11116

NATIONAL LIBRARY
OTTAWA

BIBLIOTHÈQUE NATIONALE
OTTAWA

NAME OF AUTHOR... DAVID ALAN BIRDSALL

TITLE OF THESIS... Allele and Phenotype Frequency Variation at three loci in natural populations of deer mice (Peromyscus maniculatus).

UNIVERSITY... University of Alberta

DEGREE FOR WHICH THESIS WAS PRESENTED... Ph.D.

YEAR THIS DEGREE GRANTED... 1971

Permission is hereby granted to THE NATIONAL LIBRARY OF CANADA to microfilm this thesis and to lend or sell copies of the film.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

(Signed) D. A. Birdsall

PERMANENT ADDRESS:

Fisheries Research Board of Canada
Institute of Animal Resource Ecology
University of British Columbia
Vancouver, B.C.

DATED... Nov 22 ... 1971

NL-91 (10-68)

THE UNIVERSITY OF ALBERTA

ALLELE AND PHENOTYPE FREQUENCY VARIATION OF THREE LOCI IN NATURAL POPULATIONS OF DEER MICE *(PEROMYSCUS MANICULATUS)*

by

© DAVID ALAN BIRDSALL

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF GENETICS

EDMONTON, ALBERTA

SPRING, 1972

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled Allele and Phenotype Frequency Variation in Natural Populations of Deer Mice *(Peromyscus maniculatus)*, submitted by David Alan Birdsall in partial fulfilment of the requirements for the degree of Doctor of Philosophy.

.....
Supervisor

.....

.....

.....

.....

.....
External Examiner

Date. November 19, 1971 .....

ABSTRACT

A study of the distribution of allele and phenotype frequencies of three loci in natural populations of deer mice *(Peromyscus maniculatus)* was carried out during the years 1968-71. Over 2000 deer mice were captured in woodlots in the Edmonton, Alberta area and approximately 2500 more were raised in a laboratory breeding colony.

The phenotypes of all mice were determined for three loci: the *Tf* (transferrin) locus with two alleles common in the populations; the *Pm* (blood group) locus with three alleles in the populations; and the *Ng* locus with three alleles common in the populations. The *Ng* locus, described during this study, controls the appearance of white bands on starch gels stained for esterase activity.

The results of the laboratory breeding showed examples of segregation distortion at the *Pm* and *Tf* loci. Deterministic computer simulation of populations subject to these distortions revealed that the distortion at the *Tf* locus is probably of little importance, but it would be expected that the distortion at the *Pm* locus is balanced by selection against the homozygote favored by the distortion.

Evidence for selection which is strong enough to balance the segregation at the *Pm* locus was obtained. In addition, both *Pm* locus homozygotes not favored by the distortion appear to be at a disadvantage. There appears to be selection against all three homozygotes and one heterozygote of the *Ng* locus system. Indirect evidence for selection against *Tf* locus homozygotes was obtained.

Because evidence was obtained for selection acting on the phenotypes of all of the three loci studied, it was concluded that it is unlikely that cryptic variation is frequently maintained by stochastic phenomena. Also,

because the estimates of selection are great enough so as to not be compatible with simplistic calculations of genetic load, it was suggested that theories of genetic load involving soft selection and fluctuating selection pressures are more applicable.

ACKNOWLEDGEMENTS

As the reader will discover, the work described here was to a great extent a cooperative effort between the author and Dr. David Cameron, Elizabeth Savage and Dr. Ray Canham. Without their help, this study would not have been possible.

I owe a great deal to my supervisors, Drs. David Cameron and David Nash. In retrospect, their patience appears enormous.

Special thanks are due Roxy Smiley, who did much of the work with the breeding colony. In addition, I thank all of the many persons in the Department of Genetics, the Department of Zoology, and Laboratory Animal Services, who donated time and materials so as to facilitate many aspects of the work. I would also like to thank Mrs. Kay Baert who typed the manuscript.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

So little is known about the relationship between the genetic constitution of an individual and its ecological niche, that the means whereby natural genetic variation is maintained and the distribution of variation is influenced by the environment remain controversial areas of study. Although in the past, this central problem of population biology and evolution has been approached by ecological, genetical and mathematical investigations, it is only in recent years that studies combining these three approaches have been carried out. The resulting field of study is commonly called ecological genetics.

The kinds of studies which have a bearing on ecological genetics are extremely diverse. Ecologists have measured home range, population size, migration and other population parameters, but the analysis of genetic variation in natural populations has generally been ignored as a tool for describing population structure. Similarly, geneticists have frequently studied naturally occurring variation without studying the related ecology which would allow more accurate estimates of those factors controlling the distribution of that variation.

In general, however, the *limits of action* of those factors which maintain variation and affect the distribution of that variation have been defined mathematically (e.g. Fisher, 1930; Wright, 1930, 1931; Haldane, 1932). These sophisticated mathematical analyses are rarely complemented by an understanding of what actually occurs in natural populations and when analyses are needed which will complement data from natural populations, they are generally not available.

Some combined studies of ecology and genetics began shortly after the

mathematical bases of population genetics were defined, and to some extent have followed a parallel course (review: Ford, 1964). These studies usually dealt with morphological variation (e.g. Kettlewell, 1957; Dowdeswell, 1961; Creed *et al.*, 1959; Clarke and Sheppard, 1959; Cain, King, and Sheppard, 1960), which was frequently found to be subject to strong natural selection. Ford (1964) pointed out that "advantageous qualities are frequently favoured or balanced in particular environments by far greater selection-pressures than had hitherto been envisaged, ..." For example, Ford calculates the natural selection against female butterflies *(Maniola jurtina)* with two or more spots on the underside of the hind wings to be 69% in 1959 and 74% in 1960, compared with the unspotted type (Dowdeswell, 1961). Another sort of variation has been defined as genetic polymorphism: "the occurrence together in the same locality of two or more discontinuous forms of a species in such proportions that the rarest of them cannot be maintained merely by recurrent mutation" (Ford, 1940). The best known example of polymorphism is sickle cell anaemia, a genetic disease common in negro populations in Africa and the United States (review: Harris, 1959). Although the sickle cell situation is by no means entirely understood, it is of particular interest because strong selective pressures affecting the variation have been identified (Haldane, 1949; Allison, 1954), and the specific molecular variation which causes the disease has been described (Pauling *et al.*, 1949; Ingram, 1957).

Some general conclusions can be made from these sorts of studies. Sheppard (1959) asserted that "Polymorphism will only be found where selective agencies are balanced in such a way as to maintain two or more types in a population." Sheppard's statement is based on the assumption that variation in finite populations will disappear over the long term because of stochastic phenomena, unless stabilized by selection. This disruption of stability of

variation is simply the gradual change in allele frequency, ending with fixation of one allele, which results from sampling a finite number of gametes each generation. Because of this it appears to be necessary to explain the maintenance of any particular stable variation with some form of natural selection.

The conclusion that selection is associated with most stable variation has precipitated a dilemma. As Haldane (1937) initially pointed out, variation in fitness is expensive for the population. The condition in which one type of individual is favored with respect to survival and reproduction over another type demands that those with lesser fitness suffer some death or degree of reproductive failure. If this death or reproductive failure (reduced fitness) is caused by variation at very many loci, then the cost of this natural selection (genetic load) becomes intolerable. After a mathematical analysis of this situation, Kimura and Crow (1964) concluded that populations are of necessity monomorphic at all but a few loci. More recent studies, however, using techniques (electrophoresis and immunology) which make subtle distinctions between gene products, indicate that the amount of variation present in natural populations is indeed of a level which could not exist if this relationship between genetic variation and variation in fitness is in fact obligate. The proportion of loci which are polymorphic in natural populations has been estimated to be from 14 to 42% (Lewontin and Hubby, 1966; Lewontin, 1967*a*; Harris, 1969; Dessauer and Nevo, 1969; Selander, Hunt, and Yang, 1969; Selander and Yang, 1969; Selander *et al.*, 1970), and those estimates based on electrophoretic analyses may actually underestimate the amount of allelic variation by 4:1 (Lewontin, 1967*b*).

Attempts to rationally explain this dilemma have been of two kinds.

Several authors have continued with the assumption that most variation is maintained by selection (Van Valen, 1963; King, 1967; Milkman, 1967; Sved, Reed, and Bodmer, 1967; Wallace, 1970). However, they point out that selection usually acts on a phenotype which is a result of interaction between several loci and that it is unreasonable to treat each variable locus as having a discrete effect on total fitness. Also, they point out that calculations such as those of Kimura and Crow (1964) are based on the assumption that all selection is "hard" selection, meaning that all genotypes of each variable locus have fixed probabilities of survival and reproduction. They propose that most selection is "soft" selection and that a limit is placed on the selection intensity per locus which is compatible with the capacity of the species to reproduce. Survival and reproduction are thus competitive processes, and the probability of success is related to the genotypes in the total population.

The alternative solution to the dilemma is based on the proposition that a significant proportion of variant genes have alleles which are neutral in fitness relative to each other (King and Jukes, 1969; Arnheim and Taylor, 1969; Kimura and Ohta, 1971). Thus observed variation could be the result of random changes (genetic drift) in the frequencies of alleles which are constantly being generated by mutation.

Neither the proponents of selection as the major factor in the maintenance of variation, nor those who argue for random changes in frequency of neutral alleles, have a large body of experimental or descriptive data upon which to base their conclusions. The important factors of population structure have not been quantified for most organisms. (Indeed, we are largely ignorant about which factors of population structure *are* important in affecting the distribution of variation in populations). At the genetic level, we know

little about which genes vary, how many alleles of each gene are present, and what, if any, are the functional differences between alleles. Mutation rates in higher organisms have not been well established. With such large gaps in our knowledge it is expected that controversies such as the one described will continue.

As has been discussed previously, ecological genetics has joined three approaches to attack a common problem. However, two somewhat different tactics have been used. One of these is the development of increasingly sophisticated theoretical mathematical models, and subsequent testing of natural populations for the goodness of fit to those models. The alternative is to study real populations initially and to use the information garnered for the construction of practical mathematical models. In my opinion, the most reliable analyses result from the latter method, both because of the difficulty of constructing realistically complex mathematical models, and because of the absurdity of establishing a model without enough information from real populations to construct a reasonable one.

Devising and testing hypotheses using natural populations produces a variety of problems, not the least of which are the consequences of the interference with the populations. The degree to which population processes are changed simply by the presence of an investigator, or by the removal of animals, and how the interpretation of results is thus affected, is unknown. This basic "uncertainty" principle remains an unsolvable and, to some degree, an unavoidable problem.

Another chronic problem is the definition of phenotypes and the phenotype-genotype relationships. As mentioned previously, electrophoresis probably does not detect all allelic variation, because proteins, the synthesis of which are controlled by alleles of the same gene but which have the same

isoelectric point, will probably have identical electrophoretic migration rates. Immunological techniques also fail to identify all variants, for a variety of reasons. The possibility exists that misclassification will make results meaningless. For example, Rasmussen (1964), in a study of deer mouse populations in Michigan found a 29.1% deficiency of heterozygotes at the *Pm* (blood group) locus. Selander (1970) terms this result "equivocal," pointing out that a non-reactive allele at a frequency of 0.08 would account for all of the heterozygote deficiency. Moreover, this holds in most situations, because an allele which does not obviously affect the phenotype of the heterozygote and is undetectable in homozygous form (for example, because of its rarity or lethality) could conceivably occur at most serum protein and blood group loci (e.g., a null allele for transferrin, Heilmeyer *et al.*, 1961). The possibility of immunologically non-reactive alleles can be reduced by the use of a broad spectrum of antigen donors (Rasmussen, 1969). With respect to electrophoresis, however, allelic proteins with the same migration rate cannot be detected without the use of some other technique. Null alleles can sometimes be detected in heterozygous form because of the dosage effect (Semeonoff and Robertson, 1968), but are generally only exposed if in high enough frequency to be detected in a homozygote. An alternative method of detecting null alleles is through the use of a breeding program. Heterozygotes which carry a null allele (and thus have the phenotype of a homozygote) will yield unexpected phenotypes when crossed.

The basic purpose of a breeding program is, of course, to discover if variation is inherited, and in what manner. However, situations in which variation is inherited as if controlled by alleles of nuclear genes, but occur in non-Mendelian ratios, are occasionally found. Examples of segregation distortion have been reported in *Mus* (Dunn, 1965; Dunn and Suckling, 1956;

Dunn, 1957; Lewontin and Dunn, 1960), *Nicotiana* (Cameron and Moav, 1957), *Drosophila* (Hiraizumi, Sandler, and Crow, 1960), wheat (Loegering and Sears, 1963), and *Peromyscus* (Canham, Birdsall, and Cameron, 1971; Savage, 1971). The population consequences of segregation distortion have been extensively examined (Lewontin and Dunn, 1960; Lewontin, 1968; Levin, Petras, and Rasmussen, 1969; Hartl, 1970*a*, 1970*b*).

Following definition of the phenotype-genotype relationship, and the manner of inheritance of variation, allele and phenotype frequency data must then be interpreted. This interpretation is generally made in terms of factors such as migration, mutation, selection and genetic drift. The problem is thus one of distinguishing between non-exclusive processes. Any particular frequency distribution, or any change in frequency, or any difference between frequencies, can usually be interpreted in a variety of ways. However, by studying several populations, over a number of generations, the possibility of correctly interpreting similarities, differences and changes in frequency is greatly increased. It should then be possible to formulate hypotheses which can be tested in the laboratory and by population manipulation in the field.

The interpretation of allele and phenotype frequency data in order to determine the reasons for the presence of variation is one part of the study of the genetics of natural populations. In addition, however, the distribution of variation can be used to draw conclusions about population structure. Facets of population structure such as population size, mating behaviour, migration and population subdivision can be reflected in phenotype frequencies. Because of this, frequency data can be incorporated into mathematical models in order to estimate and infer the extent and type of genetic structure in

particular populations. Rasmussen (1964) used the isolation by distance and island models of Wright (1943; 1946; 1951) to estimate the effective size of deer mouse populations. This estimate was very similar to that which Rasmussen calculated from demographic data. Petras (1967) has used similar analyses in combining ecological and genetical data to conclude that populations of house mice *(Mus musculus)* on Michigan farms were highly structured into small breeding units. Additional evidence for this subdivision was reported by Selander (1970), who found non-random spatial distributions of similar genotypes of house mice in barns. Genetical discoveries of this type can complement and support general ecological studies, providing a greater understanding of population organization.

The following study of variation in deer mouse *(Peromyscus maniculatus)* populations is an attempt to discover the reasons for specific variation, and to use the distribution of that variation to describe aspects of population structure. Also, it is aimed at the formulation of hypotheses which can be tested by future experimental population manipulations.

MATERIALS AND METHODS

Choice of Organism

Peromyscus is a geographically and ecologically widely distributed nearctic genus of cricetine rodents. Its paleontology, taxonomy, anatomy, physiology, ecology and behaviour have been extensively studied (review: King, 1968). *P. maniculatus*, the deer mouse, is the only species which occurs in central Alberta. It lives in wooded areas, which are often small and separated from each other by pastures and grain fields. Deer mice were chosen for study for the following reasons:

(a) They are easily captured alive in large numbers. Although this is often true of the sympatric microtine, *Clethrionomys gapperi* (red-backed vole), and the grassland microtine, *Microtus pennsylvanicus* (meadow vole), these species undergo drastic periodic fluctuations in numbers (Fuller, 1969; Canham, 1969; Soper, 1964). Capture of adequate numbers of the microtines is difficult when their numbers are low.

(b) They are easily handled, kept in captivity, and bred.

(c) Blood samples are easily taken with generally little effect on the animal.

(d) Electrophoresis of deer mouse serum reveals a large amount of easily resolved variation. Their serum protein concentration is higher than in many other species.

(e) Many genes have been described in deer mice (review: Rasmussen, D.I., *in* King, 1968).

Lewontin (1967*b*), in pointing out the weaknesses of *Drosophila* as an animal for population studies stated that "The best compromise would be a small mammal—*Mus* or *Peromyscus*—in which laboratory breeding is possible,

many loci can be characterized in each individual, and abundant natural populations are available and can be manipulated."

Collection of Animals

Deer mice *(Peromyscus maniculatus)* were trapped in each of nine woodlots in the St. Albert area (Figure 1) during the summers of 1968 to 1970 (with some exceptions). In addition, two woodlots in the Nisku area were trapped during 1970 and two in the St. Albert area were trapped during 1971.

All the woodlots are mixed aspen and balsam poplar, birch and spruce stands surrounded by pastures and grain fields in a productive mixed farming area. Because deer mice in the Edmonton area are forest inhabitants and are only occasionally found in fields, the populations in these woodlots are isolated to some degree.

1. SR1 - 11 acres - A U-shaped woodlot in a ravine on the north side of the Sturgeon River valley.

2. RR - 9 acres - A long, narrow woodlot running down a steep ravine on the north side of the Sturgeon River valley.

3. FMA - 16 acres - Part of the wet flatland in the valley bottom on the north side of the Sturgeon River.

4. IL2 - 50 acres - In the valley bottom directly south across the Sturgeon River from FMA.

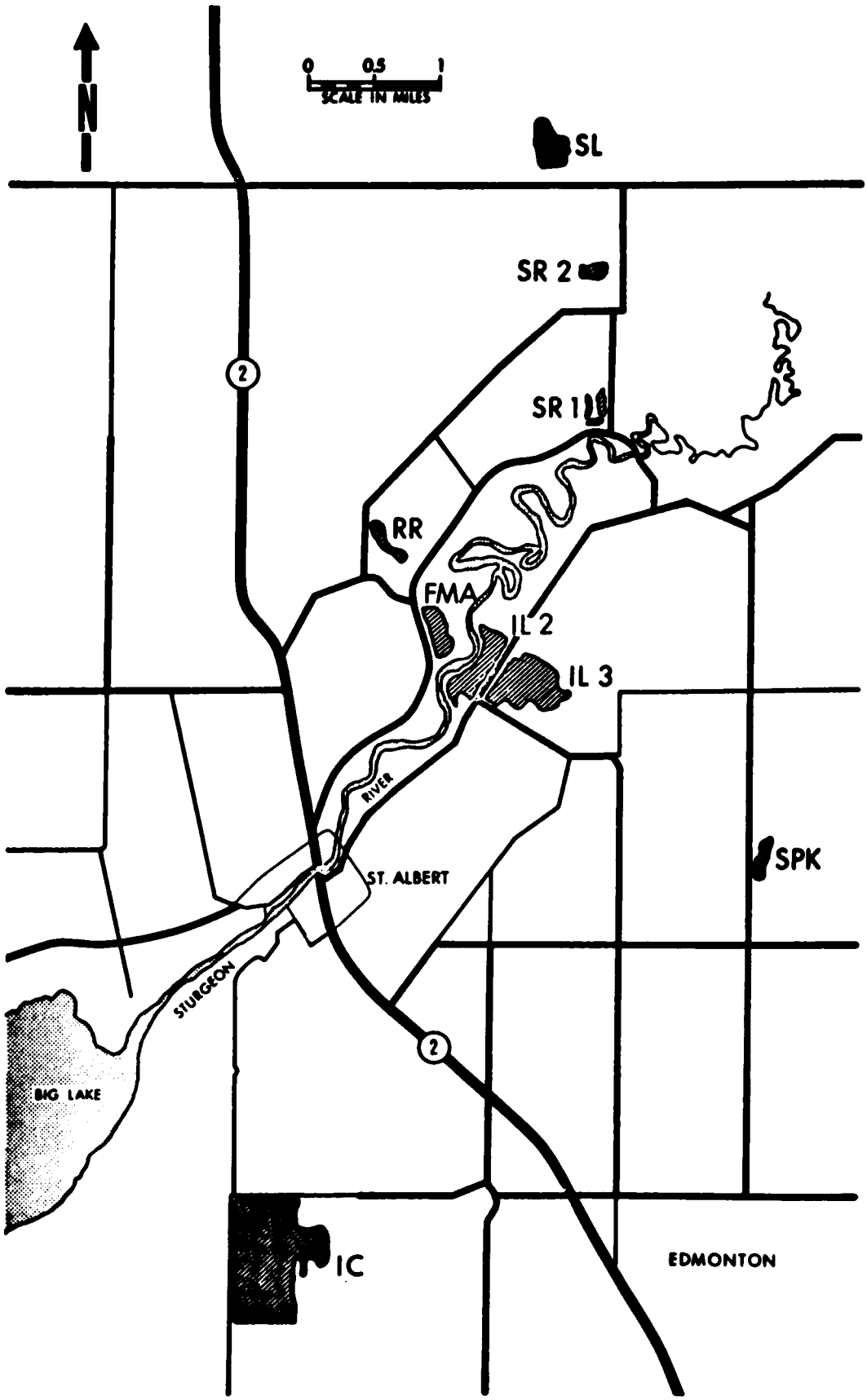
5. IL3 - 130 acres - Across a road from IL2, on the south slope of the Sturgeon River valley.

