*N*	.50	.4683	0.15	6.70**	407	0.20	11.77(8)	0.85(1)					
Intercrosses:																	
MN×MN	366	847	M, MN	N	.25	.2502	0.22	0			4.27(8)						
MN×MN	366	847	MN	M, N	.50	.4923	0.30	0.20	7,112		3.80(8)						
MN×MN	263	417	M	N	.50	.5083	0.60	0.11	8,797		2.72(4)						
Ss:																	
Backcrosses:																	
Ss×S	116	285	S	Ss	.50	.5086	0.88	0.09	5,300	0.09	7.51(5)						
Ss×s	550	1,270	Ss	s	.50	.5007	0.20	0		0.01	7.08(7)						
Ss×Pool	666	1,555	*S*	*s*	.50	.5022	0.16	0.03		0.05	11.46(8)	0.06(1)					
Intercrosses:																	
Ss×Ss	233	533	S, Ss	s	.25	.2307	0.33	1.12	793		7.05(7)						
Ss×Ss	233	533	Ss	S, s	.50	.4802	0.47	0.84	1,073		5.10(7)						
Ss×Ss	161	255	S	s	.50	.4804	0.98	0.39	1,578		1.53(3)						
Rh-C:																	
Backcrosses:																	
Cc×C	199	475	C	Cc	.50	.5241	0.53	1.11	698	0.85	3.24(6)						
Cc×c	452	1,009	Cc	c	.50	.4796	0.25	1.68	1,043	0.75	7.34(7)						
Cc×Pool	651	1,484	*C*	*c*	.50	.4938	0.17	0.22		0.04	7.18(7)	2.56(1)					
Intercrosses:																	
Cc×Cc	303	697	C, Cc	c	.25	.2309	0.25	1.43	798		5.40(6)						
Cc×Cc	303	697	Cc	C, c	.50	.4920	0.36	0.18	6,545		7.77(7)						
Cc×Cc	224	443	C	c	.50	.4693	0.73	1.30	667		5.04(3)						
Rh-D:																	
Backcross:																	
D+×d	380	835	Dd	d	.50	.4874	0.72	0.22	6,618	0.38	24.22(6)***		.4177	.4157	.50	.4866	0.22
Intercross:																	
D+×D+	893	2,040	D	d	.25	.2193	0.35	2.71	1,274		7.12(6)		.6609	.6735	.25	.2261	2.67

** Significant at .01 level of probability.

*** Significant at .001 level of probability.

TABLE 11—*Continued*

Mating Type	No. Fam.	No. Child.	Nonsegregants $1-\theta$	Segregants θ	All Sizes and Reciprocals Pooled				No. Families Required	Heterogeneity among			All Sizes and Reciprocals Pooled				
					θ_0	$\hat{\theta}$	$I_{\hat{\theta}}^{-1}\times10^{+3}$	χ^2		Reciprocals (χ^2 1 df)	Sizes (χ^2 df)	Mating Types (χ^2 df)	h_0	$\hat{h}$	θ_0	$\hat{\theta}$	χ^2 2 df
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Rh-E:																	
Backcrosses:																	
Ee×E	13	30	E	Ee	.50	.4999	8.33	0		1.25	0						
Ee×e	479	1,038	Ee	e	.50	.5027	0.24	0.03		1.24	3.49(6)						
Ee×Pool	492	1,068	*E*	*e*	.50	.5027	0.23	0.03		1.64	3.43(6)	0					
Intercrosses:																	
Ee×Ee	90	200	E, Ee	e	.25	.2249	0.87	0.72	478		6.40(4)						
Ee×Ee	90	200	Ee	E, e	.50	.3999	1.20	8.35**	42		1.61(5)						
Ee×Ee	58	80	E	e	.50	.5624	3.08	1.27	177		3.68(2)						
Kell:																	
Backcrosses:																	
Kk×k	181	428	Kk	k	.50	.5185	0.58	0.59	1,178	1.24	10.45(8)						
Kp(a+b+)× Kp(a−b+)	39	100	Kp(a+b+)	Kp(a−b+)	.50	.6299	2.33	7.24**	21	0.18	9.25(5)						
Intercrosses:																	
Kk×Kk	5	14	K, Kk	k	.25	.2142	12.03	0.11	181		0.17(1)						
Kk×Kk	5	14	Kk	K, k	.50	.3571	16.40	1.25	16		0.17(1)						
Kk×Kk	3	5	K	k	.50	.5999	48.00	0.21	56								
Duffy:																	
Backcross:																	
Fy(a+)×Fy(a−)	676	1,536	Fy(a+)	Fy(a−)	.50	.4652	0.29	4.19*	622	1.58	11.00(8)		.2656	.2658	.50	.4654	4.15
Intercross:																	
Fy(a+)×Fy(a+)	614	1,415	Fy(a+)	Fy(a−)	.25	.2580	0.34	0.19			3.86(6)		.4607	.4449	.25	.2521	0.27
Kidd-Jk[a]:																	
Backcross																	
Jk(a+)×Jk(a−)	344	806	Jk(a+)	Jk(a−)	.50	.4991	0.65	0		3.34	8.58(7)		.3460	.3973	.50	.5186	1.72
Intercross																	
Jk(a+)×Jk(a+)	540	1,262	Jk(a+)	Jk(a−)	.25	.2821	0.54	1.91	1,087		6.15(6)		.5722	.6178	.25	.3060	2.43
Kidd-Jk[b]:																	
Backcross:																	
Jk(b+)×Jk(b−)	103	211	Jk(b+)	Jk(b−)	.50	.4464	2.32	1.24	320	1.37	2.77(4)		.3172	.3106	.50	.4436	1.26
Intercross:																	
Jk(b+)×Jk(b+)	131	308	Jk(b+)	Jk(b−)	.25	.2138	1.56	0.84	599		4.44(4)		.5337	.5048	.25	.2037	0.95

* Significant at .05 level of probability.

** Significant at .01 level of probability.

TABLE 11—*Continued*

Mating Type	No. Fam.	No. Child.	Nonsegregants $1-\theta$	Segregants θ	All Sizes and Reciprocals Pooled				No. Families Required	Heterogeneity among			All Sizes and Reciprocals Pooled				
					θ_0	$\hat{\theta}$	$I_{\hat{\theta}}^{-1}\times10^{+3}$	χ^2		Reciprocals (χ^2 1 df)	Sizes (χ^2 df)	Mating Types (χ^2 df)	h_0	$\hat{h}$	θ_0	$\hat{\theta}$	χ^2 2 df
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
P:																	
Backcross:																	
P(+)×P(−)	511	1,144	P(+)	P(−)	.50	.5502	0.47	5.36*	366	0.73	2.50(6)		.3628	.3978	.50	.5641	6.62*
Intercross:																	
P(+)×P(+)	869	1,998	P(+)	P(−)	.25	.2872	0.36	3.84*	869		9.26(7)		.5939	.5543	.25	.2676	5.05
Lewis secretion:																	
Backcross:																	
LeS×nL	135	298	LeS	nL	.50	.4248	3.15	1.79	289	5.80*	9.87(4)*		.6091	.5126	.50	.3804	3.56
Intercross:																	
LeS×LeS	1,026	2,378	LeS	nL	.25	.2630	0.83	0.21			7.38(5)		.8471	.6854	.25	.1560	8.35*
ABH secretion:																	
Backcross:																	
Secr.×Nonsecr.	491	112	Secr.	Nonsecr.	.50	.5082	0.47	0.15		1.56	3.17(6)		.3424	.3163	.50	.4988	0.73
Intercross:																	
Secr.×Secr.	786	1,775	Secr.	Nonsecr.	.25	.3035	0.40	7.19**	422		18.79(6)**		.5675	.4213	.25	.2424	14.98***
Haptoglobin:																	
Backcrosses:																	
Hp2-1×Hp1-1	199	431	Hp1-1	Hp2-1	.50	.5196	0.58	0.67	1,154	0.35	4.48(6)						
Hp2-1×Hp2-2	373	749	Hp2-1	Hp2-2	.50	.4859	0.33	0.60	2,378	0	6.43(5)						
Hp2-1×Pool	572	1,180	Hp^1	Hp^2	.50	.4982	0.21	0.02		0.17	7.87(6)	1.25(1)					
Intercrosses:																	
Hp2-1×Hp2-1	248	518	Hp1-1, Hp2-1	Hp2-2	.25	.2239	0.33	2.06	462		6.62(5)						
Hp2-1×Hp2-1	248	518	Hp2-1	Hp1-1, Hp2-2	.50	.4806	0.48	0.78	1,215		6.90(5)						
Hp2-1×Hp2-1	174	249	Hp1-1	Hp2-2	.50	.4658	1.00	1.17	572		0.46(2)						
Gc:																	
Backcrosses:																	
Gc2-1×Gc1-1	484	1,055	Gc1-1	Gc1-2	.50	.5003	0.24	0		0.21	6.71(7)						
Gc2-1×Gc2-2	82	183	Gc1-2	Gc2-2	.50	.4863	1.37	0.14	2,299	0.47	7.75(5)						
Gc2-1×Pool	566	1,238	Gc^1	Gc^2	.50	.4983	0.20	0.01		0.03	6.28(7)	0.12(1)					
Intercrosses:																	
Gc2-1×Gc2-1	186	393	Gc1-1, Gc2-1	Gc2-2	.25	.2442	0.47	0.07	9,979		2.73(5)						
Cc2-1×Gc2-1	186	393	Gc2-1	Gc1-1, Gc2-2	.50	.4731	0.63	1.13	622		2.26(5)						
Gc2-1×Gc2-1	121	186	Gc1-1	Gc2-2	.50	.5161	1.34	0.19	2,403		0.16(3)						

* Significant at .05 level of probability.

** Significant at .01 level of probability.

*** Significant at .001 level of probability.

pected Hardy-Weinberg frequencies and a component due to differences between reciprocals within mating types. The first measures assortative mating, ignoring the sex of the phenotype, and the latter quantifies the sex effect on mating-type frequency.

The analysis of the number of children typed in this study by mating type of parents is also given in table 9. Because the expected number of offspring is taken to be that proportion of the total children determined by the observed mating-type frequency, the χ^2 measures the disproportionate contribution of particular mating types free of the effect of assortative mating. Only children reporting to the clinic are included; hence, a proper fertility analysis by mating type cannot be made, and comparison with the earlier study of Reed et al. [11] is not possible. Indicators of fertility, that is, abortions, stillbirths, etc., are available for each mating type and will be the subject of a publication in preparation. As with mating-type frequencies, the total χ^2 has been partitioned to quantitate the differences among mating types (reciprocals pooled) and between sexes for a phenotype within mating types.

Table 10 summarizes the analysis of parental and offspring phenotype distributions. The goodness-of-fit of observed frequencies to those expected, based on Hardy-Weinberg assumptions, is measured by χ^2. Several types of information are presented. Directional deviations due to selection, nonrandom mating, and/or the dispersive effects of sampling on phenotype frequencies in the parental and offspring generations all contribute to the χ^2 when expected frequencies are computed from the gene frequencies of the respective generations. On the other hand, any net effect of selection (change in gene frequency) may be approximated by comparing the children's χ^2 based on children's gene frequencies (table 10, expectation 1), with the χ^2 based on expectations computed using parental gene frequencies (table 10, expectation 2). This assumes that the χ^2 based on offspring gene frequencies serves as a base line, or control, for the latter χ^2. The difference between the two χ^2s is taken as a rough measure of the contribution of selection to the failure of the offspring phenotype frequencies to be predicted by parental gene frequencies and Hardy-Weinberg assumptions. For comparisons, the gene frequencies for parents and children are also given for each system in table 10.

Finally, table 11 presents the results of the analysis of the genetic ratios estimated from families. Each mating type is described with respect to phenotype (col. 1), number of families (col. 2), number of children (col. 3), and the offspring phenotypes classified as segregants or nonsegregants (cols. 4 and 5). The expected proportion of segregants is given in column 6. Columns 7–13 give an analysis of θ in which h is assigned a value of zero (no dominance) or is computed from parental gene frequencies. Columns 14–18 present the results of a simultaneous analysis of θ and h in those cases where h is not equal to zero. The estimate, $\hat{\theta}$ (col. 7), and its variance, $I_{\hat{\theta}}^{-1}$ (col. 8), is obtained from analysis of the pooled likelihood given by equation (7). The χ^2 test (eq. [11]) of the deviation of $\hat{\theta}$ from the null value, θ_0, is given in column 9. To indicate the problem of sample size, the number of families which would be necessary to reject at the .05 level the hypothesis of no selection (assuming the observed information per family as constant) was computed and is presented in column 10 of table 11. The χ^2 values for tests of the difference between estimates of θ from reciprocals (sizes pooled) (col. 11) and heterogeneity of estimates from different-size families

(reciprocals pooled) (col. 12) for each mating type are computed according to equation (12). In column 13 the heterogeneity of $\hat{\theta}$ among mating types with the same segregating phenotype (reciprocals and sizes pooled for each mating type) is also measured by the χ^2 [eq. (12)]. The degrees of freedom for each χ^2 are given in parentheses for those cases where they exceed 1.