6. SR2 - 15 acres - One mile north of SR1.

7. SL - 11 acres - One mile north of SR2.

8. SPK - 14 acres - 2 1/2 miles north of the City of Edmonton limits,

Figure 1: Map showing the nine woodlots studied in the St. Albert area.


N
0 0.5 1
SCALE IN MILES
SL
SR 2
SR 1
2
RR
FMA
IL 2
IL 3
RIVER
ST. ALBERT
SPK
STURGEON
BIG LAKE
2
IC
EDMONTON

approximately three miles southeast of IL3.

9. IC - 321 acres - Three miles west of Edmonton and 2 1/2 miles south of the Sturgeon River.

10. NO - 10 acres - and NT - 35 acres - are located one mile north of Nisku Village, on the west side of the highway. NO is north of NT and separated from it by a field 100 yards wide. Nisku is approximately 13 miles south of Edmonton and NO and NT are therefore separated from all other woodlots by the North Saskatchewan River and a minimum distance of 20 miles.

The mice were captured in Sherman live traps set in several lines, each having from 20 to 50 traps. The traps were set approximately five yards apart, and left in position for three nights, with the captured mice being removed each morning. An effort was made to set traps in all parts of each woodlot. Following capture, each mouse was brought to the laboratory so that a blood sample could be taken. Most mice were kept indefinitely in captivity for use in the breeding colony.

At the time of bleeding, each animal was weighed and its body length and total length were recorded. Its sex and reproductive condition were noted. Using these criteria, it was possible to determine accurately whether an animal had lived over the previous winter or was a young of the year.

During 1970, all wild-caught pregnant females were allowed to raise their litters in captivity. These litters were bled and typed at weaning.

The Genetic Systems

The choice of genes to use as markers in the populations under study was based on the following criteria.

(a) Two or more alleles of the gene are segregating in the populations at relatively high frequency.

(b) The inheritance of the characters is simple. Phenomena such as null alleles, polygenic inheritance, and incomplete penetrance complicate some population genetic analyses.

(c) The procedure used to determine each animal's phenotype is accurate and fast enough to allow screening of large numbers of animals.

Three genetic loci were chosen.

1. The *Tf* locus (Rasmussen and Koehn, 1966) controls the synthesis of transferrin, an electrophoretically resolvable serum protein which has been found to be polymorphic in a wide variety of species (Manwell and Baker, 1970). The primary function of transferrin is to transport ferric iron from sites of storage in the liver and elsewhere to the points of synthesis of hemoglobin, myoglobin, and various enzymes (Bearn and Parker, 1966; Giblett, 1962).

2. The *Ng* locus (Birdsall, Redfield, and Cameron, 1970) was described during the course of this study. Alleles of the *Ng* locus control the electrophoretic mobility of a substance which causes the appearance of distinct white bands, contrasting with the normal gray background of starch gels stained for esterase activity. The chemical nature and biological significance of the causative agent is unknown.

3. The *Pm* locus controls the synthesis of the erythrocytic surface antigens A and B (Rasmussen, 1961), and C (Savage, 1971).

The phenotypes of all animals, both captured and raised in captivity, were determined for these three genes.

Collection and Storage of Blood

Blood samples were taken from each of the wild-trapped animals the day following capture, and from offspring raised in captivity at from four to six weeks of age. Approximately 0.5 ml of blood could be obtained by

inserting a heparinized pipet (Natelson type, Fisher Scientific Co., Montreal) into the suborbital canthal sinus (Rasmussen, 1961) of a drugged mouse (Nembutal, Abbott Laboratories Ltd., Montreal; 45 mg/Kg body weight). Cells and serum were separated by centrifugation (5000*g* for 5 min). The serum was removed with a Pasteur pipet and stored in small plastic tubes (Microfuge tubes, Beckman Instruments, Inc., Toronto) in a freezer at -40 to -50 C. During 1968 the cells were washed twice in buffered saline (Bacto Hemagglutination Buffer, Difco Laboratories, Detroit), suspended in 20% glycerol in an aqueous solution of 5% trisodium citrate (equal volumes of glycerol-citrate solution and packed cells) and frozen. Cell samples collected after 1968 were simply suspended in an equal volume of buffered saline for blood typing.

Electrophoresis and Staining

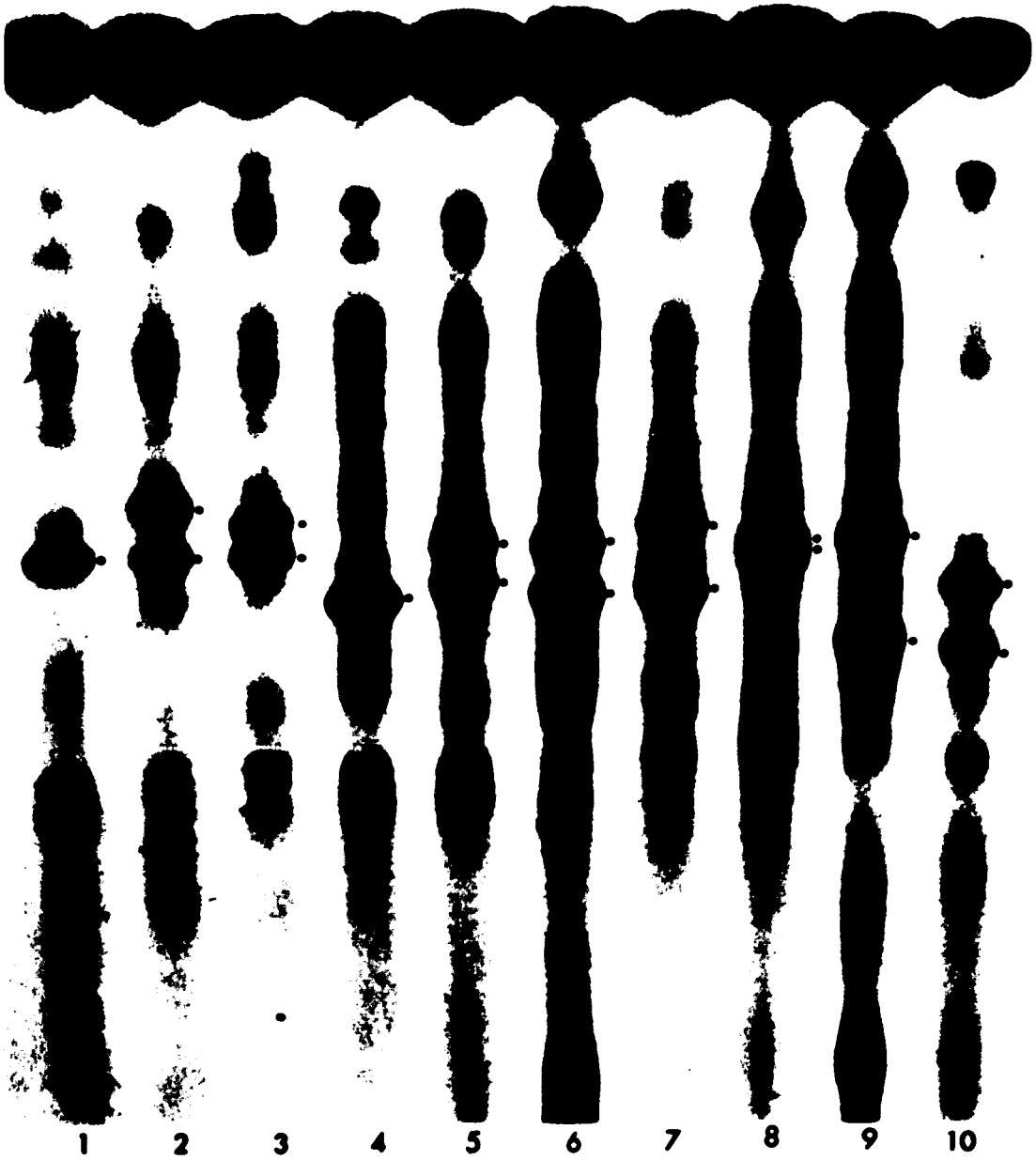
Vertical starch gel electrophoresis (Smithies, 1959) was carried out using a tris-borate discontinuous buffer system (pH 8.6) similar to that of Poulik (1957). The electrode buffer contained 11.8 g/l boric acid and 2.7 g/l lithium hydroxide, and the gel buffer contained 10.7 g/l tris (hydroxymethyl) aminomethane, 1.2 g/l citric acid and 10% electrode buffer. A 10% solution of starch (Connaught Laboratories, Toronto, or Otto Hiller Co., Madison, Wisconsin) in gel buffer (w/v) was heated, with continuous agitation, until the viscosity fell well below the initial maximum, the solution cleared, and the first minute bubbles indicative of boiling appeared. Degassing was accomplished by applying a vacuum until boiling had ceased. The starch was then poured into the gel tray (Otto Hiller Co.), the slot former cover was lowered into position so as to avoid trapping bubbles and weights were placed on the cover. Two to three hours later the weights and cover were removed.

Approximately 10 µl of whole serum was added to each of the slots in the gel, which were then covered with melted petroleum jelly. The gel was covered with "Saran Wrap" plastic film to prevent dehydration of the gel. After the end plates of the gel tray were removed, the tray was placed in a vertical position with the slots at the top. A voltage gradient was then applied to the gel, with the anode at the bottom, and the cathode at the top. The amount of voltage applied and the length of time of the run varied with each substance to be resolved.

The transferrin bands (Figure 2) were easily resolved under a wide variety of conditions. However, a voltage gradient of 500V (15.6 V/cm) was usually applied for three hours. After this time, the gel was removed from the tray and immersed for five minutes in a saturated solution of buffalo black NBR (Allied Chemical Corp., New York) in water : methanol : acetic acid (50:50:1, v/v/v). The gel background was destained with several changes of the same solvent solution, leaving bands of stained protein.

The white bands of the *Ng* locus (Figure 3) were best resolved when a voltage gradient of 600V (18.8 V/cm) was applied to the gel for three hours. The gel was stained for two hours in daylight at room temperature in the following solution: 10 ml 0.1 tris-HCl buffer (pH 7.0), 87.5 ml water, 2.5 ml 1% α-naphthyl acetate (w/v) in water acetone 1/1 (v/v), and 100 mg fast blue RR salt. The clarity of the bands was sensitive to changes in the length of time of staining, the amount of light, and the concentrations of substrate and stain. Particular lots of starch gave poorer results than others. After staining, the gels were fixed and washed in several changes of water : methanol : acetic acid, 5/5/1 (v/v/v). The bands became clearly visible after washing for 24 to 48 hr. Electrophoresis in a gel of 6 mm rather than 3 mm thickness and correspondingly lower voltage seldom produced

Figure 2: **Photograph of a starch gel showing ten phenotypes which include the nine transferrin bands found during this study:** 1. *TfJ;* 2. *TfGJ;* 3. *TfHJ;* 4. *TfM;* 5. *TfJL;* 6. *TfJM;* 7. *TfIM;* 8. *TfJK;* 9. *TfJP* and 10. *TfMQ*.


1
2
3
4
5
6
7
8
9
10

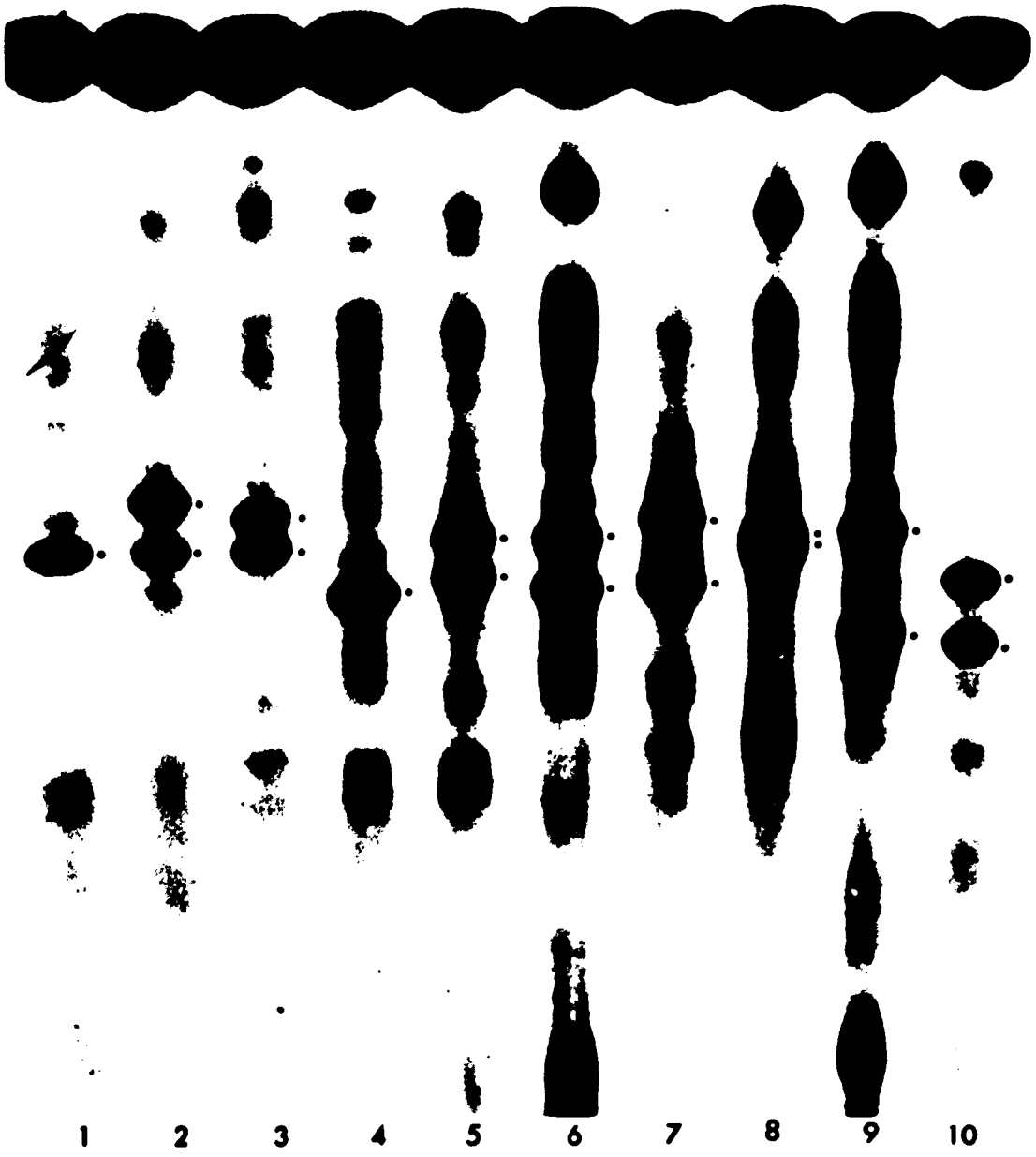
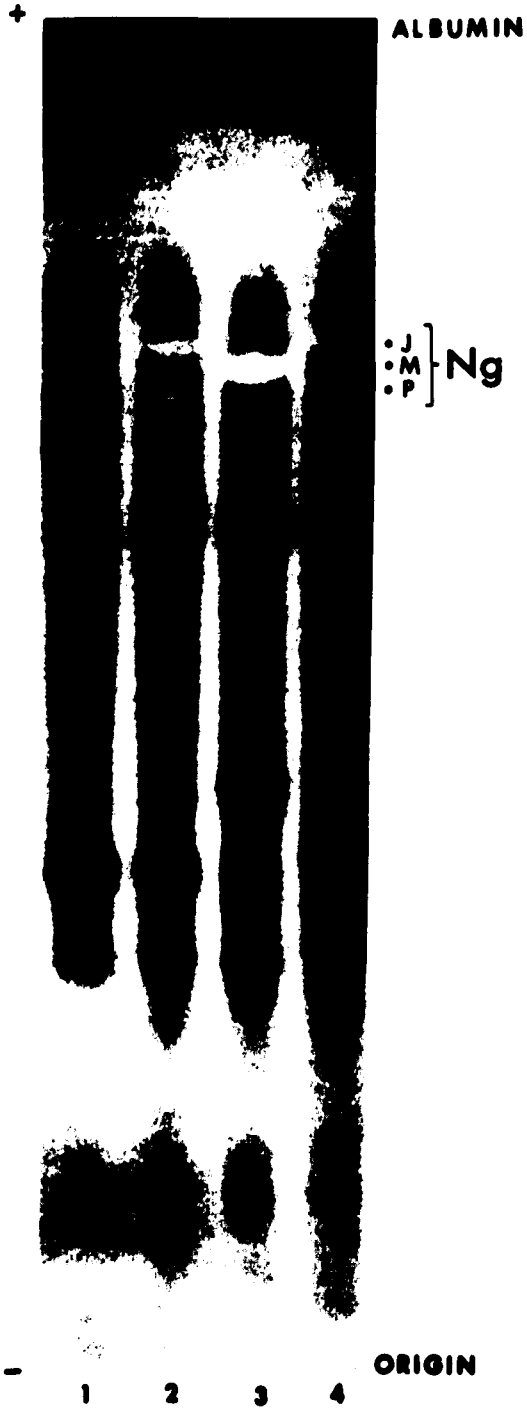

1
2
3
4
5
6
7
8
9
10

Figure 3: Photograph of the white bands of the *Mg* locus.
1. *MgMP;* 2. *MgJP;* 3. *MgM* and 4. *MgJ*.


+
ALBUMIN
J
M
P
Ng
–
ORIGIN
1
2
3
4


+
ALBUMIN
J
M
P
Ng
ORIGIN
1
2
3
4

well defined bands. Other esters, such as β-naphthyl acetate, can be used as the substrate but the bands are less easily seen because of lack of contrast.

Blood Typing

Cells from each animal were tested for agglutination in anti-A, anti-B, and anti-C antisera using the Chown capillary tube technique (Dunsgord and Bowley, 1967). Frozen cells were typed without resuspension in an aqueous solution, using an antiserum-glycerol mixture. The serology and genetics of the *Pm* locus are discussed in detail by Savage (1971).

Laboratory Breeding of Deer Mice

During 1968 and 1969, approximately 700 wild-trapped mice were paired without regard to phenotype (except for the inclusion in the breeding program of mice with rare phenotypes). These mice were found to breed readily in captivity. However, they did not breed continuously; rather, they produced up to five litters per pair during their natural breeding season. Light conditions in the animal quarters closely matched the natural photoperiod.

During the spring of 1970, a series of crosses was made, generally between opposite homozygotes at each locus, for the production of heterozygotes with defined interlocus configurations. These F_1 individuals were then crossed to wild trapped homozygotes (or different heterozygotes) in order to detect linkage. By gradually increasing the photoperiod in a manner which simulated spring light conditions, these crosses were caused to be productive during midwinter (1970-71). During midwinter (1970-71), the photoperiod was increased in a manner which simulated spring light conditions and these crosses became productive.

RESULTS AND DISCUSSION

Results of Trapping

A summary of the results of trapping is presented in Table 1. In all cases where trap success is given, it has been corrected by subtracting half the number of individuals of other species in the particular sample from the number of trap-nights. (A wide variety of small mammals other than *P. maniculatus* were captured, including representatives of the following genera: *Clethrionomys, Microtus, Mus, Eutamias, Sorex, Zapus, Mustela, Thomomys, Glaucomys*). This will correct partially for the effect of traps being filled by species other than the one(s) for which trap success is being estimated. The assumption is made that, on the average, half of the traps that will be occupied are occupied when any specific animal encounters a trap.

Although the total number of deer mice captured remained about the same during 1968 to 1970, there was a higher average trap success in 1969 than in 1968, and then a drop in trap success in 1970. The large numbers of *P. maniculatus* during 1969 were matched by even more dramatic increases in the numbers of other species, particularly the microtines *(Clethrionomys gapperi* and *Microtus pennsylvanicus), Sorex cinereus,* and *Zapus hudsonius.*

Tf Locus

Identification

Serum transferrin is commonly used as a genetic marker in studies of vertebrate populations, both because the phenotypes are easily distinguished on starch gels, and because it is commonly polymorphic (Manwell and Baker, 1970). The transferrin polymorphism in deer mice was initially described by Rasmussen and Koehn (1966), who identified the bands on starch gels using a

Table 1. The number of trap-nights, animals captured, and trap success in the study areas during the period 1968-1971.

Year	Area	Trap Nights	Deer Mice Caught*	Other Species Caught
1968	IL2	970	70 (7.5)	81 (8.7)
	IL3	410	68 (17.1)	23 (6.1)
	SR1	1654	371 (23.2)	65 (4.4)
	SR2	762	106 (14.3)	37 (5.2)
	RR	837	93 (11.6)	70 (8.9)
	IC	1047	132 (13.9)	196 (20.0)
	SL	380	34 (9.2)	21 (5.8)
	FMA	400	18 (4.8)	55 (14.1)
	SPK	420	11 (2.8)	52 (12.5)
	NO	--		
	NT	--		
Totals		6880	903 (13.7)	600 (9.3)
1969	IL2	600	60 (11.8)	185 (32.5)
	IL3	600	101 (17.8)	72 (13.1)
	SR1	765	253 (35.4)	101 (15.8)
	SR2	365	118 (36.1)	76 (24.8)
	RR	535	133 (25.8)	39 (8.3)
	IC	600	100 (19.0)	147 (26.7)
	SL	160	46 (30.2)	15 (10.9)
	FMA	165	16 (10.7)	32 (20.4)
	SPK	330	40 (13.6)	61 (19.7)
	NO	--		
	NT	--		
Totals		4120	867 (23.1)	728 (19.7)

(Cont'd)

Table 1. (Cont'd)

Year	Area	Trap Nights	Deer Mice Caught*	Other Species Caught
1970	IL2	600	47 (8.1)	34 (5.9)
	IL3	600	41 (6.9)	6 (1.0)
	SR1	600	199 (33.4)	7 (1.4)
	SR2	510	84 (16.9)	28 (6.0)
	RR	462	152 (33.5)	15 (3.9)
	IC	?	21	?
	SL	--		
	FMA	--		
	SPK	--		
	NO	600	75 (12.9)	43 (7.6)
	NT	600	127 (21.7)	30 (5.6)
Totals		4172	746 (17.7)	163 (4.3)
1971	IL2	--		
	IL3	--		
	SR1	--		
	SR2	420	22 (5.4)	20 (4.9)
	RR	--		
	IC	--		
	SL	210	35 (17.2)	12 (6.2)
	FMA	--		
	SPK	--		
	NO	--		
	NT	--		
Totals		630	57 (9.3)	32 (5.3)

Total number of trap nights - 15,172
Total number of deer mice captured - 2,573 (17.86)
Total number of other species captured - 1,523 (10.97)

*In every case, the number in brackets is the corrected trap success, in per cent.

stain which detects ferric iron bound to the transferrin. The transferrin bands described in this study were identified as such by Canham (1969), using the same staining technique. In addition, R.K. Selander and co-workers (pers. comm.) have identified deer mouse transferrin using rivanol precipitation, a technique which precipitates from serum all proteins which are visible on starch gels following staining with amidoblack, except transferrin.

Figure 2 shows the nine different transferrin bands which were found during this study. Following the nomenclature of Canham (1969), the locus is named *Tf* and the bands are given the letters G, H, I, J, K, L, M, P, and Q, in order of decreasing mobility.

The typing of the transferrin phenotypes was remarkably consistent. Sera did not change detectably in storage for up to four years; nor were there any detectable changes in the transferrins of particular animals, when serum samples were taken and typed at intervals of up to three years. However, it was discovered during the last few months of this study that minor variations in electrophoretic technique can affect the degree of separation of *TfI* from *TfJ*. Similarly, the discovery of *TfK*, and the detection of the difference between *TfP* and *TfQ* were probably because of increased expertise and better equipment acquired during the last two years of the study. Confirmation of the accuracy of detection of the transferrin phenotypes was obtained by comparing the results of typing approximately 300 mice with those independently obtained with the same mice in Dr. R.K. Selander's laboratory.

Inheritance of Transferrin Phenotypes

Mode of inheritance. The forms of transferrin have always been found to be inherited as codominant alleles of a single autosomal locus (Lush, 1966; Malecha and Tamarin, 1969). Deer mouse transferrins are no exception (Rasmussen and Koehn, 1966; Canham, 1969; Canham, Birdsall, and Cameron, 1970).

Table 2 gives the results of crosses which confirm this mode of inheritance. (The results of Canham [1969] are included in this table.) In no case was a transferrin band found in the sera from offspring of a cross that was not present in the sera from the parents. Sera from all individuals contained either one or two forms of transferrin. Although segregation data were not obtained for two of the nine forms of transferrin (*TfK* and *TfP*), it is assumed that these are controlled by alleles of the same gene.

Except for two crosses, the distribution of phenotypes in the offspring of these crosses is in ageeement with that expected according to the hypothesis of codominant alleles at a single autosomal locus. The cross TfJM x TfIJ produced a significant deficiency of the two classes receiving the *TfI* allele. This is probably attributable to the difficulty of distinguishing between *TfI* and *TfJ*. The significant ($p<0.001$) deviation from the 1:2:1 ratio of TfJ:TfJM:TfM expected from the cross of two TfJM heterozygotes cannot be so easily discounted.

Segregation distortion at the Tf *locus*. The significant deviation from the expected ratio in the offspring of the TfJM x TfJM cross was largely attributable to an excess of TfM individuals. The ratio of *TfJ* to *TfM* alleles in the offspring differed significantly ($p<0.005$) from 1:1. These deviations from expected frequencies were not restricted to a small proportion of highly aberrant litters. The phenomenon was first detected in litters from parents bred during 1966-1968, most of which were captured either in the Northwest Territories, or in the Edmonton area (Canham, 1969). The parents of the large number of 1969 crosses were captured at several locations in Colorado, Montana, Alberta, and the Northwest Territories. The excess of TfM young occurred in a majority of litters from parents captured in these widely separated locations during different years. A small number of TfJM parents,

Table 2. The results of crosses showing the inheritance of seven of the nine transferrin bands.

Phenotypes of Parents ♀	♂	Number of Offspring of the Phenotype: GJ	GM	HJ	HM	LJ	IM	J	JM	JL	M	JQ	MQ	Number of Litters	N*	Mean Litter Size	χ^2	Probability less than	Probability greater than
J	J							407						79	407	5.2	--		
J	M								225					48	225	4.7	--		
M	J								95					19	95	5.0	--		
M	M										111			23	111	4.8	--		
GJ	J	11						15						5	26	5.2	0.62	.50	.30
GM	M		1								4			1	5	5.0	**		
HM	M				8						6			3	14	4.7	0.29	.70	.50
J	IM					10			8					4	18	4.5	0.22	.70	.50
LJ	J							1						1	1	1.0	**		
J	JL							3		4				1	7	7.0	**		
J	JM							167	159					64	326	5.1	0.20	.70	.50
JM	J							166	171					75	337	4.5	0.07	.80	.70
JM	M								78		73			32	151	4.7	0.17	.70	.50
M	JM								122		119			50	241	4.8	0.04	.90	.80
MQ	J								5			6		2	11	5.5	0.09	.80	.70
M	JQ								6				6	3	12	4.0	0		
JM	JM							68	134		106			65	308	4.7	14.57	.001	
GM	JM	1	6						8		3			4	18	4.5	6.44	.10	.05
HM	JM			4	1				3		3			2	11	5.5	1.73	.70	.50
JM	LJ					3	6	15	13					8	37	4.6	10.46	.02	.01
MQ	JM								1		2	4	1	2	8	4.0	**		
Totals		12	7	4	9	13	6	842	1028	4	427	10	7	491	2369	4.82			

*Total number of offspring
**Sample size too small for meaningful calculations

themselves born in captivity, also produced an excess of TfM young.

The cause of the observed deviation from the expected phenotype frequencies is not immediately obvious. However, some possibilities can be eliminated. Any form of selective gametogenesis can be ruled out, because no other type of mating involving TfJM individuals showed distorted ratios.

The mean size of litters from TfJM x TfJM crosses was similar to that from other types of matings. Canham, Birdsall and Cameron (1970) report that a count of corpora lutea from 112 pregnancies in deer mice indicated that a mean of 6.70 ova were released at oestrus. The number of selectively eliminated TfJ and TfJM individuals which would account for the observed deviation from a 1:2:1 ratio is 116, or 1.78 per litter. The mean number of ova released per pregnancy is just slightly more than the sum of the mean litter size (4.7) and the calculated number of selectively eliminated individuals per litter. This would not allow for the occurrence of random prenatal death, and unfertilized ova. Of course, if the proportion of ova which remain unfertilized is extremely small, it is possible (but improbable) that the normally random lethalities in other crosses do not occur in the TfJM x TfJM cross, and in that cross the difference between the mean number of ova released and the mean litter size is largely attributable to selective elimination of TfJ and TfJM individuals. So, unless the number of ova released per pregnancy is an underestimation, it is unlikely that selection against zygotes, embryos, foetuses, or young up to the time of weaning is the explanation.

Because the number of TfJM young from the TfJM x TfJM cross differed significantly from half of the total ($p<0.05$), selective transport of sperm can be ruled out as a possible cause. The remaining reproductive stage at which irregularities could cause the distorted ratios is fertilization.

Bateman (1960) has analyzed the various forms of selective fertilization. From that analysis, the only explanation which accounts for the observations, is that TfM sperm are able to fertilize a majority of TfM ova in the reproductive tract of heterozygous (TfJM) females.

Ng Locus

Discovery and Characteristics

The discovery of white bands on starch gels stained for serum esterase activity (Figure 3) occurred because of an accidental overstaining. A wide range of experimentation with electrophoretic and staining variations was necessary in order to standardize the technique. An account of this technique and some preliminary breeding data has been published (Birdsall, Redfield, and Cameron, 1970).

The nature of the substance causing the appearance of the white bands is unknown. Brewer (1967) has reported inherited electrophoretic variation in achromatic regions on tetrazolium-stained starch gels. There seems to be little relationship between the white bands on gels stained for esterase and the achromatic regions on tetrazolium-stained gels. The bands on gels stained for esterase do not appear when the staining procedure used by Brewer is applied. In fact, direct observation of the achromatic regions on tetrazolium-stained gels resulting from electrophoresis of tissue extracts from mice used in this study revealed that their phenotypes were quite different than their previously described "white band" phenotypes determined from gels stained for esterase. (Gels were provided for comparison by Dr. R.K. Selander.) Unlike the regions described by Brewer, the formation of white bands on gels stained for esterase is not inhibited by cyanide (10^{-3}M). Whereas α-naphthyl acetate is necessary for clear visualization of the bands on gels stained for esterase, the achromatic regions on

tetrazolium-stained gels are not dependent upon the presence of coenzymes or substrates.

Numerous hypotheses can be constructed to account for the bands. The most likely of these are: (a) a change in the gel which alters the normal deposition or retention of the naphthol-fast blue RR salt complex; (b) a breakdown of the complex, inhibition of its formation, or any change which would render it soluble in the wash solution; (c) the presence of a substance which inhibits esterase activity; or (d) the formation of a different substance which appears as a white precipitate in the gel.

Because of the appearance of the bands, and other reasons (such as the effect of varying substrate concentration), I would suggest that (b) is most probably correct. However, no systematic biochemical analysis has been performed, and definition of the chemical origin of the bands and the biological role of the gene products involved must await that analysis.

Inheritance of the *Ng* Locus Phenotypes

Table 3 shows the results of crosses demonstrating the inheritance of the white bands. These data are all in agreement with the hypothesis that the appearance of the bands is controlled by codominant alleles at a single autosomal locus. In no case was there a significant deviation from the expected Mendelian ratios. In accordance with general protein nomenclature (Braend, 1965; Ashton *et al.*, 1967), the locus has been named *Ng*. The bands found to date have been given the letters H, J, K, M, P, and R, H being fastest and R slowest in electrophoretic migration rate. In no case has serum from a deer mouse failed to produce the bands, nor has any individual's serum produced more than two bands. No bands have been produced by sera from offspring that were not present in the sera from the parents.

Table 3. The results of crosses showing the inheritance of bands J, M, P, and H; *Hg* locus.

Phenotypes of Parents ♀	♂	Number of Offspring of the Phenotype: J	JM	M	MP	P	JP	HM	HP	Number of Pairs	Number of Litters	N*	Mean Litter Size	χ^2	Probability less than	Probability greater than
J	M		18							3	4	18	4.5	--		
M	J		13							2	4	13	3.3	--		
J	P						20			2	4	20	5.0	--		
M	P				56					5	10	56	5.6	--		
P	M				20					2	4	20	5.0	--		
M	M			254						25	53	254	4.8	--		
P	P					18				3	4	18	4.5	--		
J	JM	5	11							3	4	16	4.0	2.25	.20	.10
JM	J	23	17							4	8	40	5.0	0.90	.50	.30
J	JP	5					7			3	3	12	4.0	0.30	.70	.50
J	MP		7				9			3	5	16	3.2	0.25	.70	.50
MP	J		4				4			1	2	8	4.0	0		
M	MP			129	124					29	51	253	5.0	0.10	.80	.70
MP	M			65	50					14	24	115	4.8	1.95	.20	.10
M	JP		25		34					8	13	59	4.5	1.37	.30	.20
JP	M		27		20					6	11	47	4.3	1.04	.50	.30
M	JM		73	66						16	26	139	5.3	0.35	.70	.50
JM	M		67	70						18	31	137	4.4	0.07	.80	.70
P	JM				6		7			2	3	13	4.3	0.08	.80	.70
JM	P				10		13			3	5	23	4.6	0.48	.50	.30
P	JP					7	8			2	3	15	5.0	0.07	.80	.70
JP	P					8	8			2	3	16	5.3	0		
P	MP				17	18				5	7	35	5.0	0.03	.90	.80
MP	P				20	20				6	9	40	4.4	0		
M	HM			1				4		1	1	5	5.0	**		
HM	P				6				8	1	3	14	4.7	0.29	.70	.50
MP	M				8			12		1	4	20	5.0	0.80	.50	.30
JM	JM	12	23	10						7	10	45	4.5	0.07	.98	.95
JP	JP	2					1			1	1	3	3.0	**		
MP	MP			51	90	33				21	39	174	4.5	3.93	.30	.20
JM	MP		19	18	17		11			8	13	65	5.0	2.38	.50	.30
MP	JM		27	25	29		34			11	23	115	5.0	1.56	.70	.50
JM	JP	16	20		16		13			5	13	65	5.0	1.52	.70	.50
JP	JM	3	4		6		3			1	3	16	5.3	1.25	.80	.70
MP	JP		9		16	17	10			6	11	52	4.7	3.85	.30	.20
JP	MP		4		6	6	7			4	6	23	3.8	0.83	.90	.80
HM	MP			2	1			5	2	1	2	10	5.0	**		
Totals		66	368	691	552	127	155	21	10	235	420	1990	4.74			

*Total number of offspring
**Sample size too small for meaningful calculations

Pm Locus

History

The initial genetical investigation of antigens on red blood cells of deer mice was carried out by Rasmussen (1961, 1964). He detected two antigenic types which were inherited as antithetical alleles of a single autosomal locus, the *Pm* locus. He named the alleles *PmA* and *PmB* and the corresponding antigens A and B. By testing for agglutination of red cells by anti-A and anti-B antisera, he described the phenotype of each mouse as being A, AB, or B. Savage (1971) has expanded the genetics and serology of the *Pm* locus using mice which were trapped during this study. She discovered that the *Pm* locus is complex, with at least four, and possibly six alleles determining the presence of antigens A, B, C, and X. The X antigen is not an alternate antigen to A, B, and C, so that the six phenotypes A, AB, B, BC, C, and AC can be described by testing cells for agglutination in anti-A, anti-B, and anti-C antisera.

Inheritance of the *Pm* Locus Antigens

All serological observations were consistent with the hypothesis that the reagents anti-A, anti-B, and anti-C were detecting a blood group system which was under the control of three codominant alleles of one autosomal locus (Savage, 1971). In addition, the results of crosses made to define the inheritance of the three antigens A, B, and C (Table 4) were consistent with that hypothesis.[1]

Segregation Distortion

In general, the distributions of phenotypes in the offspring of crosses

---

[1]A portion of the data included in Tables 4 and 5 has been reported by Savage and Cameron (1971) and Savage (1971). Because the most complete interpretation of the population data requires an understanding of the inheritance and characteristics of the genetic systems used, these have been repeated here.

Table 4. The results of crosses showing the inheritance of antigens A, B, and C; controlled by alleles of the *Rw* locus.