For the analysis when $h \neq 0$, θ and h are estimated simultaneously from data of a mating type when reciprocals and sizes are pooled. The χ^2 tests [eq. (13)] for goodness-of-fit of the two-valued vector $(\hat{\theta}, \hat{h})$ to the corresponding null vector (θ_0, h_0) are presented in column 18.

The results of tests of significance from the analyses presented in tables 9, 10, and 11 indicate that, of 397 tests, 23 (5.79%) were significant at the 5% level of probability, and 9 (2.27%) were significant at the 1% level of probability. Hence, slightly more tests were significant than would be expected if no true deviation from the null hypothesis were present and the observed significant χ^2 values were attributable to chance alone. *Therefore, although some fraction of this number of tests yielding apparently significant results is surely due to chance, there seems to be no recourse but to scrutinize each case, searching for consistencies within a system which would give some confidence in our interpretation of a test.*

ABO. The analyses of mating-type frequencies and the number of tested offspring by mating type of parents show no deviation from expectation (table 9). Phenotype frequencies of parents and children indicate no significant deviations from proportions expected under Hardy-Weinberg equilibrium (table 10). There were no significant deviations from the expected genetic ratios among offspring of any ABO mating (table 11).

The significant A_1 incompatibility effect reported by Morton et al. [12], Chung et al. [13], and Matsunaga and Itoh [14] and B incompatibility reported by Peritz [15] were not observed. In fact, there was a deficiency of O offspring from incompatible matings with an O mother. The proportion of incompatible children in the pooled matings of O mothers by A_1, A_2, or B fathers, was .48, which is not significantly different from one-half and which was precisely equivalent to the estimate from the pooled reciprocals, that is, O fathers by A_1, A_2, or B mothers. There were deficiencies of B offspring in A_1B male by O female matings ($\hat{\theta} = .38$) and in A_2B male by O female matings ($\hat{\theta} = .42$) which were, however, in these small samples (20 and nine families, respectively) not significant in either case.

MNSs. The mating-type frequencies and number of children tested by mating type of parents show no significant deviation from expectation (table 9). Likewise, phenotype frequencies of parents and children (table 10) do not deviate from Hardy-Weinberg expectations. As noted earlier, the difference between the χ^2 based on parental gene frequencies, 5.87 (2 df), and that based on offspring gene frequencies, .39 (1 df), can be used as a rough measure of selection. The difference of 5.48 (1 df, $.025 < P < .01$) is due to an increased M frequency in children. With respect to this MN effect, in the pooled backcross data (table 11) we observe a 3.17% excess of M gametes ($\chi^2 = 6.70$, $.005 < P < .01$). The departure from expectation is in the same direction for both backcrosses but is significant only in the MN $\times$ M mating. However, the intercross data reveal no corresponding excess of the M gamete. The de-

parture in the backcrosses is apparently responsible for the increase in M gene frequency in the children's generation (table 10). Letting

$$(1 - \hat{\theta}) = \frac{1}{(2 - s)} \quad \text{and} \quad \hat{\theta} = \frac{(1 - s)}{(2 - s)} \tag{16}$$

represent the observed frequencies of M and N, respectively, from the pooled backcrosses, the selection coefficient, s, against N is .12. We note, however, that in these data there was no significant age trend with respect to the MNSs system [1]. In fact, there was an apparent increase in M in the oldest decades; hence one would have to postulate a reversal in the direction of selection in three generations if this result were regarded as a valid effect.

There were no significant deviations from expectation in any analysis of Ss. Analysis of the above M effect in the MN $\times$ M matings according to the Ss type of offspring indicated this effect is independent of Ss type: children S, $\hat{\theta} = .4073$ ($\chi^2 = 0.96$); children Ss, $\hat{\theta} = .4434$ ($\chi^2 = 1.49$); and children s, $\hat{\theta} = .4017$ ($\chi^2 = 4.50$). Although the estimates deviate greatly from $\theta_0 = .5$ in all cases, the number of families with only SS or only Ss children, nine and 50, respectively, was not sufficient to reject the null hypothesis.

This M effect is different from the much-discussed excess of heterozygotes in MN backcrosses [16, 17]. Furthermore, the results are not compatible with an effect attributable to technical errors [18], since the observed increase in M in children obtains in backcross but not intercross matings. It is more likely that if this effect is present in other studies it has been overlooked, for this may be the first study to include the necessary number of matings (710) and children (1,614) to obtain a significant result. The excess of the M gamete, regardless of the genotype of the child (MN or MM), supports the suggestion of Hiraizumi [19] that a heterotic model may be an inappropriate biological interpretation.

Rh. All matings involving C^w and D^u were excluded. We note that the E-e phenotype frequencies of the children depart from Hardy-Weinberg equilibrium at the .05 level for both comparisons (table 10). There is an excess of heterozygotes, significant at the .01 level of probability, in the Ee $\times$ Ee intercross (table 11). This effect is in the same direction as the departure of children's phenotype frequencies from Hardy-Weinberg equilibrium expectations based on children's gene frequencies; hence the two results tend to reinforce one another. On the other hand, the absence of any suggestion of a deviation from expected genetic ratios in the E-e backcrosses certainly casts suspicion on any real biologic effect. We note that the selection coefficient against homozygotes from the intercross is $s = .33$.

Table 11 also reveals, for the first time in this analysis, a "family-size effect," that is, there is heterogeneity among genetic ratios estimated from families of differing size in the D+ $\times$ dd mating. Examination of the Rh-D analysis in depth reveals neither a difference between results from D+ father by dd mother as compared to the dd father by D+ mother or the parity effect expected with Rh-isoimmunization in the matings with dd mothers. Thus the heterogeneity does not fit the established facts concerning Rh-isoimmunization and defies a consistent interpretation.

Kell. For purposes of this analysis, the K-k and Kp^a-Kp^b alternatives of the Kell

system are considered separately in the same fashion as for the component sets of factors in the MNSs and Rh systems. This strategy of analysis seemed justified because the number of informative matings defined by any two sets of factors was small in most instances. For Kell, the K^b and k^a allele frequencies were near zero; hence few informative matings for each set of factors were available. For neither set of factors is there a significant deviation from Hardy-Weinberg expectation of mating-type frequencies (table 9) or of phenotype frequencies of parents or children (table 10), nor did number of children tested per mating type deviate significantly from expected (table 9).

A significant deficiency (.01 level) of Kp(a+b+) offspring was observed in the progeny of Kp(a+b+) by Kp(a−b+) marriages (table 11). Taken at face value, this distortion requires a selection against the Kp^a gene, in heterozygotes, of .41. Selection coefficients of this magnitude, even if directed toward coadapted complexes of which Kp^a is a part, seem unlikely and suggest sampling error (39 families) may be contributing disproportionately to this observation. The finding of Reed et al. [11] that K+ women had more pregnancies is not confirmed by any distortion of the genetic ratios.

Duffy. Parental mating frequencies and representation of children among mating types are as expected (table 9). No simple test of goodness-of-fit based on phenotype tests exists when gene frequencies are estimated from the two phenotypic classes available in single-factor systems with dominance (table 10). However, deviations of children's phenotypes from expectations based on parental gene frequencies are appropriately tested with a χ^2 with 1 df. There is a significant deviation (.01 level) which is due in part to the increase in the frequency of the Fy^a allele from the parent to offspring generations (table 10).

In the backcross (table 11), there is a significant excess (.05 level) of the Fy(a+) phenotype, shown to be due primarily to a deficiency of Fy(a−) phenotypes in the Fy(a−) male by Fy(a+) female matings ($\hat{\theta} = .44$). The effect is in the opposite direction to that expected from maternal-fetal incompatibility. This distortion of the genetic ratio is consistent with the gene-frequency change (table 10) but is not confirmed by the intercross data in table 11. The studies by Reed et al. [11] in this population and Reed [20] in California indicated no mortality differentials associated with the Duffy system, whereas Morton et al. [12] found a decrease in living children from incompatible matings. But, as here, the latter study did not detect a deficiency of children with the incompatible phenotype.

Kidd. A portion of the individuals in the study were typed for Jk^a and the remainder for Jk^b. In the group of individuals tested with anti-Jk^b, there is a departure in children's phenotype frequencies from those predicted using parental gene frequencies, significant at the .05 level of probability (table 10). The estimated increase in the frequency of the Jk^b allele between the two generations is .05, an improbably large change to be attributable to selection and a discouragingly large estimate of the stochastic forces at work in data of this sort. With anti-Jk^a there is a smaller increase, .01, in the Jk^b gene frequency (also table 10). The combined results do not differ significantly from expected ($\chi^2 = 3.79$, 1 df).

There were no significant deviations of genetic ratios (table 11) for samples tested

with either anti-Jk^a or anti-Jk^b. Differential mating frequencies and number of tested children per mating (table 9) did not deviate from expected values in the sample tested with anti-Jk^b although observed differences were in the direction suggested by the shift in gene frequencies. However, there was a significant (at .05 level of probability) difference in frequency of the reciprocal backcrosses in the group tested for Jk^a sera (table 9).

P. The analyses of mating type and recorded offspring per mating type (table 9) reveal no significant effects. There is a small, but nonsignificant, decrease in the frequency of the P^1 allele in the children (table 10), which is consistent with an effect to be noted below.

In table 11, the analysis of genetic ratios indicates a significant excess of P(−) which is apparent in both the backcross and intercross. Previous studies have noted a tendency for misclassification of P(+) individuals as P(−), especially in the younger age groups (referred to in [21]). This effect is present in this study. We found [1] a highly significant deficiency of the P^1 allele and the P(+) phenotype in the zero-to-nine age group. Such an age effect is apparently responsible for the significant excess of P(−) phenotypes seen in table 11. The genetic ratio for backcrosses based on children who were in the first decade is .5866, as contrasted with .4963 when only children 10 years or older are included. Therefore it seems clear that misclassification could account for the distortions of genetic ratios in table 11.

Lewis secretion. Neither mating-type frequencies nor offspring per mating deviate significantly from expected (table 9). However, there is a significant decrease (at .05 level of probability) in the *Le* allele in children (table 10). The results of the genetic-ratio analysis indicate that the average genetic ratios—$\hat{\theta} = .4248$ for the backcross, and $\hat{\theta} = .2630$ for the intercross—are not significantly different from expected. There is, however, heterogeneity among sibship sizes and between reciprocals in the backcross. The 76 secretor ♂♂ by nonsecretor ♀♀ matings gave an estimate of $\hat{\theta} = .3140$ (significant at .01 level), whereas the 59 reciprocals gave an estimate of .5755 (not significant at .05 level). The heterogeneity among estimates from different size families is, in the main, associated with the former mating. For 72 of these backcross matings with at least four children, the results were as shown in table 12. It is more than a little difficult to ascribe a reasonable interpretation to the large deviations in the families of sizes one and two. The 1.3% excess of nonsecretor offspring in the

TABLE 12

ANALYSIS OF THE Le BACKCROSS RATIO FOR VARIOUS FAMILY SIZES

Size of Family	No. Families	$\hat{\theta}$	χ^2 (1 df)
1	29	.2645	2.65
2	23	.2125	6.75
3	15	.4692	0.04
4	5	.4916	0.00

1,026 intercross matings can easily account for the 1.87% increase in the nonsecretor allele frequency recorded in table 10.

ABH secretion. There is no distortion of expected parental mating-type frequencies or representation of children among mating types (table 9). In the children, there is a deficiency of secretors significant at the 5% level of significance (table 10). There is a highly significant deficiency of secretors among progeny from the intercross (table 11), which substantiates the deficiency of secretors detected in the total group of children.

Simultaneous estimates of θ and h resulted in a poorer fit ($\chi^2 = 14.98$, $P < .001$) to the data than when only θ was estimated, indicating that the genetic ratio is disturbed, or the frequency of nonsegregating matings is not as expected, or both. The a priori method, which ignores all families with no segregants (hence h is not involved), gave a χ^2 of 0.004. This result would suggest that the departure of the genetic ratio is not contributing to the poor fit above but that the deviation occurs because parental gene frequencies are not accurately predicting the expected proportion of nonsegregating matings.

Haptoglobin. No significant effects are found in table 9. However, there is a deficiency of the Hp 2-2 phenotype in children when the expectation is computed from parental gene frequencies. The decrease in the Hp^2 gene frequency in children is significant ($\chi^2 = 5.28$, $P < .025$). Although the proportion of Hp 2-2 offspring does not differ significantly from expected in any of the crosses (table 11), there is in each case a deficiency of segregants carrying Hp^2.