Phenotypes of Parents ♀	♂	Number of Offspring of the Phenotype: A	AB	B	BC	C	AC	Number of Pairs	Number of Litters	N	Mean Litter Size	χ^2	Probability less than	greater than
A	A	155						19	34	155	4.6	--		
A	B*		199					23	44	199	4.3	--		
B	B			8				1	2	8	4.0	--		
B	C*				3			1	1	3	3.0	--		
A	C*						57	4	10	57	5.7	--		
A	AB	186	179					57	77	365	4.7	0.13	.80	.70
AB	A	190	159					34	78	349	4.5	2.75	.10	.05
A	BC		30				26	5	12	56	4.7	0.29	.70	.50
BC	A		34				28	4	12	62	5.2	0.52	.50	.30
A	AC	18					24	5	8	42	5.3	0.86	.50	.30
AC	A	26					33	7	13	59	4.5	0.83	.50	.30
B	AB		45	34				6	14	79	5.6	1.53	.30	.20
AB	B		44	25				7	15	69	4.6	5.23	.05	.02
B	BC			8	2			1	2	10	5.0	**		
B	AC		15		15			3	6	30	5.0	0		
AC	B		4		6			2	3	10	3.3	**		
C	AB				2		2	1	1	4	4.0	**		
AB	C				14		14	1	4	28	7.0	0		
C	BC				6	8		1	4	14	3.5	0.29	.70	.50
C	AC					3	5	1	2	8	4.0	**		
AB	AB	33	68	29				14	25	130	5.2	0.52	.80	.70
AC	AC	6				10	26	5	7	42	6.0	3.36	.20	.10
AC	AB	36	29		28		30	10	27	123	4.6	1.26	.80	.70
AB	AC	21	13		17		26	7	17	77	4.5	4.82	.20	.10
BC	AB		28	13	16		16	5	13	73	5.6	7.27	.10	.05
AB	BC		9	13	9		10	5	8	41	5.1	1.05	.80	.70
AC	BC		5		6	3	10	2	5	24	4.8	4.33	.30	.20
Totals		671	861	130	124	24	307	231	444	2117	4.77			

*Reciprocal crosses combined
**Sample size too small for meaningful calculations

Table 5. Segregation ratios from heterozygotes.

Phenotype of Mate	Phenotype of Heterozygote											
	Female AB		Male AB		Female BC		Male BC		Female AC		Male AC	
	A	B	A	B	B	C	B	C	A	C	A	C
A	190	159	186	179	34	28	30	26	26	33	18	24
B	44	25	45	34			8	2	4	6	15	15
C	14	14	2	2			6	8			3	5
AB					41	32			65	58	34	43
BC	19	22	44	29			22	19	15	9		
AC	47	30	66	57			11	13				
Totals	314	250	343	301	75	60	77	68	110	106	70	87
χ^2	7.26		2.74		1.67		0.56		0.07		1.84	
P	<.01		<.10		<.20		<.50		<.80		<.20	

involving AB females have lower p values than their reciprocal crosses. A comparative analysis of the transmission of alleles by heterozygotes (Table 5) reveals that PmAB females transmit significantly more *PmA* than *PmB* alleles to their offspring. The phenomenon was not restricted to a small proportion of PmAB females producing highly aberrant segregation ratios. In the initial series of crosses, nearly all PmAB females transmitted more *PmA* than *PmB* alleles to their offspring. The magnitude of the discrepancy was less in subsequent crosses; however, again more offspring received the *PmA* allele.

Various factors could cause this aberrant segregation ratio. These have been discussed in detail by Savage (1971), who concluded that the distortion was not attributable to either mistyping of offspring with non-specific antisera, or post-implantation selection. The segregation distortion can thus be attributed to differential production of ova, differential fertilization, or differential implantation.

Results of Crosses Made to Detect Linkage

Table 6 shows the results of crosses of F_1 heterozygotes with homozygotes or different heterozygotes in order to detect linkage. Because of the wide variety of configurations of alleles among the F_1, the offspring have simply been classed as having received a parental or nonparental pair of alleles with respect to each pair of loci. No significant differences from a 1:1 ratio of parental to nonparental types were found, so it is assumed that the three loci are part of different linkage groups.

Allele and Phenotype Frequencies in the Populations

The *Tf* Locus

Nine alleles of the *Tf* locus were found in the natural populations. In every population, however, the common alleles were *TfJ* and *TfM*. No other alleles were found in a frequency greater than 0.05, nor were any homozygotes

Table 6. Results of crosses made to detect linkage between the *Tf*, *Pm*, and *Ng* loci.

Loci	Number of Offspring		Number of Pairs	Number of Litters	χ^{2}*	Probability	
	Parental	Non-parental				Less than	Greater than
Tf - *Pm*	141	132	36	61	0.30	.60	.50
Tf - *Ng*	84	67	14	30	1.91	.20	.10
Pm - *Ng*	64	82	13	29	2.22	.10	.05

*The χ^{2} values result from tests for goodness of fit to a 1:1 ratio.

other than *TfJ* and *TfM* found. In order to avoid the complication of considering rare phenotypes separately, for the following analyses all alleles other than *TfJ* and *TfM* will be considered to be the same as the most common allele in each particular sample.

The number of animals with each transferrin phenotype in the population samples is given in Table 7. The frequency of *TfJ* (p) is given for each sample.

Each sample of greater than 20 animals has been tested for goodness of fit to the Hardy-Weinberg distribution, using the exact method of Levene (Dobzhansky and Levene, 1948) to calculate the expected number of each phenotype. This method gives the expected values

$$\frac{2pN(2pN-1)}{4N-2}, \quad \frac{4pqN^2}{2N-1}, \text{ and } \frac{2qN(2qN-1)}{4N-2}$$

for one homozygote, the heterozygote, and the other homozygote respectively, where p is the frequency of one allele, q is the frequency of the other, and N is the total number of individuals in the sample. The use of the exact method of Levene negates a bias toward heterozygote excess which is inherent in the application of the simple binomial expansion to small samples. The null hypothesis that the phenotypes in the population are distributed according to the Hardy-Weinberg distribution is rejected when the χ^2 value (one degree of freedom) has a significantly small probability of occurring. Two samples were found to deviate significantly. The sample from IL2 (1969) included a significant excess of heterozygotes, and the sample from RR (1970) showed a significant deficiency of heterozygotes. Although the inclusion of animals from more than one generation violates a condition of Hardy-Weinberg equilibrium, the results of tests for goodness of fit to Hardy-Weinberg proportions are given here for the total samples, as a rough approximation and to show the generally close fit of observed to expected.

Table 7. The distribution of the transferrin phenotypes in the natural population samples. p is the frequency of *TfJ*. The χ^2 values for goodness of fit to Hardy-Weinberg proportions are given.

Year	Area	N*	Phenotypes: J	JM	M	p	χ^2	Probability Less than	Greater than
1968	IL2	60	13	30	17	.47	0.00		.90
1969	IL2	59	7	37	15	.43	4.28	.05	.02
1970	IL2	44	10	17	17	.42	2.11	.20	.10
1968	IL3	59	19	29	11	.57	0.00		.90
1969	IL3	98	25	56	17	.54	2.07	.20	.10
1970	IL3	40	14	18	8	.58	0.34	.70	.50
1968	SR1	305	94	147	64	.55	0.24	.70	.50
1969	SR1	223	77	101	45	.57	1.34	.30	.20
1970	SR1	199	71	96	32	.60	0.00		.90
1968	SR2	100	40	43	17	.62	0.94	.50	.30
1969	SR2	106	43	42	21	.60	3.31	.10	.05
1970	SR2	83	24	38	21	.52	0.66	.50	.30
1971	SR2	17	3	10	4	.47	**		
1968	RR	82	11	37	34	.36	0.06	.90	.80
1969	RR	120	21	61	38	.43	0.13	.80	.70
1970	RR	133	35	53	45	.46	5.44	.02	.01
1968	IC	121	64	53	4	.75	3.02	.10	.05
1969	IC	96	46	42	8	.70	0.10	.80	.70
1970	IC	21	5	10	6	.48	0.10	.80	.70
1968	SL	43	12	23	8	.55	0.20	.70	.50
1969	SL	45	21	20	4	.69	0.03	.90	.80
1971	SL	31	6	19	6	.50	1.36	.30	.20
1968	FMA	17	3	7	7	.38	**		
1969	FMA	15	0	8	7	.27	**		
1968	SPK	11	4	5	2	.59	**		
1969	SPK	40	6	23	11	.44	0.97	.50	.30
1970	NO	70	27	32	11	.61	0.13	.80	.70
1970	NT	128	33	73	22	.54	2.70	.20	.10
					$\bar{X}$	.52			

*N is the total number of animals in the sample.
**Sample size too small for meaningful calculation.

For comparison of individual allele frequencies with one another, independent of phenotype frequency, a rows x columns test of independence using the G-test was performed (Sokal and Rohlf, 1969). The computations were done using program A, given in Appendix I. For each sample, the number of alleles of each type was tested. The G-value for the total sample array was 161.086, which is highly significant ($\chi^2_{.01(28)}$=48.278). No individual sample, nor any group of samples from one woodlot, when removed resulted in a homogeneous subset.

In order to visually represent the heterogeneity of the samples, they have been plotted on a de Finetti diagram (Crow and Kimura, 1970) (Figure 4). The relative proportion of each phenotype is the distance from the plot to the appropriate base, when the altitude of the triangle is unity. The bell-shaped curve represents Hardy-Weinberg proportions. Those falling above the curve exhibit a heterozygote excess, and those below, a heterozygote deficiency. The frequency of *TfJ* (p) of a particular plot can be estimated from a perpendicular to the horizontal base.

Table 8 gives the population results for the two age classifications, overwintered and juvenile, for those samples which exceed 15 individuals. These were tested for changes in allele frequency using a 2 x 2 contingency χ^2 test. There were no significant differences observed, either between generations, or between juvenile and overwintered samples of one generation.

No consistent differences were found between the sexes in allele or phenotype frequency, so in general both sexes will be considered together for subsequent calculations.

The *Pm* Locus

The number of animals of each *Pm* locus phenotype in the population samples is presented in Table 9. Each sample of greater than 20 animals has

Figure 4: A de Finetti plot of the proportions of the transferrin phenotypes in the samples from the natural populations. See text for explanation.

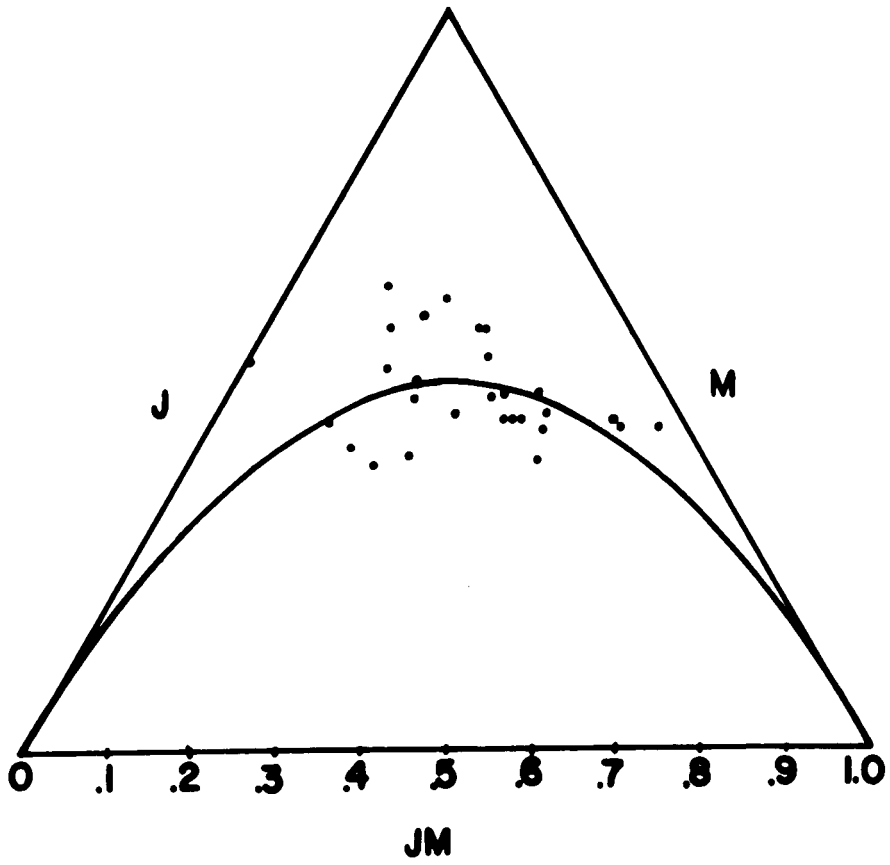

J
M
0
.1
.2
.3
.4
.5
.6
.7
.8
.9
1.0
JM

Table 8. The distribution of the transferrin phenotypes in overwintered (O) and juvenile (J) mice in all samples of greater than 15 individuals from the natural populations. p is the frequency of *Tf J*.

Area	Year	Age Class	Number of Animals of the Phenotype:			p
			J	JM	M	
IL2	1968	O	5	11	7	.46
IL2	1968	J	8	19	10	.47
IL2	1969	O	5	21	6	.48
IL2	1969	J	2	16	9	.37
IL2	1970	J	9	16	13	.45
IL3	1968	O	6	8	3	.59
IL3	1968	J	13	21	8	.56
IL3	1969	O	7	19	5	.53
IL3	1969	J	18	37	12	.55
IL3	1970	J	14	18	8	.58
SR2	1968	J	36	41	14	.62
SR2	1969	O	18	21	8	.61
SR2	1969	J	25	21	13	.60
SR2	1970	O	12	22	11	.51
SR2	1970	J	13	16	9	.55
RR	1968	O	5	11	7	.32
RR	1968	J	6	26	17	.39
RR	1969	O	6	25	14	.41
RR	1969	J	15	36	24	.44
RR	1970	O	4	10	11	.36
RR	1970	J	31	43	35	.48
SR1	1968	O	42	50	25	.57
SR1	1968	J	52	97	39	.54
SR1	1969	O	44	65	32	.54
SR1	1969	J	33	36	13	.62
SR1	1970	O	26	39	11	.60
SR1	1970	J	45	57	21	.60
IC	1968	J	58	52	4	.74
IC	1969	J	42	33	7	.71
SL	1968	O	4	15	2	.55
SL	1968	J	8	8	6	.55
SL	1969	O	12	12	2	.69
SL	1969	J	9	8	2	.68
NO	1970	O	25	27	11	.61
NT	1970	O	29	65	20	.54

Table 9. The number of animals of each *Pm* locus phenotype in the natural population samples. p, q, and r are the frequencies of *PmA*, *PmB*, and *PmC*, respectively.

Year	Area	N*	Number of Animals of the Phenotype: A	AB	B	BC	C	AC	p	q	r	$\chi^{2\dagger}$	Probability Less than	Greater than
1968	IL2	55	17	22	14	1	0	1	.52	.46	.02	1.67	.70	.50
1969	IL2	59	11	28	12	4	0	4	.46	.48	.07	0.68	.90	.80
1970	IL2	44	18	13	3	1	0	9	.66	.23	.11	2.43	.50	.30
1968	IL3	49	20	14	7	5	0	3	.58	.34	.08	5.52	.20	.10
1969	IL3	97	55	26	8	1	2	5	.73	.22	.05	18.38	.001	
1970	IL3	40	13	21	5	1	0	0	.59	.40	.01	1.97	.70	.50
1968	SR1	268	128	80	33	9	2	16	.66	.29	.06	13.79	.01	.001
1969	SR1	218	107	70	23	7	0	11	.68	.28	.04	5.80	.20	.10
1970	SR1	199	129	54	3	3	0	10	.81	.16	.03	1.58	.70	.50
1968	SR2	93	44	29	7	1	1	11	.69	.24	.08	3.10	.50	.30
1969	SR2	87	42	30	4	9	0	2	.67	.27	.06	17.52	.001	
1970	SR2	83	44	26	5	3	0	5	.72	.24	.05	1.19	.80	.70
1971	SR2	17	7	6	2	2	0	0	.59	.35	.06	**		
1968	RR	69	28	16	8	3	3	11	.60	.25	.15	7.61	.10	.05
1969	RR	118	57	35	5	7	1	13	.69	.22	.09	1.38	.80	.70
1970	RR	133	24	36	9	19	12	33	.44	.28	.29	0.96	.90	.80
1968	IC	99	26	36	10	12	1	14	.52	.34	.14	1.25	.80	.70
1969	IC	97	30	35	9	3	4	16	.57	.29	.14	6.07	.20	.10
1970	IC	21	8	12	1	0	0	0	.67	.33	0	1.45	.70	.50
1968	SL	28	11	11	3	0	0	3	.64	.30	.05	1.67	.70	.50
1969	SL	45	12	20	12	1	0	0	.49	.50	.01	1.45	.70	.50
1971	SL	31	11	9	1	8	0	2	.53	.31	.16	12.50	.01	.001
1968	FMA	12	2	4	2	0	0	4	.50	.33	.17	**		
1969	FMA	15	11	3	1	0	0	0	.83	.17	0	**		
1968	SPK	8	1	3	1	2	0	1	.38	.44	.19	**		
1969	SPK	40	7	21	9	1	0	2	.46	.50	.04	1.06	.80	.70
1970	NO	70	18	18	0	11	5	18	.52	.21	.28	4.79	.20	.10
1970	NT	28	43	45	12	11	2	15	.57	.31	.12	0.68	.90	.80
								$\bar{x}$	.60	.31	.09			

*N is the total number of animals in the sample.
**Sample size too small for meaningful calculations.
†The χ^2 values are for goodness of fit to Hardy-Weinberg proportions (3 degrees of freedom).

been tested for goodness of fit to Hardy-Weinberg proportions, which were calculated using the extension of the exact test of Levene given by Crow and Kimura (1970). Four of the 24 samples (IL3, 1969; SR1, 1968; SR2, 1969; and SL, 1971) showed significant deviations from the expected proportions. All of these four samples included less PmAB individuals than expected.

A comparison of individual allele frequencies with the G-test showed that the samples are highly heterogeneous (χ^2=392.734; $\chi^2_{.01(56)}$=83.514). In order to visually demonstrate this heterogeneity, the allele frequencies have been plotted on a triangle (Figure 5) similar to the de Finetti diagram described previously. However, when three alleles are present, allele frequencies rather than phenotype frequencies are plotted. It should be noted that the distribution of points in the de Finetti plot will be close to the line representing Hardy-Weinberg proportions, whereas the allele frequencies plotted in Figure 5 have no *a priori* expected distribution.

Table 10 shows the results from the population samples, divided into overwintered and juvenile age classes. Consecutive age classes from each woodlot were tested with a 2 x 3 contingency χ^2 test for differences in allele frequency. In contrast to the *Tf* locus, there were significant differences in five cases. Three of these cases occurred between samples from one woodlot (IL2), in which there were large fluctuations in the allele frequencies of all three alleles. The change in the frequency of *PmC* between the juveniles of 1969 and the 1970 overwintered mice in RR was very dramatic, and is of interest because it occurred within one generation, and because the juveniles of 1970 also showed the high frequency of *PmC*.

There were no consistent differences between the sexes in allele or phenotype frequency at the *Pm* locus.