Gc. There is a significant decrease in children's Gc^1 frequency ($\chi^2 = 4.20$, $P < .05$) in table 10. However, no other significant deviations were observed in any of the analyses.

DISCUSSION

Since 1960, the people who live in, or in close proximity to, the town of Tecumseh, Michigan, have been the subject of a total community health study. The primary objective has been to define the nature and distribution of disease in a natural community. It is hoped that the knowledge derived can be used to develop approaches to the prediction (and ultimate prevention) of the onset of major illness in healthy individuals. From the first, these studies have been designed to detect the role that genetic factors play in the etiology of disease. Consequently, data pertinent to a genetic analysis were collected in conjunction with the variables of epidemiological interest. This paper has reported the results of studies of 11 genetic systems of the erythrocyte and blood serum to determine whether there are departures from random distributions of sets of well-defined genes and genotypes in families, such as might suggest the operation of selective factors in this population.

Although there is ample documentation of the role the gene may play in determining certain disease states, for none of the really common diseases is the nature of the genetic component at all clear. Conversely, the biological relevance of inherited variation at a large portion of the genetic loci has become a major issue in modern biology (see [22]). Estimates in man [23, 24], Drosophila [25, 26], and other animals [27] indicate that 30%–40% of the loci may be polymorphic in outbreeding populations (two or more alleles, each with a frequency greater than .01). This variability, if main-

tained by selection, could be the basis for differential disease susceptibilities which would, in turn, be reflected in abnormal genetic ratios.

Some have recently argued [28, 29] that much of the observed genetic variation at the molecular level is selectively neutral and exists in the population because of mutation and stochastic forces (random genetic drift). This reasoning has been encouraged by (1) lack of evidence for selective differentials (see [30] for a review of human studies), and (2) theoretical considerations of the consequences to the average fitness of the interbreeding population if the polymorphic alleles are being maintained by selective forces (discussed in [25, 31, 32]). Needless to say, it is most difficult to present data which could prove the hypothesis that most of the observed genetic variation does not contribute to the adaptive characteristics or evolutionary success of the species. We must recognize that inferential studies can only fail to reject this hypothesis of no contribution of selection—a point not emphasized by all advocates of the "neutral hypothesis."

The alternate point of view holds that it is premature to retreat from determinism simply because current formulations cannot accommodate so much inherited variation, and argues that the biological system is not as loosely organized as the proponents of "non-Darwinian evolution" would suggest. This alternate contention, that selection plays a larger role in the maintenance of the polymorphisms, is equally void of satisfactory data which could reject the random hypothesis. The difficulties in obtaining the necessary data are manyfold: (1) the effects necessary to drive gene frequencies are probably much smaller than experimental error [33], and sample size has been inadequate [22]; (2) differing selective coefficients in varying environments over time [34] make assignment of fitness values impossible; and (3) there has been scant success in recognizing and measuring how molecular variation contributes to the phenotypic array which is subject to selection.

The objective, then, of this study has been an attempt to obtain comprehensive estimates of selection effects on each marker system from a sample drawn from a human population which has been better defined than any heretofore available. By analyzing a number of systems and contrasting the results from the analysis of each system, we hoped to generate some statement as to which marker systems in this population are unlikely to be undergoing selection, which could be subject to selection, and which are most likely to be determined by selective forces.

The proper test for departures of the genetic ratio from expected values has been the subject of much discussion. No method has been completely satisfactory in that it takes into account finite samples and the effect of binomial variability in small families. In this analysis there were 76 tests of $\theta = .25$ (or .50) based on mating types with reciprocal designations and size classifications pooled (table 13). We have compared the test of agreement of $\hat{\theta}$ with expectation by the approach in this paper (eq. 11), using the expected variance evaluated at $\hat{\theta}$ (and $\hat{h}$ when appropriate), with the approach of Morton (eq. 14) which utilizes the expected variance of $\hat{\theta}$ evaluated at θ_0, that is, K^{-1}. The two different χ^2 tests for each contrast were concordant as to significance for 74 (97%) of the 5% tests which were statistically independent. In 59 cases the χ^2 based on U and K (eq. 14) was larger, while the method utilized here gave a larger χ^2 in 15 cases and equivalence in the two remaining tests. Both proce-

TABLE 13

NUMBER OF INDEPENDENT TESTS OF GENETIC RATIOS (DESIGNATED BY COLUMN IN TABLE 11) SIGNIFICANT AT THE .05 LEVEL OF PROBABILITY

System	θ_0 (Col. 9)	Test of Heterogeneity for θ		Percentage Significant at	
		Reciprocal (Col. 11)	Sizes (Col. 12)	.05	.01
ABO	0:27	0:18	0:24	0.0	0.0
MNSs	1:10	0:4	0:10	4.2	4.2
Rh	1:12	0:5	1:12	6.9	6.9
Kell	1:5	0:2	0:4	10.0	10.0
Duffy	1:2	0:1	0:2	20.0	0.0
Kidd	0:4	0:1	0:4	0.0	0.0
P	2:2	0:1	0:2	40.0	0.0
Lewis secretion	0:2	1:1	1:2	40.0	0.0
ABH secretion	1:2	0:1	1:2	40.0	40.0
Haptoglobln	0:5	0:2	0:5	0.0	0.0
Gc	0:5	0:2	0:5	0.0	0.0
Percentage significant at .05	9.21	2.63	4.17	5.91	.....
Percentage significant at .01	5.26	0.00	1.39	.....	2.15

NOTE.—The number of tests which are significant is given before colon and the total number of tests is given after colon.

dures were concordant for the four cases significant at the 1% level of significance. For small samples (number of families less than 10), $I_{\hat{\theta}}^{-1}$ was greater than K^{-1}, whereas for larger samples the reverse was true. The results of tests for heterogeneity of $\hat{\theta}$ based on U scores (eq. 15) versus those based on our weighted analysis (eq. 12) were concordant as to significance in 96.5% of the cases. Comparisons of the alternate methods when both parameters (θ_0, h_0) were tested simultaneously were concordant in all cases.

Little is known concerning the statistical bias or power of alternate tests of the genetic ratio from family data. It would seem from the above considerations that the χ^2 based on the deviations from the null and observed variances is somewhat more conservative (at least in studies of this size) than the corollary based on U and K scores. A study in depth to compare the statistical properties of the two χ^2 tests discussed here (to include also other methods which have been suggested for the treatment of the problem) is presently under way.

The results of the analyses suggest a few generalizations. First, mating-type frequencies and phenotype frequencies of parents did not deviate from Hardy-Weinberg expectations based on parental gene frequencies (the reciprocal mating-frequency effect in the Kidd-B analysis was the only exception). This is not surprising in view of the general insensitivity of Hardy-Weinberg perturbations [35–37]. There was no evidence that the number of children presenting for examination was a function of parental mating type for any system. In contrast, a significant χ^2 was associated with eight of the 17 analyses of children's phenotype frequencies (MN, Rh-E, Duffy, Kidd-b sera, Lewis secretion, ABH secretion, haptoglobin, and Gc; table 10). In one case, Rh-E, an excess of heterozygotes was significant when expected phenotype fre-

quencies were based on the children's gene frequencies. The other significant deviations occurred when expected phenotype frequencies for children were computed from parental gene frequencies. The Duffy, Kidd-b, Lewis-secretion and ABH-secretion results could be due to deviations from Hardy-Weinberg expectations and/or a difference in gene frequencies between parents and children. The MN, haptoglobin, and Gc indicated a significant increase in the χ^2 associated with the substitution of parental gene frequencies to compute expected phenotype frequencies of the children. There was, more or less, a correspondence of these results with the analyses of genetic ratios given in table 11. The significant deviations in genetic ratios [deficiency of N in the MN $\times$ M mating, Fy(a+) excess in the Fy(a+) $\times$ Fy(a$-$) backcross, and secretor deficiency in the ABH-secretion intercross] could account for the observed change in gene frequency in children. For the others, the deviations of genetic ratios were indicative of the magnitude and direction of the difference between gene and phenotype frequencies of parental and offspring generations. These apparent changes in gene frequencies between parents and offspring were not detected as an age effect in the analysis of the total population [1].

A closer look at the tests to detect departures of genetic ratios from expectations (summarized in table 13) reveals that, in general, neither effects of reciprocal mating types nor effects of family size were statistically significant. One exception, the heterogeneity among estimates from reciprocal matings and families differing in number of children in the Lewis-secretion system, stands out but defies interpretation at this time. In contrast to the few significant tests for heterogeneity, the pooled genetic ratio estimated for each mating type (table 11, col. 7) gave χ^2 values (table 11, col. 9) which exceeded the 5% and 1% critical values, 9.21% and 5.26%, respectively (table 13). Hence, overall, the test of the genetic ratio is significant two to three times more often than expected by chance alone. However, these effects are distributed among six of 11 systems studied, and within those systems showing a significant deviation, the effect is not present for all mating types (P is an exception but must be disregarded as a deviation due to possible misclassification of younger children). Of the five remaining systems, four (MN, Rh-E, Kell, and ABH secretion) have χ^2 values sufficiently large to be expected in less than 1% of tests by chance alone. But the effects are not consistent for all mating types. The MN effect is apparent in backcrosses but not intercrosses; the Kell effect occurs for only the 39 Kp(a+b+) $\times$ Kp(a+b$-$) matings, and the Rh-E and ABH secretion effects are present only in intercrosses, while backcrosses deviate very little from expectation.

In an effort to discern any trend in the estimates of genetic ratios, the pooled backcrosses for each system were examined (table 14). The frequencies of heterozygotes were also estimated in cases with no dominance. We note that although the average ratio for all systems is very near the expected 0.50, individual systems vary considerably. The variation in the estimated selection coefficient necessary to account for the corresponding deviation in the ratio is even more impressive. Likewise, the average frequency of heterozygotes is essentially one-half. One might conclude from such average values (as did Morton et al. [12]) that results of the analysis of genetic ratios do *not* point to evidence for selection. An alternate interpretation might be that the observed variation in genetic ratios (13.3% of backcross ratios significant at the 5%

TABLE 14

POOLED ESTIMATES OF $\hat{\theta}$ AND AVERAGE FREQUENCY OF HETEROZYGOTES WITH χ^2 TESTS OF THE NULL HYPOTHESIS OF EQUAL PROPORTIONS IN OFFSPRING OF BACKCROSS MATINGS

System	$\hat{\theta}$	s†	χ^2	Frequency of Heterozygotes	χ^2
ABO	.4853	+.057	1.75	.....	.....
MN	.4683	+.119	6.70**	.4819	2.12
Ss	.5022	−.009	0.03	.5022	0.02
Rh-C	.4939	+.024	0.22	.5216	2.77
Rh-D	.4874	+.049	0.22	.....	.....
Rh-E	.5027	−.011	0.03	.4974	0.02
Kell	.5409	−.178	3.58	.4592	3.57
Duffy	.4652	+.130	4.19*	.....	.....
Kidd-Jk^a	.4991	+.004	0.00	.....	.....
Kidd-Jk^b	.4464	+.194	1.24	.....	.....
P††	.4963	+.015	0.23	.....	.....
Lewis secretion	.4248	+.261	1.79	.....	.....
ABH secretion	.5083	−.034	0.15	.....	.....
Haptoglobin	.4982	+.007	0.02	.5161	1.22
Gc	.4983	+.007	0.01	.5024	0.02
Weighted mean	.4918	.029	.....	.5004	.....

* Significant at .05 level of probability.
** Significant at .01 level of probability.
† See text, equation (16).
†† Based on children 10 years of age or older.

level of probability) is indicative of nonrandom, that is, deterministic, forces operating in this population on at least a portion of these polymorphic loci.

This alternate point of view recognizes several considerations. As we have emphasized, the studies to date are capable of detecting at the 1% probability level only effects of a magnitude which seem inherently improbable. The effects observed in the various studies have not been consistent. At this time, the data do not exclude the possibility that in several of the systems under examination, there are indeed distortions of genetic ratios, indicative of selective differentials amounting to 0.02 or 0.03—detected in some studies because of sampling error which moves them in a direction favorable to detection, not detected in other studies because of sampling error in the other direction. In this connection, it is sobering to note that none of the studies to date has demonstrated an effect on the reproductive pattern of D incompatibility between husband and wife, although the existence of such an effect is incontestable.

Even more perplexing is the interpretation of very large, statistically significant deviations in genetic ratios (MN, s = .12; Rh-E, s = .33; and Kell, s = .41). Complete loss of the segregant class (s = 1) from a mating type is equivalent to 50% mortality of the offspring (percentage mortality = 100 $s/2$). It is unlikely that the cumulative nonrandom mortality until the age of inclusion in this study is greater than 10%. While a differential mortality of 6% for the MN backcrosses may be a possibility, certainly selection coefficients of .33 and .41 would imply unrealistic losses even though the proportion of such matings in the population might be small. Such an interpretation suggests that nonbiological deviations may also play a role in exag-

gerating the genetic ratio and reinforces the need to examine the distribution of apparent selection coefficients among genetic systems.