Figure 5: A triangular plot of the frequencies of the alleles of the *Pm* locus in the natural populations. The distances from the sides labelled p, q and r are the frequencies of the alleles *PmA*, *PmB* and *PmC*, respectively, where the altitude of the triangle is unity.

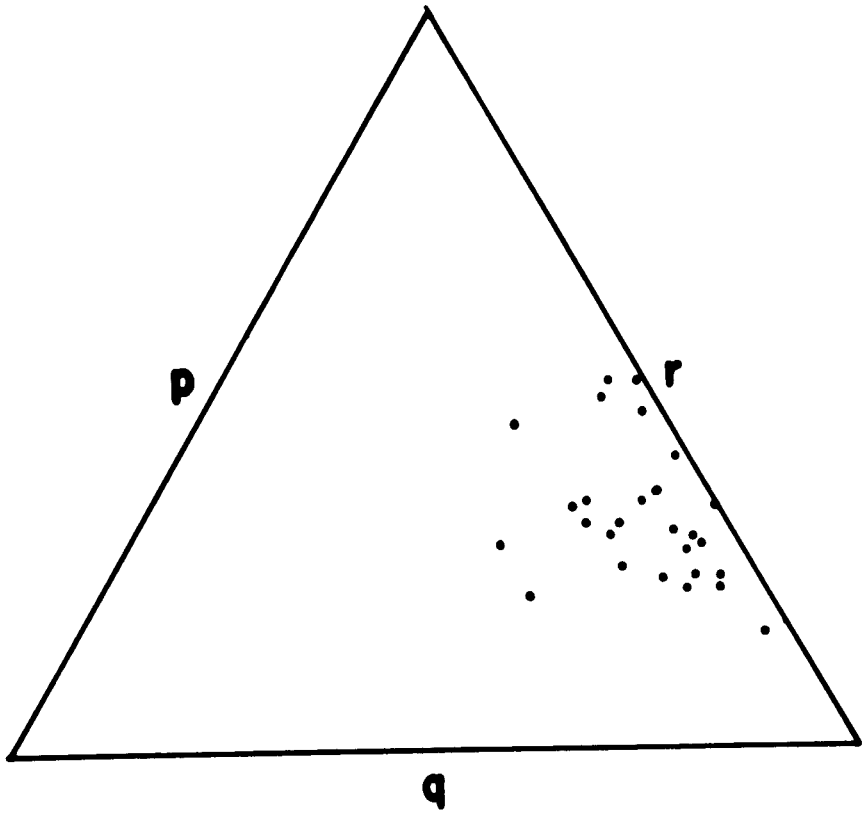

p
r
q

Table 10. The number of overwintered (O) and juvenile (J) mice of each *Pm* locus phenotype in the natural population samples. p, q, and r are the frequencies of *PmA*, *PmB*, and *PmC*, respectively. The χ^2 values are given for those age classes from particular woodlots between which there are significant differences.

Area	Year	Age Class	Number of Animals of the Phenotype: A	AB	B	BC	C	AC	p	q	r	
IL2	1968	O	7	9	2	1	0	1	.60	.35	.05	χ^2= 6.617
IL2	1968	J	10	13	12	0	0	0	.47	.53	0	p < .05
IL2	1969	O	7	14	6	2	0	3	.49	.44	.07	χ^2= 6.052
IL2	1969	J	4	14	6	2	0	1	.43	.52	.05	χ^2=14.846
IL2	1970	J	16	12	1	1	0	8	.69	.20	.11	p < .001
IL3	1968	O	5	3	3	1	0	0	.54	.42	.04	
IL3	1968	J	15	11	4	4	0	3	.60	.31	.09	
IL3	1969	O	13	12	2	1	0	3	.66	.28	.06	
IL3	1969	J	42	14	6	0	2	2	.76	.20	.04	χ^2=11.221
IL3	1970	J	13	21	5	1	0	0	.59	.40	.01	p < .01
SR2	1968	J	40	26	6	1	1	10	.69	.23	.08	
SR2	1969	O	20	19	2	3	0	2	.66	.28	.06	
SR2	1969	J	22	11	2	6	0	0	.67	.26	.07	
SR2	1970	O	28	12	1	1	0	3	.79	.17	.04	
SR2	1970	J	16	14	4	2	0	2	.63	.32	.05	
RR	1968	O	11	6	2	1	2	7	.60	.19	.21	
RR	1968	J	17	10	6	2	1	4	.60	.30	.10	
RR	1969	O	23	10	2	4	1	4	.68	.21	.11	
RR	1969	J	34	25	3	3	0	9	.69	.23	.08	χ^2=20.395
RR	1970	O	4	4	3	3	3	7	.40	.27	.33	p < .001
RR	1970	J	20	32	6	16	9	26	.45	.28	.27	
SR1	1968	O	51	35	14	3	0	2	.66	.32	.02	
SR1	1968	J	77	45	19	6	2	14	.65	.27	.08	
SR1	1969	O	67	45	14	4	0	8	.68	.28	.04	
SR1	1969	J	40	25	9	3	0	3	.68	.29	.03	
SR1	1970	O	46	22	2	1	0	5	.78	.18	.04	
SR1	1970	J	83	32	1	2	0	5	.83	.15	.02	
IC	1968	J	24	34	10	12	1	12	.51	.36	.13	
IC	1969	J	27	30	6	2	4	13	.59	.27	.14	
SL	1968	O	4	5	1	0	0	1	.64	.31	.05	
SL	1968	J	7	6	2	0	0	2	.59	.35	.06	
SL	1969	O	4	15	6	1	0	0	.44	.54	.02	
SL	1969	J	8	5	6	0	0	0	.55	.45	0	
NO	1970	O	15	18	0	10	4	16	.51	.22	.27	
NT	1970	O	36	41	12	11	2	12	.55	.33	.12	

The *Ng* Locus

Table 11 gives the number of animals of each *Ng* locus phenotype in the population samples. Six of these samples were significantly different than the expected Hardy-Weinberg distribution (IL3, 1968; SR1, 1970; RR, 1968; RR, 1969; IC, 1969; and SL, 1968).

A comparison of allele frequencies using the G-test showed that the samples are highly heterogeneous (χ^2=284.133; $\chi^2_{.01(56)}$=83.514). The allele frequencies from all the samples in Table 11 have been plotted on a triangle (Figure 6) in the same manner as has been described for the *Pm* locus.

Table 12 includes the results from the population samples, divided into overwintered and juvenile age classes. Consecutive age classes from each woodlot were tested with a 2 x 3 contingency χ^2 test for differences in allele frequency. The three cases in which there were significant differences are designated in the table. It is interesting to note that all three cases in which there were significant changes in allele frequency at the *Ng* locus also showed significant changes at the *Pm* locus.

There were no consistent differences in allele or phenotype frequency between the sexes.

Population Consequences of Segregation Distortion

The ultimate consequence of a persistent segregation distortion, in the absence of balancing selection is, of course, the elimination of alleles other than the one being favored by the distortion. Thus it is expected that, if segregation distortion is found associated with a locus which is polymorphic in a natural population, there will be gradual increases in the frequency of the favored allele (which would mean that the polymorphism is transient) or there will be selection acting to stop that increase, creating a stable polymorphism. The best example of such a polymorphism is at the *T* locus in

Table 11. The number of animals of each *Mg* locus phenotype in the natural population samples. p, q, and r are the frequencies of *MgJ*, *MgM*, and *MgP*, respectively. χ^2 is calculated for goodness of fit to Hardy-Weinberg proportions (3 d.f.).

Year	Area	N*	Number of Animals of the Phenotype: J	JM	M	MP	P	JP	p	q	r	χ^2	Probability Less than	Probability Greater than
1968	IL2	59	1	3	29	16	5	5	.09	.65	.26	6.87	.10	.05
1969	IL2	59	1	8	18	20	9	3	.11	.54	.35	1.67	.70	.50
1970	IL2	44	0	8	15	16	3	2	.11	.61	.27	1.46	.70	.50
1968	IL3	58	2	3	30	11	7	5	.10	.64	.26	15.66	.01	.001
1969	IL3	98	3	24	34	20	8	9	.20	.57	.23	3.43	.50	.30
1970	IL3	40	4	2	8	14	5	7	.21	.40	.39	7.27	.10	.05
1968	SR1	286	2	45	162	61	5	11	.11	.75	.14	1.18	.80	.70
1969	SR1	222	1	51	95	57	6	12	.15	.67	.18	4.71	.20	.10
1970	SR1	199	5	55	76	54	5	4	.17	.66	.17	10.44	.02	.01
1968	SR2	81	7	14	34	15	4	7	.22	.60	.19	7.41	.10	.05
1969	SR2	104	6	22	42	19	4	11	.22	.60	.18	3.55	.50	.30
1970	SR2	83	2	21	30	23	3	4	.18	.63	.20	1.48	.70	.50
1971	SR2	17	1	7	5	3	0	1	.30	.59	.12	**		
1968	RR	81	5	19	26	15	0	16	.28	.53	.19	11.00	.02	.01
1969	RR	120	12	37	33	14	9	15	.32	.49	.20	8.71	.05	.02
1970	RR	132	1	30	36	49	9	7	.15	.57	.28	7.51	.10	.05
1968	IC	121	10	26	27	31	7	20	.27	.46	.27	1.50	.70	.50
1969	IC	96	4	29	19	25	12	7	.23	.48	.29	8.21	.05	.02
1970	IC	21	2	4	2	6	2	5	.31	.33	.36	0.36	.95	.90
1968	SL	43	0	4	23	7	4	5	.11	.66	.23	10.44	.02	.01
1969	SL	45	3	10	11	9	5	7	.26	.46	.29	1.64	.70	.50
1971	SL	31	2	6	14	7	0	2	.19	.66	.15	2.10	.70	.50
1968	PMA	17	2	3	11	0	0	1	.24	.74	.03	**		
1969	PMA	15	3	6	1	0	2	3	.50	.27	.23	**		
1968	SPK	11	0	1	7	2	0	1	.09	.77	.14	**		
1969	SPK	40	0	12	17	9	1	1	.16	.69	.15	2.68	.50	.30
1970	NO	70	0	9	32	22	5	2	.08	.68	.24	1.16	.80	.70
1970	NT	128	2	21	49	42	9	5	.12	.63	.25	1.35	.80	.70
								$\bar{x}$	.20	.58	.22			

*N is the total number of animals in the sample
**Sample size too small for meaningful calculations

Figure 6: A triangular plot of the frequencies of the alleles of the *Ng* locus in the natural populations. See Figure 4 for explanation. p, q and r are the frequencies of *NgJ*, *NgM* and *NgP*, respectively.

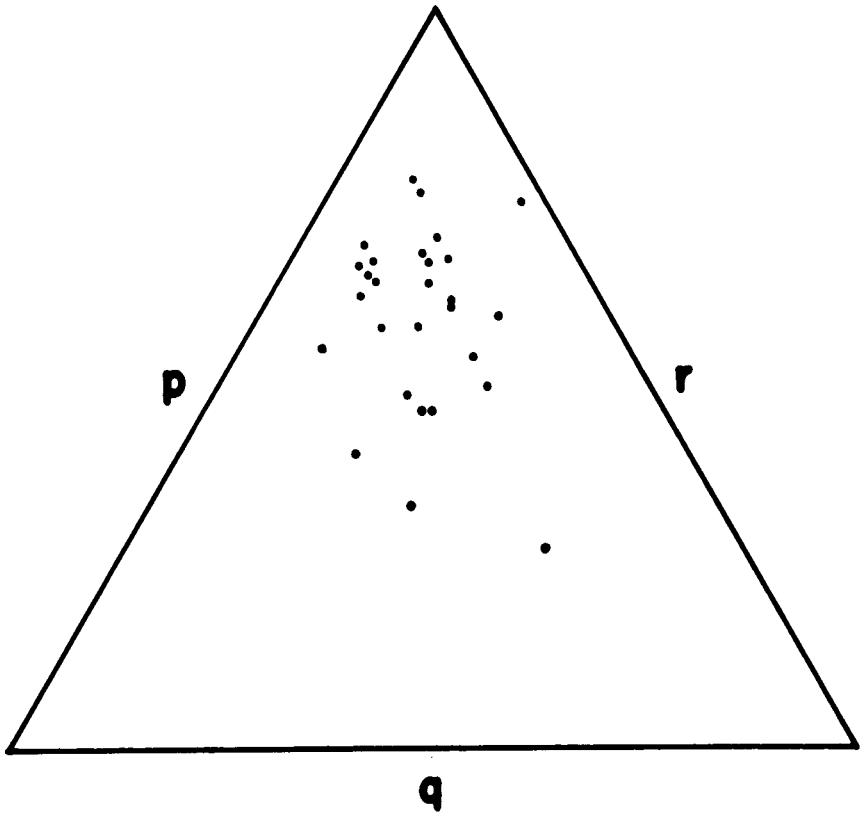

p
r
q

Table 12. The number of overwintered (O) and juvenile (J) mice of each *Hg* locus phenotype in the natural population samples. p, q and r are the frequencies of *Hg J*, *Hg N* and *Hg P*, respectively. The χ^2 values are given for those age classes from particular woodlots between which there are significant differences.

Area	Year	Age Class	Number of Animals of the Phenotype: J	JN	N	NP	P	JP	p	q	r	
IL2	1968	O	1	1	8	6	3	4	.15	.50	.35	$\chi^2 = 8.812$ $p < .02$
IL2	1968	J	0	2	21	10	2	1	.04	.75	.21	$\chi^2 = 7.385$ $p < .05$
IL2	1969	O	1	4	9	12	5	1	.11	.53	.36	
IL2	1969	J	0	4	9	8	4	2	.11	.56	.33	
IL2	1970	J	0	8	12	13	3	2	.13	.59	.28	
IL3	1968	O	0	1	6	5	2	2	.09	.56	.35	
IL3	1968	J	2	2	24	6	5	3	.11	.67	.22	
IL3	1969	O	1	9	11	7	1	2	.21	.61	.18	
IL3	1969	J	2	15	23	13	7	7	.20	.55	.25	
IL3	1970	J	4	2	8	14	5	7	.21	.40	.39	
SR2	1968	J	7	13	31	12	4	6	.23	.60	.17	
SR2	1969	O	1	10	19	8	2	6	.20	.61	.19	
SR2	1969	J	5	12	23	11	2	5	.23	.60	.17	
SR2	1970	O	2	13	13	12	2	3	.22	.57	.21	
SR2	1970	J	0	8	17	11	1	1	.12	.70	.18	
RR	1968	O	1	11	8	5	0	8	.32	.49	.19	
RR	1968	J	4	8	18	10	0	8	.25	.56	.19	
RR	1969	O	8	14	8	6	3	6	.40	.40	.20	
RR	1969	J	4	23	25	8	6	9	.27	.54	.19	$\chi^2 = 7.862$ $p < .02$
RR	1970	O	0	4	4	11	1	4	.17	.48	.35	
RR	1970	J	1	26	32	28	8	3	.16	.60	.24	
SR1	1968	O	1	21	58	21	1	5	.13	.74	.13	
SR1	1968	J	1	24	104	40	4	6	.09	.76	.15	
SR1	1969	O	0	31	62	39	2	6	.13	.69	.18	
SR1	1969	J	1	20	33	18	4	6	.17	.64	.19	
SR1	1970	O	1	19	33	22	0	1	.15	.70	.15	
SR1	1970	J	4	36	43	32	5	3	.19	.63	.18	
IC	1968	J	9	24	25	30	7	19	.27	.46	.27	
IC	1969	J	4	25	16	21	12	3	.22	.48	.30	
SL	1968	O	0	3	10	2	4	2	.12	.60	.28	
SL	1968	J	0	1	13	5	0	3	.09	.73	.18	
SL	1969	O	2	6	8	4	2	4	.27	.50	.23	
SL	1969	J	1	4	3	5	3	3	.24	.50	.26	
NO	1970	O	0	7	30	19	5	2	.07	.68	.25	
NT	1970	O	2	21	44	36	7	3	.12	.64	.24	

house mice (Lewontin and Dunn, 1960). Males which carry the recessive *t* alleles in heterozygous form (+/*t*) transmit a high proportion (up to 95%) of *t* alleles to their offspring. This is balanced by the lethality or male sterility of *t*/*t* homozygotes and, as a result, *t* alleles are common in most house mouse populations.

As has been discussed previously, two instances of segregation distortion were discovered in laboratory crosses during this study. At the *Pm* locus, heterozygous females (AB) segregate a majority (55-60%) of *PmA* alleles. Crosses between animals heterozygous (JM) at the *Tf* locus yield progenies with an excess of TfM individuals. Assuming that these distortions are persistently characteristic of the segregation of alleles at the two loci, then it is of interest to describe the selection pressures which might balance the distortions. For this purpose, a deterministic computer simulation model was used, because no suitable algebraic solutions for equilibria were available. Prout (1953) gave a solution which is correct only when the distortion modifies segregation equally in both sexes. Bruck (1957) has given solutions for the equilibrium allele frequency and the equilibrium frequency of heterozygotes for the case in which the segregation ratio is unequal in one sex only and the homozygous recessive genotype is lethal (the *T* locus system). Recently, Hartl (1970*b*) has given a series of more general solutions which will be discussed later.

The programs used for the simulations are given in Appendix I. They are written in FOCAL and were run on a DEC PDP-8. Their effect is to calculate the change in the proportions of the genotypes each generation, given appropriate values for (a) the initial value of p (the frequency of the allele favored by the distortion; q=1-p); (b) the amount of segregation distortion, K, expressed as the effective proportion of gametes which carry the allele

favored by the distortion; and (c) the values of S_1 and S_2, the selection coefficients acting on the homozygous genotype for the allele favored by the distortion and the other homozygote, respectively.

Initially, the situation at the *Pm* locus was examined. In this system heterozygotes of one sex (females) give distorted ratios no matter what the genotype of the second parent. If the allele favored by the distortion is designated *A* and the other allele is designated *a*, then the proportions of the genotypes *AA*, *Aa* and *aa*, following one generation of random mating, are given by the expressions

$$(p^2+2pqK)\ (p^2+pq)\ (1-S_1), \quad \ldots\ldots\ldots \quad [1]$$

$$(p^2+2pqK)\ (q^2+pq) + (p^2+pq)\ (q^2+2pq[1-K]) \quad \ldots\ldots\ldots \quad [2]$$

and $$(q^2+2pq[1-K])\ (q^2+pq)\ (1-S_2), \quad \ldots\ldots\ldots \quad [3]$$

respectively. The change each generation in the frequency of the allele *A* is thus easily calculated by subtracting the new value of p obtained from the expressions above (the sum of the value of expressions [1] and 1/2 of [2] divided by the sum of the values of expressions [1], [2] and [3]) from the initial value. The computer repeated these calculations until the allele frequency change was very small, indicating that an equilibrium allele frequency ($\hat{p}$) had been reached. Each value of $\hat{p}$ was calculated by setting the initial value of p arbitrarily, obtaining a value for $\hat{p}$, then re-running the program with a different starting value for p, usually such that, if, in the first simulation $\hat{p}>p$, then in the second $\hat{p}<p$, or *vice versa*. The paired values of $\hat{p}$ produced by such means were invariably identical at the level of accuracy required by the program. The program set $\hat{p}$ when the allele frequency change was less than 0.00001, which generally resulted in the two values of $\hat{p}$ being identical to three decimal places. In some cases in which the rate of approach to equilibrium was slow enough that the difference in the paired

values of $\hat{p}$ were different by greater than 0.001, the program was changed so that $\hat{p}$ was set when the allele frequency change was less than 0.000001.

The first sets of conditions tested were when K=0.6 (approximately the case at the *Pm* locus, 0.7 and 0.9, and when S_2=0, for values of S_1, in intervals of 0.05, from 0 to 1. The results of these simulations are presented graphically in Figure 7.

The following points should be noted:

(a) For the three values of K there is a wide range of values of S_1 over which a balanced polymorphism can be maintained. However, for each value of K there is a value of S_1 below which the segregation effectively overwhelms selection, and in a stochastic situation would, presumably, drive *A* to fixation (p=1.0), despite the selective disadvantage of the *AA* individuals. This value of S_1 is not small, being approximately 0.1 when K=0.6, 0.2 when K=0.7 and 0.4 when K=0.9.

(b) The intercept of each line on the right vertical axis in Figure 7 represents the *T* locus situation (S_1=1.0). Values of $\hat{p}$ at those points obtained by solution of the formula given by Bruck (1957) agree with those obtained with the computer simulation.

(c) For the situation when K=0.6, and $\hat{p}$ is equal to the mean frequency of *PmA* in the natural populations (approximately 0.6), the selection pressure against *PmA* homozygotes which would balance the segregation distortion is approximately 0.15.

A second set of conditions was tested in order to examine the consequences of heterozygote advantage. The results presented graphically in Figure 8 shows values of $\hat{p}$ when K=0.6 and S_2=0.0, 0.2 and 0.5, for values of S_1 from 0 to 1. As would be expected if there were no segregation distortion, selection against both homozygotes results in stable equilibria, the only

Figure 7: Equilibrium values of p ($\hat{p}$), with segregation distortion (K), and selection (S_1) against the homozygote favored by the distortion.

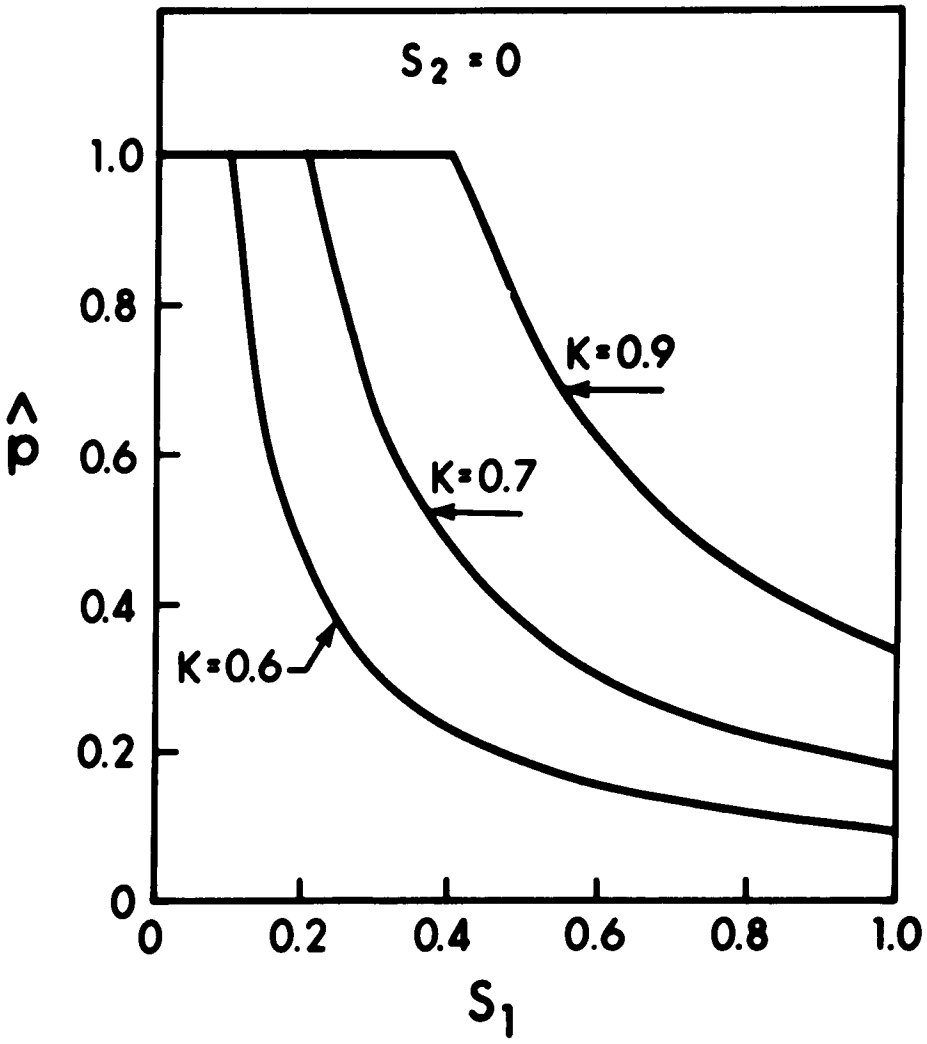
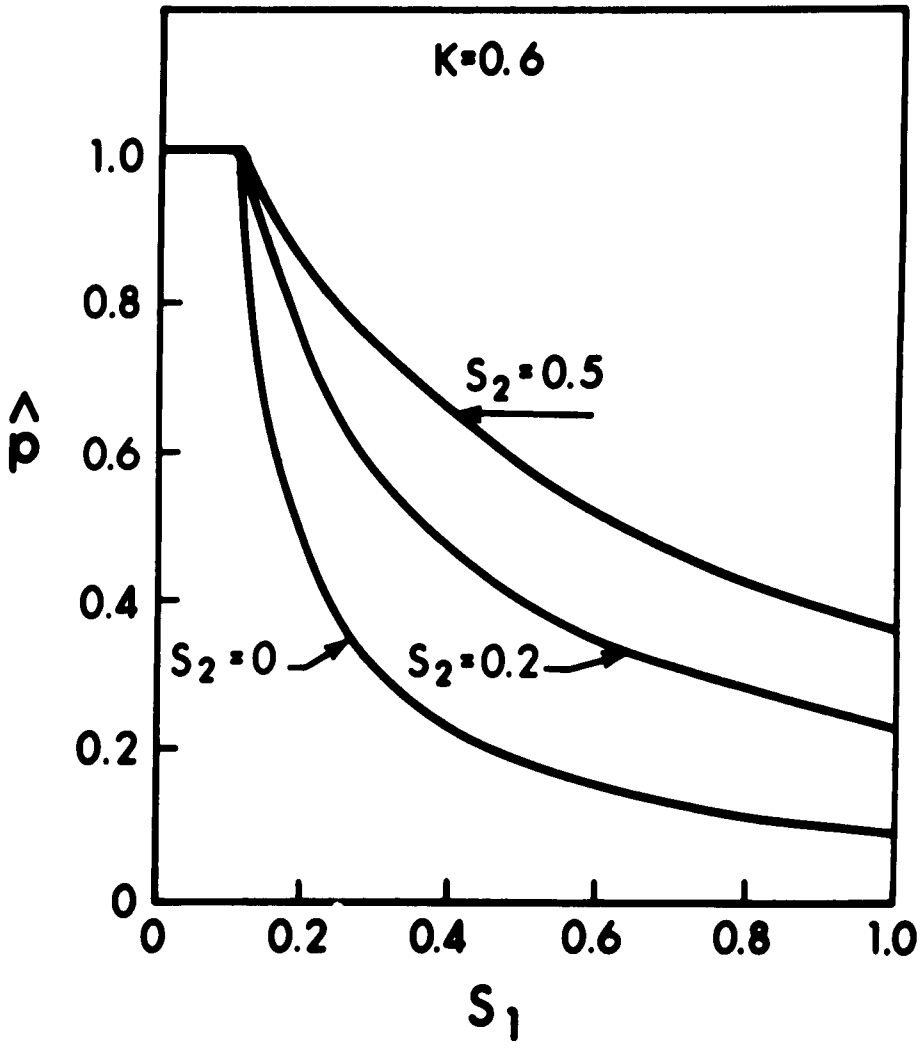

$S_2 = 0$
1.0
0.8
0.6
0.4
0.2
0
$\hat{p}$
K=0.9
K=0.7
K=0.6
0
0.2
0.4
0.6
0.8
1.0
S_1

Figure 8: The effect of heterozygote advantage on the equilibrium values of p ($\hat{p}$), given a segregation distortion (K) of 0.6. S_1 is the coefficient of selection against the homozygote favored by the distortion; S_2 is the coefficient of selection against the other homozygote.


K=0.6
1.0
0.8
0.6
0.4
0.2
0
$\hat{p}$
S_2=0.5
S_2=0
S_2=0.2
0
0.2
0.4
0.6
0.8
1.0
S_1

limiting condition being that the minimum value of S_1 which results in a stable equilibrium is unchanged by the additional type of selection.

Since these simulations were completed, Hartl (1970*b*) has given a set of algebraic solutions which describe the conditions necessary for stable equilibrium in a wide variety of cases. He points out that a stable equilibrium exists if

$$1<w_1(K+k)<2w_1-w_0,$$

where w_0 and w_1 are the fitnesses of the homozygote favored by the distortion and the heterozygote relative to the other homozygote, and K and k are the segregation ratios in the two sexes. Renormalizing Hartl's fitness values w_0 and w_1 to fit the model used here, this becomes

$$1<(K+k)/(1-S_2)<(2/(1-S_2))-((1-S_1)/(1-S_2)).$$

Since these simulations tested only cases where k=0.5, K>0.5 and $S_2>0$, the expression $(K+k)/(1-S_2)$ is always greater than one. Thus a stable equilibrium exists if

$$(K+0.5)/(1-S_2)<(1-S_1)/(1-S_2).$$

When the common denominator is removed, this reduces to simply $S_1>K-0.5$. The maximum value of S_1 which will not result in a stable equilibrium can thus be obtained by solving the formula $S_1=K-0.5$ for the appropriate values of K. For example, when K=0.6, 0.7 and 0.9, S_1=0.1, 0.2 and 0.4, which confirms the values obtained by the simulations (see Figure 7). Thus for the model used here, the conditions for stable equilibria depend on the relative values of K and S_1, independent of S_2. However, as shown in Figure 8, changes in S_2 affect $\hat{p}$.

The segregation distortion at the *Tf* locus requires a different mathematical simulation because the distortion is restricted to just the

heterozygote-heterozygote cross. The proportions of the genotypes *AA*, *Aa* and *aa*, following each generation of random mating, are given by the expressions

$$p^2(p^2+2pq+4q^2k^2), \quad \ldots\ldots\ldots \quad [4]$$

$$2pq(p^2+pq+q^2+4pqk(1-k)) \quad \ldots\ldots\ldots \quad [5]$$

and $$q^2(q^2+2pq+4p^2(1-k)^2), \quad \ldots\ldots\ldots \quad [6]$$

respectively. Figure 9 is a graph of the equilibrial conditions when k=0.56 (approximately the case at the *Tf* locus) and S_2=0, for values of S_1, in intervals 0.05, from 0 to 1. The following points should be noted:

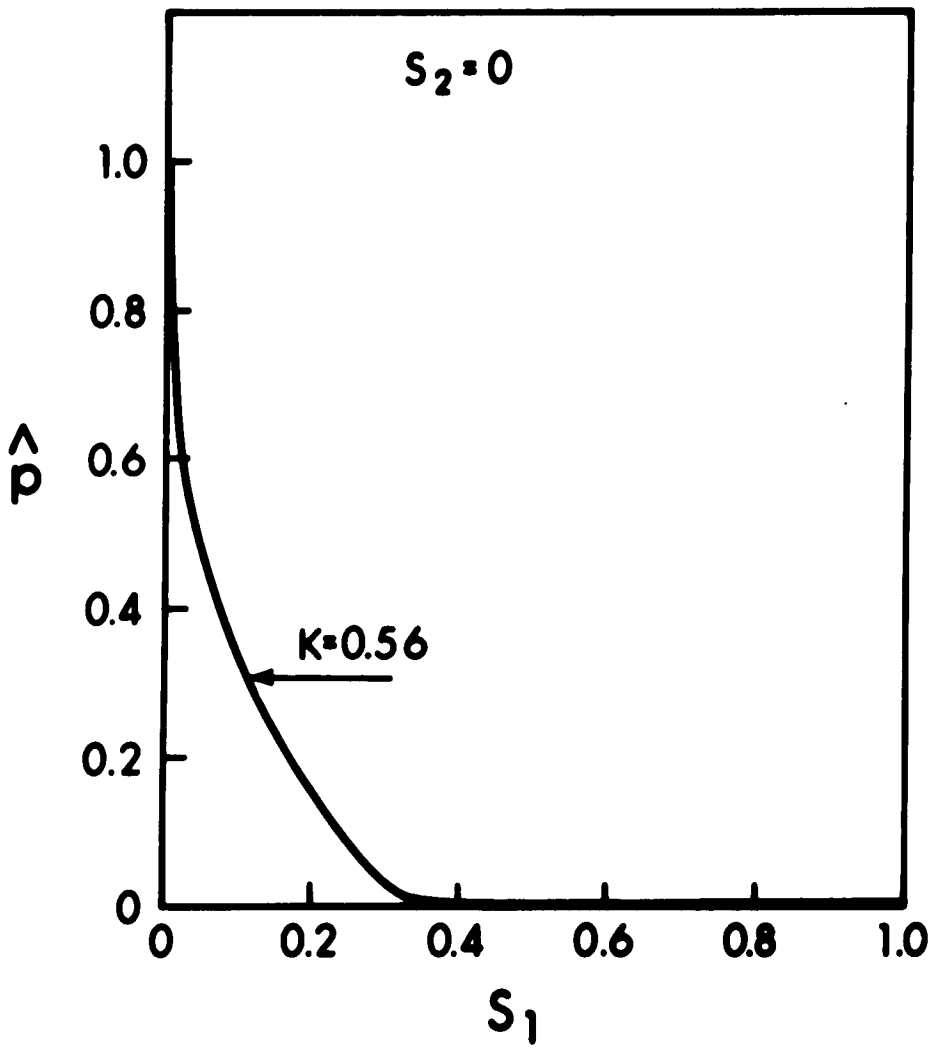
(a) If S_1 exceeds 0.3, $\hat{p}$=0.

(b) Minor variations in S_1 result in major changes in $\hat{p}$, particularly when S_1 is small ($p<0.6$).

(c) It was noted during the computer simulations that the increment of change in p each generation was small, compared with the model which describes the situation at the *Pm* locus, even when p and $\hat{p}$ were quite different. Because of the narrow range of values of S_1 which results in stable equilibrium, the large changes in p which result from small changes in S_1 and the slow rate of return to equilibrium when p is different from $\hat{p}$, it seems unlikely that the segregation distortion at the *Tf* locus could be a major factor in the maintenance of the polymorphism. However, such a distortion would modify allele frequencies and over a long term would have to be balanced or fixation would result.

No evidence has been obtained which shows that the distortions at either of the two loci occur or do not occur in natural populations, although an attempt was made to detect the results of the distortions in the offspring from females pregnant when captured (see page 71).

Figure 9: Effect of segregation distortion in the heterozygote-heterozygote cross, opposed by selection against the homozygote favored by the distortion. Labelling as in Figures 7 and 8.


$S_2 = 0$
1.0
0.8
$\hat{p}$
0.6
0.4
K=0.56
0.2
0
0
0.2
0.4
0.6
0.8
1.0
S_1

The computer simulation of the segregation distortion at the *Pm* locus has shown that if the distortion occurs in the natural populations, there must be selection against PmA homozygotes in order for alleles other than *PmA* to be maintained in the population. This selection should be detectable if the proportion of PmA homozygotes removed by the selection is greater than the proportion added by the segregation distortion. If K=0.6, the amount of selection (S_1) which would cause a stable equilibrium at $\hat{p}$=0.6 is approximately 15% (or greater, if S_2>0). The increase in PmA homozygotes each generation because of the distortion would be $(p^2+2pqK)(pq+p^2)-(p^2+pq)^2$, which is approximately 3% when p=0.6. The conclusion is that if the distortion occurs in the natural populations and if the polymorphism is stable, the results of a 12% or greater selective pressure against PmA homozygotes should be manifest in the phenotype frequencies.

Detection of Natural Selection: Some Theoretical Considerations

An examination of the data for the purpose of evaluating the theories of selectively maintained and neutral variation must begin with a search for the effects of natural selection. This may appear to reveal a preexisting bias in the approach to the conflict; however, the theory of neutral allelic variation can only be supported (in a study of this kind) by the absence of evidence for selection.

Detection of the effects of selection requires that an estimate be obtained of the genotype frequencies which existed prior to the action of the selection. This estimate can be obtained by two different methods. One is to compare allele or phenotype frequencies with those expected based on the previous generation. However, in order to detect selection which does not alter allele frequencies between generations, but only changes the

proportions of the genotypes within one generation, a mathematical model must be used which can predict the relative proportions of the genotypes when all that is known are the allele frequencies. The Hardy-Weinberg rule describes this allele-genotype relationship (Hardy, 1908; Weinberg, 1908; Castle, 1903).

Briefly, when individuals in an infinitely large, closed population mate at random, in the absence of mutation and selection, the genotype frequencies are predicted by the terms resulting from the expansion of the multinomial $(p_1+p_2...p_n)^2$, where p_1, p_2,...p_n are the allele frequencies. When genotype frequencies fit this distribution, they are commonly described as being in Hardy-Weinberg equilibrium.

The conditions which must be met for the Hardy-Weinberg relationship to apply are certainly restrictive enough to eliminate any real population. Further, it has been pointed out (Cameron, 1966; Neel *et al.*, 1964) that a good fit to the Hardy-Weinberg model does not mean that the conditions are met, but rather that factors causing deviations from the model fortuitously cancel each other. Nevertheless, the converse case, where deviations from Hardy-Weinberg proportions are detected, may provide clues to those forces which are acting.

Information is available on the potential effect of factors which affect Hardy-Weinberg proportions. Before the data can be considered with respect to selection, it is necessary to take into account those factors as they apply to the populations of deer mice studied here:

(a) *Mutation:* Because of its rarity, mutation can be eliminated as a factor which could cause any appreciable deviations from Hardy-Weinberg proportions.

(b) *Migration:* Deviations from Hardy-Weinberg proportions are caused

when animals move between populations having different allele frequencies, or when certain genotypes have higher probabilities of being migrant. Without the aid of an extensive mark-release program it is not possible to estimate the proportion of animals captured which are migrant. There is little doubt that migrant animals were occasionally captured, since some deer mice are known to be "wanderers" (Stickel, 1968). However, because the heterogeneity of allele frequencies between woodlots with respect to all three loci suggests that migration is limited, it is assumed that the effect of including occasional migrant animals in a sample is small.

(c) *Sampling error:* Any finite sample will show random deviations from exact Hardy-Weinberg proportions, the magnitude of which is a function of sample size and is thus predictable. The χ^2 test, which is used to determine the probability of random occurrence of a particular deviation, takes into account the effect of sample size.