As future studies unfold, we suggest it is improper to pool numbers across systems [38] presumably for the purpose of obtaining an impressive total. Despite our predilection for determinism as a general explanation of the polymorphisms, we are prepared to concede that at some polymorphic loci the alleles may have always been neutral, and at others the original selective force responsible for the polymorphisms has since disappeared. Given the existence of some polymorphic loci at which selection is *not* occurring, pooling across loci can only obscure the findings at those loci at which selection is occurring. Furthermore, it is very likely that those polymorphs which are being maintained by nonrandom forces are not all a result of heterozygote excess [39]. If (as Dempster [40] states so clearly) diversity is being maintained to some degree by variation of selection coefficients in time and space, we must treat the data in a way so as to recognize the systems of interaction which have been emphasized by the work of Wright (see [41] for a discussion).

In the current study we have attempted to detect temporal trends and gross departures from expected genetic ratios. Fertility effects have been considered only in terms of numbers of living children presenting for examination, a less than completely satisfactory measure. Thus, fertility differences associated with differential infant mortality, or reproductive compensation related to genotype, would probably not be detected in these studies. And, of course, if frequency- or density-dependent selection exists in man, as is being increasingly demonstrated for experimental organisms (for review, see [42–45]), one study can well be "negative" where another is "positive."

It seems clear, viewing past efforts to detect selection (including our own), that the studies of the "next round" in the effort to understand the forces maintaining the polymorphisms must be both far more extensive and far more detailed than those to date. If we accept the reasonable hypothesis that selective coefficients are not unrelated to the environment in which the genotype finds itself, there is also a need for studies in diverse environments. We must recognize that in those areas where it is feasible to perform the necessary large-scale studies, man in his numbers and his technology may have so altered the environment that the original selective pressures can no longer be detected. We must also recognize the composition of the study population. If genetic coadaptation exists, then estimates of selection based on populations where major hybridization has occurred recently may have no relevance to populations where only minor hybridization has occurred in the ancestry. Thus, the results of studies such as Morton et al. [12] in Brazil can scarcely be taken as a basis to challenge the results of others in different types of populations.

As we consider the next round in such studies, in which one would surely hope to detect selective differentials of the order of 0.02, it is important to examine closely two factors which could limit such studies. We calculate that in Tecumseh, utilizing the data from the MN system according to the procedure suggested by Schacht and Gershowitz [2], there are 3.82% of discrepancies between legal and biological parentage. Since in the total material there were 3.76% of the children with paternity exclusions, this implies that 98% of the discrepancies have been detected. Those missed will be cases in which the discrepancy did not introduce a usually detectable antigen, as in

O in the ABO system, d of the Rh system, k of the Kell system, or Fy^b of the Duffy system. This could be confused with selection against the gene responsible for a detectable antigen. Of equal importance may be typing errors. It has previously been demonstrated that with single typings, the differences between two competent laboratories may range from 0.2% for the MN system to 6.0% for the Duffy system (Gershowitz 1970, personal communication). Discrepancy rates are especially high in those systems in which the Coombs antiglobulin reaction is part of the test. Since here the Coombs sera and the conditions of testing are so critical, simple duplicate tests employing the same sera and reaction circumstances may not be as critical as they appear at first glance. Rather, tests in different laboratories, employing different lots of antisera, seem indicated, thus adding greatly to the labor of the study.

There may, however, be one important factor offsetting these limitations. The analyses of the past have been directed toward first-order effects. But since selection acts on the total phenotype—as envisioned by some, on the most disadvantaged tail of a distribution of fitnesses—future studies must be concerned with interacting gene systems. In a future paper, we will make an initial attempt at that approach with these data. Such studies had best be prospective—associations may be expected to emerge more strongly among the small fraction dying than among the large fraction surviving.

Given these reservations, the future of studies of this type is moot. It is clear that, conducted on the present scale, they can only implicate or exclude selection coefficients which in their magnitude are improbably large (table 11) save for the occasional exceptional case. Inspection of the number of families required to detect a departure from the expected genetic ratio (table 11, col. 10) suggests that it will require studies roughly five to 10 times the size of the present one to demonstrate selective differentials between genotypes of .01 to .02, even when the gene frequency is favorable. Material permitting generalizations is not apt to emerge from the operations of blood banks, maternity hospitals, or military induction centers. Special efforts are required. Even if funds were available, the high probability of an inconclusive outcome and the magnitude of the necessary funding are apt to discourage even the most dedicated investigators.

In fact, the only spur to investigators at this point is the fundamental nature of the question concerned. One is accordingly led to wonder whether future studies can be combined with other large-scale genetic investigations to which society finds itself committed. Should, for instance, systems of continuous monitoring for increased mutation rates be established in designated areas, involving studies of a large series of proteins derived from cord bloods from newborn infants with corresponding studies of maternal and paternal specimens, here might well be the basis for a *prospective* study of selection on the requisite scale.

SUMMARY

Analyses of 2,507 nuclear families for 11 genetic systems are presented. Mating-type frequencies and phenotype frequencies of parents did not deviate from Hardy-Weinberg expectations for any genetic system. There was no evidence that parental mating type was related to the number of children of a family that was included for study. The differences in gene frequencies between parents and children were, more

or less, concordant with deviations in genetic ratios. Significant deviations in the genetic ratio occurred two to three times more often than expected by chance alone. However, these effects were distributed among six of the 11 systems; and, within those systems showing a significant deviation, the effect was not present for all mating types. Four systems (MN, Rh-E, Kell, and ABH secretion) had departures from expected ratios which were significant at the 1% level of probability. The MN and Kell effects are apparent in backcrosses but not intercrosses. The Rh-E and secretion effects are present only in intercrosses, while backcross ratios are as expected.

ACKNOWLEDGMENTS

We wish to express our gratitude to Barbara Eggleston for her dedication to the details of manipulating this large body of data for the analyses presented here. The advice and counsel of Dr. W. J. Schull throughout this study is gratefully appreciated.