(d) *Lack of Panmixis:* Deviations from random mating in a population can be of three general types: (i) The breeding systems in which like individuals interbreed more or less frequently than would be expected with random mating produces an excess or a deficiency of homozygotes, respectively. Because of the covert nature of the blood variants studied here, it would appear to be a reasonable assumption that assortative mating is not a factor in affecting the phenotype distributions. (ii) If animals more frequently mate with related individuals than expected (inbreeding), then alleles which are common by descent have a greater probability of being found together, causing a deficiency of heterozygotes. This heterozygote deficiency is usually quantified with the inbreeding coefficient F. (iii) If a sample includes individuals from two or more populations which are partially or completely isolated and have different allele frequencies, then there will be a deficiency,

or in some circumstances an excess, of heterozygotes. This has been called the Wahlund effect by Li (1955*a*). The Wahlund effect is a result of structuring of populations and therefore varies in magnitude depending on the type of population structure and the population size. When there are two alleles at a locus the heterozygote deficiency because of inbreeding is equivalent to the heterozygote deficiency because of population subdivision (Li, 1955*a*). This equivalence can be expressed as $\sigma^2=Fp(1-p)$, where σ^2 is the variance of the allele among the subgroups of the total population, F is the inbreeding coefficient (the correlation coefficient between the uniting gametes) and p is the average frequency of the allele among the subgroups. Li (1969) has suggested that the symbol F be used to denote the degree of association between uniting gametes as caused by inbreeding, but that the symbol ϕ be used in the future to designate the degree of association resulting from subdivision. The distinction is based on Li's demonstration that when more than two alleles of a locus are considered, no correlation coefficient relating to subdivision can be assigned to all heterozygotes because each allele can have its own variance among the subgroups. In fact, the indices of association between multiple alleles can be positive or negative depending on how the alleles are distributed among the subgroups. Calculation of a generalized value of $F+\phi$ from observed heterozygote deficiencies can be done in a variety of ways, but Li (1969) points out that there is some justification for using a simple unweighted mean of the correlation coefficients for each of the heterozygotes. This method will be used in the calculations which follow.

(e) *Allele frequency differences between the sexes:* Robertson (1965) has pointed out that if the frequencies of alleles in a population differ between the sexes, then a heterozygote excess is expected in the next generation. Because random differences in allele frequency would be expected to

occur between the sexes, independent of any larger or consistent differences generated deterministically, small heterozygote excesses would be expected to be a characteristic of finite populations. In this study, allele frequency differences between the sexes appeared to be distributed normally, so that heterozygote excesses were probably of the smaller (stochastic) type.

(f) *Selection:* Any deviation from Hardy-Weinberg proportions could be caused by natural selection. However, some characteristics of selective processes (and their effects) can be used to estimate the likelihood that selection has caused a particular deviation: (i) Selection against heterozygotes is disruptive of stable polymorphism (Li, 1955*b*). If a particular polymorphism occurs over a wide geographical range, then it would appear to be unlikely that there is heterozygote disadvantage. (ii) Selection would not be expected to have similar effects on all homozygotes or heterozygotes of a multi-allelic system. This is different from inbreeding, which causes similar decreases in the proportions of all the heterozygotes of all variable loci, and the Wahlund effect, which would usually cause general heterozygote deficiencies.

In order to detect general deviations from the expected Hardy-Weinberg proportions, the distribution of phenotypes in each pooled sample of overwintered mice (O) and juvenile mice (J) from each year has been compared with the expected. Because the calculation of Hardy-Weinberg proportions directly from the pooled samples themselves creates an artificial Wahlund effect, usually resulting in apparent heterozygote deficiencies, the expected numbers of each phenotype have been determined by calculating the expected numbers based on the samples from each woodlot and then pooling them. For each phenotype in each pooled sample, the ratio of the observed to the expected number (O/E) has been calculated to show the proportional excess or deficiency.

Thus an O/E ratio of greater than one occurs because of an excess and an O/E ratio of less than one occurs because of a deficiency. The observed number, expected number and the O/E ratio for the phenotypes of the *Tf*, *Pm* and *Ng* loci are given in Tables 13, 14 and 15, respectively. The χ^2 values are given for goodness of fit of the observed numbers to the expected.

Natural Selection and Maintenance of the Three Polymorphisms

The *Pm* locus

As described previously, if there is selection against PmA individuals which balances the segregation distortion discovered in the laboratory crosses, then there should be a deficiency of PmA individuals in the population samples. If there is no selection against PmB and PmC individuals then this deficiency can be detected by comparing the O/E ratios of PmA individuals with the O/E ratios of PmB and PmC individuals. The alternative comparisons of the PmA O/E ratios with the heterozygote O/E ratios are unsuitable because of the heterozygote deficiency which might occur because of inbreeding and/or population subdivision. Table 16 shows the PmA O/E ratios (from Table 14) compared with the O/E ratios calculated from the total of PmB and PmC individuals. The 1968 overwinter (O) and juvenile (J) and 1969 J samples show deficiencies of PmA individuals when compared to PmB and PmC individuals. In addition, the total J, total O and grand total samples all include a relative deficiency of PmA individuals. The amount of selection against PmA animals which would account for the differences in the O/E ratios of PmA and PmB+PmC individuals has been calculated (Table 16) using the formula:

$$S=\frac{[E(A)][O(B+C)/E(B+C)]-O(A)}{O(A)},$$

where O(A) and E(A) are the observed and expected numbers of PmA individuals,

and O(B+C) and E(B+C) are the observed and expected numbers

Table 13. A comparison of the observed and expected number of the *Tf* locus phenotypes in the total samples of overwintered (O) and juvenile (J) mice. The expected numbers are the sums of the expected numbers from each sample, calculated using the Hardy-Weinberg distribution. The χ^2 and p values (one degree of freedom) are given for the goodness of fit of the observed to the expected numbers.

Year	Age Class		*Tf* locus Phenotypes: J	JM	M	
1968	O	Observed	72	98	57	χ^2=1.98
		Expected	66.7	108.6	51.7	p < .20
		Obs./Exp.	1.08	0.90	1.10	
1968	J	Observed	181	264	98	χ^2=0.54
		Expected	185.2	255.6	102.2	p < .50
		Obs./Exp.	0.98	1.03	0.96	
1969	O	Observed	96	172	68	χ^2=0.58
		Expected	99.5	165.1	71.5	p < .50
		Obs./Exp.	0.96	1.04	0.95	
1969	J	Observed	150	218	97	χ^2=0.01
		Expected	149.6	218.9	96.6	p < .95
		Obs./Exp.	1.00	1.00	1.00	
1970	O	Observed	101	169	69	χ^2=0.10
		Expected	102.5	166.3	70.5	p < .80
		Obs./Exp.	0.99	1.02	0.98	
1970	J	Observed	119	168	92	χ^2=3.90
		Expected	109.5	187.3	82.5	p < .05
		Obs./Exp.	1.09	0.90	1.12	
Total	O	Observed	269	439	194	χ^2=0.00
		Expected	268.7	440.0	193.7	p < .98
		Obs./Exp.	1.00	1.00	1.00	
Total	J	Observed	450	650	287	χ^2=0.43
		Expected	443.3	661.8	281.3	p < .60
		Obs./Exp.	1.01	0.98	1.02	
Total	Total	Observed	719	1089	481	χ^2=0.28
		Expected	713.0	1101.8	475.0	p < .70
		Obs./Exp.	1.01	0.99	1.01	

Table 14: A comparison of the observed and expected numbers of the *Pm* locus phenotypes in the pooled samples of overwintered (O) and juvenile (J) mice. The expected numbers are the sums of the expected numbers from each sample, calculated using the Hardy-Weinberg distribution. The χ^2 value and p values (three degrees of freedom) are given for the goodness of fit of the observed to the expected numbers.

Year	Age Class		*Pm* locus Phenotypes: A	AB	B	BC	C	AC	
1968	O	Observed	82	61	23	6	2	12	χ^2=4.02
		Expected	75.7	71.9	17.8	5.6	1.4	13.6	p < .30
		Obs./Exp.	1.08	0.85	1.29	1.07	1.43	0.88	
1968	J	Observed	190	145	59	25	5	45	χ^2=8.60
		Expected	175.8	170.4	46.8	24.0	4.0	48.1	p < .05
		Obs./Exp.	1.08	0.85	1.26	1.04	1.25	0.94	
1969	O	Observed	137	120	35	16	1	23	χ^2=1.89
		Expected	133.4	124.5	34.7	12.2	1.6	25.6	p < .70
		Obs./Exp.	1.03	0.96	1.01	1.31	0.63	0.90	
1969	J	Observed	195	148	48	17	6	30	χ^2=6.22
		Expected	186.0	157.9	44.8	16.4	2.7	35.8	p < .20
		Obs./Exp.	1.05	0.94	1.07	1.04	2.22	0.84	
1970	O	Observed	138	102	20	26	9	44	χ^2=0.75
		Expected	137.0	100.9	22.0	23.3	8.9	47.3	p < .90
		Obs./Exp.	1.01	1.01	0.91	1.12	1.01	0.93	
1970	J	Observed	159	123	18	23	10	46	χ^2=3.52
		Expected	164.4	110.7	24.8	22.0	9.8	47.7	p < .40
		Obs./Exp.	0.97	1.11	0.73	1.05	1.02	0.96	
Total	O	Observed	357	283	78	48	12	79	χ^2=3.00
		Expected	346.1	297.3	74.5	41.1	11.9	86.5	p < .50
		Obs./Exp.	1.03	0.95	1.05	1.17	1.01	0.91	
Total	J	Observed	544	416	125	65	21	121	χ^2=4.60
		Expected	526.2	439.0	116.4	62.4	16.5	131.6	p < .30
		Obs./Exp.	1.03	0.95	1.07	1.04	1.27	0.92	
Total	Total	Observed	901	699	203	113	33	200	χ^2=6.72
		Expected	872.3	736.3	190.9	103.5	28.4	218.1	p < .10
		Obs./Exp.	1.03	0.95	1.06	1.09	1.16	0.92	

Table 15. A comparison of the observed and expected numbers of the *Ng* locus phenotypes in the total samples of overwintered (O) and juvenile (J) mice. The expected numbers are the sums of the expected numbers from each sample, calculated using the Hardy-Weinberg distribution. The χ^2 and p values (three degrees of freedom) are given for the goodness of fit of the observed to the expected numbers.

Year	Age Class		*Ng* locus Phenotypes: J	JM	M	MP	P	JP	
1968	O	Observed	3	37	90	39	10	21	χ^2=9.33
		Expected	6.1	39.0	84.3	48.3	9.5	12.7	p < .05
		Obs./Exp.	0.49	0.95	1.07	0.81	1.05	1.65	
1968	J	Observed	23	74	236	113	22	46	χ^2=13.00
		Expected	17.0	96.5	218.7	125.1	21.2	35.5	p < .01
		Obs./Exp.	1.35	0.77	1.08	0.90	1.04	1.30	
1969	O	Observed	13	78	120	80	15	29	χ^2=0.99
		Expected	16.1	73.4	121.9	80.8	15.4	27.4	p < .90
		Obs./Exp.	0.81	1.06	0.98	0.99	0.97	1.06	
1969	J	Observed	17	115	149	93	39	36	χ^2=13.67
		Expected	19.8	102.9	144.8	113.5	25.6	42.4	p < .01
		Obs./Exp.	0.86	1.12	1.03	0.82	1.52	0.85	
1970	O	Observed	6	65	128	106	17	17	χ^2=4.22
		Expected	7.2	57.1	136.5	97.0	18.8	22.6	p < .30
		Obs./Exp.	0.83	1.14	0.94	1.09	0.90	0.75	
1970	J	Observed	10	85	120	120	24	19	χ^2=9.65
		Expected	10.3	72.4	132.7	107.6	24.2	31.9	p < .05
		Obs./Exp.	0.97	1.17	0.90	1.12	0.99	0.61	
Total	O	Observed	22	180	338	225	42	67	χ^2=2.94
		Expected	29.4	169.5	342.7	226.1	43.7	62.7	p < .50
		Obs./Exp.	0.75	1.06	0.99	1.00	0.96	1.07	
Total	J	Observed	50	274	505	326	85	101	χ^2=2.78
		Expected	47.1	271.8	496.2	346.2	71.0	109.0	p < .50
		Obs./Exp.	1.06	1.01	1.02	0.94	1.20	0.93	
Total	Total	Observed	72	454	843	551	127	168	χ^2=2.84
		Expected	76.5	441.3	838.9	572.3	114.7	171.7	p < .50
		Obs./Exp.	0.94	1.03	1.00	0.96	1.11	0.98	

Table 16. Comparison of the ratios of observed to expected numbers of PmA and PmB+PmC animals. S is the amount of selection against PmA animals required to account for the differences in the O/E ratios. N is the observed number of animals of each phenotype.

Year	Age Class	PmA		PmB+PmC		S
		O/E	N	O/E	N	
1968	O	1.08	82	1.30	25	.20
1968	J	1.08	190	1.26	64	.17
1969	O	1.03	137	0.99	36	*
1969	J	1.05	195	1.14	54	.09
1970	O	1.01	138	0.94	29	*
1970	J	0.97	159	0.81	28	*
Total	O	1.03	357	1.04	90	.01
Total	J	1.03	544	1.10	146	.06
Total	Total	1.03	901	1.08	236	.05

*The PmA O/E ratio is lower than the PmB+PmC O/E ratio, so no selection against PmA animals can be calculated.

of the total of PmB and PmC individuals.

This formula is based on the assumption that the preselection homozygote O/E ratios are the same. Any selection against PmB and/or PmC animals would reduce the estimates of selection against PmA animals obtained by this method. Because there is evidence (which will be discussed later) that there is overwinter selection against PmB and/or PmC individuals, the best estimate of selection against PmA animals is obtained from the total juvenile sample. The observed difference between the total juvenile PmA O/E ratio and the PmB+PmC O/E ratio indicates that there has been approximately 6% selection against PmA animals prior to the collection of the juvenile samples.

The two intrageneration comparisons that are possible (1968 J to 1969 O; 1969 J to 1970 O) can be used to detect whether other components of selection against the *Pm* locus phenotypes act during the overwinter period. The homozygote O/E ratios for these samples are shown in Table 17, together with estimates of the amount of selection necessary to account for the observed intrageneration changes. These estimates of selection were calculated using the formula: $S=\frac{(OJ/EJ)(EO)-OO}{OO}$

where EJ and OJ are the expected and observed numbers of juveniles, and EO and OO are the expected and observed numbers of overwintered mice.

This method simply uses the juvenile O/E ratio to predict the preselection number of animals in the overwinter class. Although there are large reductions in the PmC O/E ratios during the two overwinter periods, the number of PmC animals is too small for meaningful calculation of selection coefficients. The other two homozygous classes both appear to be subject to selection. The estimates of overwinter selection are .04 and .05 against PmA animals and .20 and .18 against PmB animals. This results, of course, in increases in

Table 17. The amount of selection (S) required to account for the observed intrageneration changes in the O/E ratios of PmA and PmB animals. N is the observed number of each phenotype.

Year	Age Class	*Pm* Locus Phenotype: A			B			C	
		N	O/E	S	N	O/E	S	N	O/E
1968	J	190	1.08	.05	59	1.26	0.20	5	1.25
1969	O	137	1.03		35	1.01		1	0.63
1969	J	195	1.05	.04	48	1.07	0.18	6	2.22
1970	O	138	1.01		20	0.91		9	1.01

the O/E ratios of all three heterozygous classes during the overwinter period. Particularly dramatic are the excesses of PmBC individuals in the 1969 O, 1970 O and total O samples (see Table 14).

In summary, there appears to be: (i) selection of approximately 6% against PmA individuals prior to the collection of the juvenile samples; (ii) further overwinter selection of approximately 5% against PmA individuals; and (iii) overwinter selection against PmB individuals of approximately 20%. If there is additional selection against PmB and/or PmC individuals prior to the collection of the juvenile samples, then (i) above is perhaps an underestimate. If there is selection against PmC individuals, then the resulting picture is one of general overdominance at the *Pm* locus.

The *Ng* locus

The phenotypes of the *Ng* locus usually showed greater deviations from the expected Hardy-Weinberg proportions (Table 15) than did the *Pm* locus phenotypes. In addition, inspection of the O/E ratios reveals that the deviations are not caused by a consistent heterozygote excess or deficiency. For example, the 1968 O and J samples show excesses of NgJP individuals and deficiencies of the other heterozygotes, whereas the 1970 O and J samples show deficiencies of NgJP individuals and excesses of the other heterozygotes. Table 18 shows the selection pressures which would cause the intrageneration differences for the two comparisons which can be made. Note that: (i) the O/E ratios of the 1968 J sample are different from the 1969 J sample, suggesting that there were different forces acting prior to the collection of the juvenile samples; (ii) the same phenotypes are favored in both the overwinter periods; and (iii) there are large differences in the selection coefficients between the two overwinter periods. However, the total O and total J samples show that the net result of the fluctuations of the O/E ratios

Table 18. The amount of selection (S) required to account for the observed intrageneration changes in the O/E ratios of the *Ng* locus phenotypes. N is the observed number of each phenotype.

Year	Age Class	*Ng* Locus Phenotypes														
		J			JM		M			MP		P			JP		
		N	O/E	S	N	O/E	N	O/E	S	N	O/E	N	O/E	S	N	O/E	S
1968	J	23	1.35	.40	74	0.77	236	1.08	.09	113	.90	22	1.04	.06	46	1.30	.19
1969	O	13	0.81		78	1.06	120	0.98		80	.99	15	0.97		29	1.06	
1969	J	17	0.86	.03	115	1.12	149	1.03	.09	93	.82	39	1.52	.40	36	0.85	.12
1970	O	6	0.83		65	1.14	128	0.94		106	1.09	17	0.90		17	0.75	

is that there is an overall excess of all three heterozygote classes in the overwintered animals compared to the juvenile animals. The amount of selection which would account for the differences in the phenotype frequency distributions of overwintered and juvenile mice is high against NgJ and NgP homozygotes and low against NgM, an observation which would account for *NgM* being the most frequent allele in all populations sampled. This overwinter selection against the three homozygotes is manifest in the O/E ratios of the phenotypes in the total O and total J samples (see Table 15). The three lowest O/E ratios in the total J sample are those of the heterozygotes, while the reverse is true of the total O sample. The overwinter selection against NgJP animals is not manifest in a reduced O/E ratio in the total O sample (relative to the other heterozygotes), which suggests that there may have been selection against NgJ and NgP animals prior to the collection of the juvenile samples. In any case, the general picture at the *Ng* locus is one of general overdominance.

The *Tf* locus

Except for one sample (1970 J) the *Tf* locus phenotypes show a close fit to the expected Hardy-Weinberg proportions. The question arises as to whether the good fit is because of the absence of factors causing deviations, or whether two or more factors have balanced each other prior to the collection of samples. Because the juvenile samples show heterozygote deficiencies at the *Pn* and *Ng* loci, which are presumably a result of inbreeding and/or population subdivision, a similar heterozygote deficiency would be expected to occur at the *Tf* locus. The conclusion is that there has been selection against the *Tf* homozygotes prior to the collection of juvenile mice. In addition, both intrageneration comparisons show small increases in the heterozygote O/E ratio, and the total overwinter sample has a higher TfJM O/E ratio than the total juvenile sample.