REFERENCES

1. SHREFFLER DC, SING CF, NEEL JV, et al.: Studies on genetic selection in a completely ascertained Caucasian population. I. Frequencies, age and sex effects, and phenotype associations for 12 blood group systems. *Amer J Hum Genet* 23:150–163, 1971
2. SCHACHT LE, GERSHOWITZ H: Frequency of extra-marital children as determined by blood groups, in *Proceedings of the 2d International Congress of Human Genetics*, Rome, 1961, 1963, pp 894–897
3. FISHER RA: The effect of methods of ascertainment upon the estimation of frequencies. *Ann Eugen* 6:13–25, 1934
4. BAILEY NTJ: *Introduction to the Mathematical Theory of Genetic Linkage.* Oxford, Oxford Univ. Press, 1961
5. MORTON NE: Genetic tests under incomplete ascertainment. *Amer J Hum Genet* 11:1–16, 1959
6. MORTON NE: Segregation and linkage, in *Methodology in Human Genetics*, edited by BURDETTE J, San Francisco, Holden Day, 1962, pp 17–52
7. MI MP: Segregation analysis. *Amer J Hum Genet* 19:313–321, 1967
8. ELANDT-JOHNSON RC: Segregation analysis for complex modes of inheritance. *Amer J Hum Genet* 22:129–144, 1970
9. ELANDT-JOHNSON RC: *Probability Models and Statistical Methods in Genetics.* New York, Wiley. In press
10. REED TE, SCHULL WJ: A general maximum likelihood estimation program. *Amer J Hum Genet* 20:579–580, 1968
11. REED TE, GERSHOWITZ H, SONI A, et al: A search for natural selection in six blood group systems and ABH secretion. *Amer J Hum Genet* 16:161–179, 1964
12. MORTON NE, KRIEGER H, MI MP: Natural selection on polymorphism in northeastern Brazil. *Amer J Hum Genet* 18:153–171, 1966
13. CHUNG CS, MATSUNAGA E, MORTON NE: The ABO polymorphism in Japan. *Jap J Genet* 5:124–134, 1960
14. MATSUNAGA E, ITOH S: Blood groups and fertility in a Japanese population. *Ann Hum Genet* 22:111–131, 1958
15. PERITZ E: A statistical study of intrauterine selection factors related to the ABO system. *Ann Hum Genet* 30:259–271, 1967
16. MORTON NE, CHUNG CS: Are the MN blood groups maintained by selection? *Amer J Hum Genet* 11:237–251, 1959
17. CHUNG CS, MATSUNAGA E, MORTON NE: The MN polymorphism in Japan. *Jap J Genet* 6:1–11, 1961
18. WIENER AS, DIDIEGO N, SOKOL S: Studies on the heredity of the human blood groups. The M-N types. *Acta Genet Statist Med* (Basel) 2:391–398, 1953

19. HIRAIZUMI Y: Are the MN blood groups maintained by heterosis? *Amer J Hum Genet* 16:375–379, 1964
20. REED TE: Research on blood groups and selection from the Child Health and Development Studies, Oakland, California. III. Couple mating type and reproductive performance. *Amer J Hum Genet* 20:129–141, 1968
21. GERSHOWITZ H, NEEL JV: The blood groups and secretor types in five potentially fatal diseases of Caucasian children. *Acta Genet Statist Med* (Basel) 15:261–308, 1965
22. NEEL JV, SCHULL WJ: On some trends in understanding the genetics of man. *Perspect Biol Med* 11:565–602, 1968
23. HARRIS H: Enzyme polymorphisms in man. *Proc Roy Soc* [*Biol*] 164:298–310, 1966
24. LEWONTIN RC: An estimate of average heterozygosity in man. *Amer J Hum Genet* 19:681–685, 1967
25. LEWONTIN RC, HUBBY JL: A molecular approach to the study of genic heterozygosity in natural populations. II. Amount of variation and degree of heterozygosity in natural populations of *Drosophila pseudoobscura*. *Genetics* 54:595–609, 1966
26. O'BRIEN SJ, MACINTYRE RJ: An analysis of gene-enzyme variability in natural populations of *D. melanogaster* and *D. simulans*. *Amer Naturalist* 103:97–113, 1969
27. STONE WS, KOJIMA K, JOHNSON FM: Enzyme polymorphisms in animal populations (abstract). *Jap J Genet* 2:153, 1969
28. KIMURA M: The rate of molecular evolution considered from the standpoint of population genetics. *Proc Nat Acad Sci USA* 63:1181–1188, 1969
29. KING JL, JUKES TH: Non-Darwinian evolution. *Science* 164:788–798, 1969
30. REED TE: The evidence for natural selection due to blood groups, in *Proceedings, United Nations World Population Conference*, Belgrade, Yugoslavia, 1965, vol 2, 1967, pp 498–502
31. KIMURA M, CROW JF: The number of alleles that can be maintained in a finite population. *Genetics* 49:725–738, 1964
32. SVED JA, REED TE, BODMER WF: The number of balanced polymorphisms that can be maintained in a natural population. *Genetics* 55:469–481, 1967
33. EWENS WJ: *Population Genetics*. London, Methuen, 1969
34. LEVINS R: *Evolution in Changing Environments*. Princeton, N.J., Princeton Univ. Press, 1968
35. WALLACE B: The comparison of observed and calculated zygote distributions. *Evolution* 12:113–115, 1958
36. LEWONTIN RC, COCKERHAM CC: The goodness-of-fit test for detecting natural selection in random mating populations. *Evolution* 13:561–564, 1959
37. WARD RH, SING CF: A consideration of the power of the χ^2 test to detect inbreeding effects in natural populations. *Amer Naturalist* 104:355–365, 1970
38. MORTON NE: Genetic structure of northeastern Brazilian populations. Symposium 45, Wenner-Gren Foundation. In press
39. FORD EB: Polymorphism and taxonomy, in *The New Systematics*, edited by HUXLEY J, Oxford, Clarendon, 1940, pp 493–513
40. DEMPSTER E: Maintenance of genetic heterogeneity. *Cold Spring Harbor Sympos Quant Biol* 20:25–32, 1955
41. WRIGHT S: *Evolution and Genetics of Populations*, vol 2, Chicago, Univ. Chicago Press, 1969
42. EHRMAN L, SPIESS EB: Rare-type mating advantage in *Drosophila*. *Amer Naturalist* 103:675–680, 1969
43. KOJIMA KI, YARBROUGH KM: Frequency-dependent selection at the esterase-6 locus in *Drosophila melanogaster*. *Proc Nat Acad Sci USA* 57:645–649, 1967
44. WALLACE B: *Topics in Population Genetics*. New York, Norton, 1968
45. WALLACE B: *Genetic Load, Its Biological and Conceptual Aspects*. Englewood Cliffs, N.J., Prentice-Hall, 1970