The magnitude of the selection against *Tf* homozygotes can be calculated if an estimate of the preselection heterozygote deficiency is obtained from the *Pm* and *Ng* loci. The unweighted mean values of F+ϕ(=1-O/E) from the heterozygotes of the *Pm* and *Ng* loci in the total juvenile samples are .03 and .04, respectively. The amount of selection against the *Tf* homozygotes which would cause a preselection 3 to 4% heterozygote deficiency to be changed to the observed total O heterozygote O/E ratio of 1.00 is approximately 4%, if the selection is the same against both homozygotes.

The Distribution of Phenotypes in the Offspring of Females Pregnant When Captured

As has been discussed at length, two instances of segregation distortion were discovered as a result of an extensive laboratory breeding program. The possibility exists that these distortions were caused by laboratory breeding conditions. This possibility could be dismissed if distorted segregation ratios could be demonstrated in the offspring of females pregnant when captured. For this purpose, all pregnant females captured during 1970 were allowed to raise their litters in captivity. The distribution of phenotypes in these mothers and offspring is presented in Table 19.

Cooper (1968) has analyzed the ratios which are expected in litter data when the phenotype of one parent is unknown (which he calls incomplete family data; I.F.D.). There are two types of ratios which might be applied to show the effects of distorted segregation ratios:

(a) If two alleles are segregating, the ratio of homozygotes to heterozygotes in the offspring of heterozygous females should be 1:1 (which Cooper calls the 1:1 ratio). The 1:1 ratio can only be applied to the *Tf* locus data. Ignoring phenotypes including rare alleles, this ratio in the offspring of TfJM females is exactly 1:1 (109:109). However, the sensitivity of a test for a deviation from the 1:1 ratio because of distorted ratios in just the one cross is low, because it is expected that less than half of litters from heterozygous females will be sired by heterozygous males (when $p \neq q$).

(b) The second type of I.F.D. ratio which might be applied to demonstrate segregation distortion Cooper calls the gametic ratios. If there are three genotypes A_1A_1, A_1A_2 and A_2A_2, and the frequency of the A_1 allele is p, then

Table 19. Distribution of phenotypes in the offspring of females pregnant when captured.

Phenotype of Mother	Number of Litters	Number of Offspring	Number of Offspring of the Phenotype:								
			Tf locus								
			HM	J	JM	JK	K	KM	M	JP	JQ	MQ
J	31	151		81	63	7						
JK	2	5		2	2			1				
JM	47	229	1	60	109	3		1	49	1	4	1
KM	1	5					3	2				
M	22	102			63				39			
MP	2	11			4					7		
JQ	1	3			2							1
MQ	1	5			2				1		1	1
			Pm locus								
			A		AB		B		BC		C	AC
A	52	246	182		45							19
AB	32	151	55		69		19		5			3
B	5	25			17		8					
BC	8	36			6		3		12		10	5
AC	10	53	14		4				3		10	22
			Mg locus									
			J		JM		M		MP		P	JP
JM	19	92	5		27		34		15			11
M	46	223			39		140		44			
MP	31	143			9		50		58		12	14
P	5	21							6		11	4
JP	6	32	5		5				8		6	8

(i) A_1A_1 and A_1A_2 offspring from A_1A_1 mothers,

(ii) A_1A_1 and A_2A_2 offspring from A_1A_2 mothers

and (iii) A_1A_2 and A_2A_2 offspring from A_2A_2 mothers

should all be in the ratio p:1-p. Gametic ratio (ii) is the one which applies to the case when the distortion affects the offspring of heterozygous females. If the segregation distortions discovered at the *Pm* and *Tf* loci cause deviations from this ratio then one would expect there to be excesses of *PmA* and *TfM* when compared with p:1-p for each case. If p is the frequency of *PmA*, calculated from the distribution of phenotypes in the female parents (assuming the gene pools to be identical in the male and female parents) with *PmC* considered to be identical to *PmB*, then the ratio p:1-p is .68:.32, or 2.1. The ratio of PmA to PmB offspring from PmAB mothers is 55:24, or 2.3, which represents an excess of between one and two PmA individuals. Similarly, the ratio of p:1-p at the *Tf* locus (when p is the frequency of *TfM*) is .44:.56, or 0.79, whereas the ratio of TfM offspring to TfJ offspring from TfJM mothers, if all rare alleles are considered to be the same as *TfJ*, is 49:68, or 0.72. This ratio represents a deficiency of approximately three TfM individuals. Thus Cooper's I.F.D. ratios cannot be used to accept or reject the null hypothesis that the segregation ratios are normal among litters conceived in the woods.

Although the collection of data from females pregnant when captured proved futile in detecting the effects of segregation distortion, a discovery was made which is an interesting example of how genetic information can help to describe population processes. An examination of Table 19 shows that in every case the phenotypes of the offspring and that of the mother are in agreement with the hypotheses of inheritance which have been substantiated by laboratory breeding. However, within 11 litters the genotypes of the individual mice involved indicate that three different alleles of a single locus

have been derived from male parents. The phenotypes of the mother and offspring of these litters are given in Table 20. Any errors in the interpretation of the phenotypes or inheritance patterns are improbable explanations because of the large number of crosses made in captivity. The only reasonable explanation of this phenomenon is polyandry. Other potential explanations can be eliminated: The distribution of phenotypes in the aberrant litters is not explicable by the presence of null alleles (at the *Tf* and *Ng* loci) or a non-reactive allele (at the *Pm* locus). Mutation is not a likely cause, because the minimum number of mutational events which would account for the aberrant phenotypes among the 511 offspring is 12. Spontaneous mutation at such a high frequency is not compatible with the fact that no postulated mutational events occurred among over 2000 offspring from laboratory crosses.

The Probability of Detection of Multiple Insemination

Given that evidence for multiple insemination has been obtained in 11 of 107 litters, it is of interest to determine the frequency of occurrence which this proportion might represent. This has been done by calculating the probability of detection of multiple insemination, determining the number of litters which would be expected to show the phenomenon if all fertilizations were polyandrous, and comparing that with the observed number.

Two general conditions exist for the detection of multiple insemination:

(a) The males involved must carry at least three alleles of the locus in question.

(b) At least three paternal alleles must occur in the offspring in combination with maternal alleles so that they can be confirmed as being paternal.

In calculating the specific probabilities for (a) and (b) it has been assumed that only three alleles, distributed in Hardy-Weinberg proportions, are present

Table 20. The phenotypes of the mother and offspring of those litters which include three paternal alleles.

# of Mother	Locus	Phenotype of Mother	Phenotypes of Offspring
2209	*Tf*	JM	J, JM, HM, J, M, JM, M.
2218	*Pm*	AB	A, B, AB, AB, AC.
2193	*Pm*	AB	A, A, A, B, AB, BC.
2215	*Pm*	A	A, A, A, AB, AC, AC.
2232	*Pm*	AB	A, B, AB, AB, BC.
2322	*Pm*	A	A, A, A, AB, AC.
2835	*Pm*	A	A, AB, AB, AC, AB.
2151	*Ng*	P	P, JP, MP, P.
2201	*Ng*	M	M, M, MP, JM.
2235	*Ng*	MP	P, JP, M, MP, M, JP.
2639	*Ng*	M	M, JM, M, JM, MP, MP.

in the population and that sperm from all males involved in the multiple insemination have equal probabilities of success in fertilization.

The probability of occurrence of the first of the two conditions is least when there are only two males involved, and increases as the number of males increases. When the number of males is large, then the male gamete pool has effectively the same constitution as the population gene pool, and the probability of three different alleles being present in the males approaches one. The probability of two males carrying three different alleles is 12pqr, where p, q and r are the frequencies of the three alleles. This solution is shown in Table 21. Thus the probability of the males carrying three different alleles is between 12pqr and one, depending on the number of males involved.

The probability of occurrence of the second condition, that is that three paternal alleles be present and detectable in the offspring, depends on the maternal genotype, the particular combination of maternal and paternal genotypes, the litter size, and the number of males involved. Those cases which have been examined and the respective probabilities are presented in Table 22. Following is an explanation of how these were derived.

When the mother is homozygous, there will be three offspring classes which will occur in the ratio with which the three alleles are distributed among the males. This ratio is 2:1:1 when two males are involved, because if there are three different alleles among the four carried by two males, there must be two copies of one of them. When a large number of males contribute, the ratio is p:q:r. The presence of each of the three offspring classes will confirm the presence of one paternal allele, so all three offspring classes are necessary in one litter. The probability that any litter of size n from a homozygous female will include all three offspring classes is obtained by

Table 21. Determination of the probability of choosing two males carrying three alleles from a population in which three alleles of a gene are distributed in Hardy-Weinberg proportions. The genotypes in the population are A_1A_1, A_1A_2, A_2A_2, A_2A_3, A_3A_3 and A_1A_3, and p, q and r are the frequencies of A_1, A_2 and A_3, respectively.

Genotype of First Male	Genotype of Second Male	Probability
A_1A_1	A_2A_3	$p^2(2qr)$
A_1A_2	A_1A_3, A_2A_3, A_3A_3	$2pq(2pr + 2qr + r^2)$
A_2A_2	A_1A_3	$q^2(2pr)$
A_2A_3	A_1A_1, A_1A_2, A_1A_3	$2qr(p^2 + 2pq + 2pr)$
A_3A_3	A_1A_2	$r^2(2pq)$
A_1A_3	A_1A_2, A_2A_2, A_2A_3	$2pr(q^2 + 2pq + 2qr)$
	Total Probability	12 pqr

Table 22. The probability that three alleles will be present and detectable in the offspring. See text for explanation.

Maternal Genotype	Expansion Used to Obtain Probability	Number of Males Involved	Special Case[a]	Probability of each Offspring Class	Litter Size	Theoretical Probability (P)		
Homozygous	$(a+b+c)^{n\dagger}$	2		a = 1/2 b = 1/4 c = 1/4	3 4 5 6 7	0.19 0.38 0.53 0.64 0.73		
		∞		a = p b = q c = r	3 4 5 6 7	P per gamete 0.22 0.45 0.62 0.74 0.83	P per locus 0.10 0.20 0.29 0.37 0.44	P per locus 0.15 0.31 0.44 0.55 0.64
Heterozygous	$(e+f+g+h)^{n\dagger}$	2	A	e = 3/8 f = 1/4 g = 1/8 h = 1/4	3 4 5 6 7	0.05 0.13 0.23 0.33 0.42		
		2	B	e = 1/4 f = 1/8 g = 1/8 h = 1/2	3 4 5 6 7	0.05 0.12 0.20 0.27 0.31		
A_1A_3		∞		e = (p+q)/2 f = p/2 g = q/2 h = r	3 4 5 6 7	P per gamete 0.06 0.15 0.26 0.36 0.45	P per locus 0.03 0.07 0.14 0.21 0.28	P per locus 0.04 0.11 0.19 0.28 0.36
A_1A_3		∞		e = (p+r)/2 f = p/2 g = r/2 h = q	3 4 5 6 7	0.06 0.15 0.26 0.36 0.45	0.03 0.07 0.12 0.17 0.22	0.04 0.09 0.15 0.21 0.23
A_2A_3		∞		e = (r+q)/2 f = q/2 g = r/2 h = p	3 4 5 6 7	0.06 0.15 0.26 0.36 0.45	0.03 0.06 0.10 0.14 0.15	0.04 0.11 0.19 0.28 0.37

[a]A is the case when the allele carried twice by the males is present in the female parent; B is the case when it is not.

†n is litter size; a, b and c are the probabilities of obtaining offspring carrying paternal alleles A_1, A_2 and A_3, which occur in the population with the frequencies p, q and r, respectively.

†e is the probability of an individual having the same genotype as the mother, f and g are the probabilities of an individual being homozygous for each of the alleles carried by the mother, and h is the sum of the probabilities of an individual carrying the allele not present in the mother.

adding those terms of the expansion of $(a+b+c)^n$ which include a, b, and c, where a, b and c are the probabilities of obtaining each offspring class. This probability has been calculated for litter sizes from 3 to 7, when two males are involved, and when an infinitely large number of males are involved. For the latter case, probabilities have been calculated for the maximal situation (p=q=r) and for values of p, q and r which are the mean frequencies of the alleles of the *Pm* and *Ng* loci from the natural population samples. Because of the difficulty of estimating the frequency of alleles other than *TfJ* and *TfM*, because of their rarity, this has not been done for the *Tf* locus.

When the female parent is heterozygous and there are two males involved, the probability of detection of three paternal alleles when the allele which is present twice in the male is also present in the female parent is different than when it is not. In either case, the probability is obtained by expanding the expression $(e+f+g+h)^n$, adding those terms which include f, g and h, and substituting the appropriate probabilities, where e is the probability of an individual having the same genotype as the mother, f and g are the probabilities of an individual being homozygous for each of the alleles carried by the mother, and h is the sum of the probabilities of an individual carrying the allele not present in the mother. When the paternal allele which is present twice is also carried by the female parent then e=3/8, f=h=1/4 and g=1/8. (Note that the probabilities of obtaining the two genotypes which include the allele not carried by the female parent can be treated as one, because the presence of either or both of them provides confirmation of the presence of that allele.)

When the female parent is heterozygous and a large number of males are involved, the probability of detecting three different paternal alleles is

again obtained from the expansion of $(e+f+g+h)^n$. However, the probabilities are different for each different heterozygous female parent (unless p=q=r), as shown in Table 22. For each of the heterozygotes, the probabilities have been calculated when p=q=r, and for values of p, q and r which are the mean frequencies of the alleles of the *Pm* and *Ng* loci from the natural population samples.

Using these probability calculations, the proportion of all litters of a given size which could be confirmed as having more than one male parent, if there are two males involved in all cases, or if there is a large number of males involved, can be estimated. For any litter size n, we get $\pi_n=\Sigma MFP$, where π_n is the probability of detection of multiple insemination, M is the probability of the males involved carrying three different alleles, F is the frequency of occurrence of each case and P is the probability of detection of the three different alleles (from Table 22). These calculations are shown in Table 23.

Thus the two values of π_n for each locus are estimates of the maximum and minimum values of the range in which the observed proportion of litters showing multiple insemination would fall if more than one male is involved in every case. Using these values, the observed number of litters demonstrating multiple insemination can be compared with the expected, as shown in Table 24. Pooling the expected and observed numbers from the two loci, there were ten litters observed, and the range of the expected number of litters is from 23.50 to 51.43. The actual frequency of occurrence of insemination by more than one male can be estimated from these figures as being from 0.43 to 0.19. Although the lower limit of the range is unrealistic because it represents the case when an infinite number of males are involved, it has not been determined which portion of the range is most probably accurate. In addition, the

Table 23. Calculation of π_n, the proportion of all litters of size n in which could be confirmed as having more than one male parent, for the cases where there are two males involved and a large number of males involved, for the conditions which describe the *Pm* and *Ay* loci.

Maternal Genotype	Number of Males Involved	Special Case**	F	M	P for Litter Size: 3	4	5	6	7	M x F x P 3	4	5	6	7
				Pm locus										
Homozygous	2		0.53*	$12pqr = 0.20$	0.19	0.38	0.53	0.64	0.73	0.02	0.04	0.06	0.07	0.08
Heterozygous	2	A	$2pq+2pr+2qr - 6p^2q^2r^2 = 0.45$	0.20	0.10	0.20	0.29	0.37	0.44	0.01	0.02	0.03	0.03	0.04
	2	B	$6p^2q^2r^2 = 0.02$	0.20	0.10	0.20	0.29	0.37	0.44	0.00	0.00	0.00	0.00	0.00
										π_n = 0.03	0.06	0.09	0.10	0.12
Homozygous	-		0.53*	1	0.10	0.20	0.29	0.37	0.44	0.05	0.11	0.15	0.20	0.23
Heterozygous														
AB	-		$2pq = 0.32$	1	0.03	0.07	0.14	0.21	0.28	0.01	0.02	0.04	0.07	0.09
AC	-		$2pr = 0.10$	1	0.03	0.07	0.12	0.17	0.22	0.00	0.01	0.01	0.02	0.02
BC	-		$2qr = 0.05$	1	0.03	0.06	0.10	0.14	0.15	0.00	0.00	0.01	0.01	0.01
										π_n = 0.06	0.14	0.21	0.30	0.35
				Ay locus										
Homozygous	2		0.48*	$12pqr = 0.31$	0.19	0.38	0.53	0.64	0.73	0.03	0.06	0.08	0.10	0.11
Heterozygous	2	A	0.48	0.31	0.15	0.31	0.44	0.55	0.64	0.02	0.05	0.07	0.08	0.10
	2	B	0.04	0.31	0.15	0.31	0.44	0.55	0.64	0.00	0.00	0.01	0.01	0.01
										π_n = 0.05	0.11	0.16	0.19	0.22
Homozygous	-		0.48*	1	0.16	0.31	0.44	0.55	0.64	0.07	0.15	0.21	0.26	0.31
Heterozygous														
JH	-		$2pq = 0.21$	1	0.04	0.11	0.19	0.28	0.36	0.01	0.02	0.04	0.06	0.08
JP	-		$2pr = 0.08$	1	0.04	0.09	0.15	0.21	0.23	0.00	0.01	0.01	0.02	0.02
HP	-		$2qr = 0.23$	1	0.04	0.11	0.19	0.28	0.37	0.01	0.03	0.04	0.06	0.09
										π_n = 0.09	0.21	0.30	0.40	0.50

*These estimates of the proportion of homozygous females were calculated from the distribution of phenotypes among the female parents (Table 19).
**A is the case when the allele carried twice by the males is present in the female parent; B is the case when it is not.

Table 24. The expected and observed numbers of litters showing the effect of insemination by more than one male.

Litter Size (n)	Πn*		Total Number of Litters**	Expected Number of Polyandrous Litters		Observed Number of Polyandrous Litters
	Two Males	Many Males		Two Males	Many Males	
			Pm locus			
3	0.03	0.06	8	0.24	0.48	0
4	0.06	0.14	18	1.08	2.52	0
5	0.09	0.21	43	3.87	9.03	4
6	0.10	0.30	29	2.90	8.70	2
7	0.12	0.35	2	0.24	0.70	0
Totals				9.33	21.43	6
			Mg locus			
3	0.05	0.09	8	0.40	0.72	0
4	0.11	0.21	18	1.98	3.78	2
5	0.16	0.30	43	6.88	12.90	0
6	0.19	0.40	29	5.51	11.60	2
7	0.20	0.50	2	0.40	1.00	0
Totals				15.17	30.00	4

*from Table 23.
**This is the number of litters of each size from females pregnant when captured.

effect of males having unequal probabilities of success in siring offspring has not been considered. In any case, it appears that litters must frequently include offspring sired by more than one male.

The discovery that the deer mice in the populations sampled during this study are promiscuous contradicts a common belief that mating patterns among deer mice are highly restrictive. For example, King (1967) reports that monogamy is not rigid among deer mice, "but usually results from the sedentary habits of the pair occupying the same nest and home range." Although comparisons between studies are difficult, the review of population studies of *Peromyscus* by Terman (1967) indicates that deer mice usually are found in much lower densities than were found during this study. It is possible, then, that the promiscuity discovered was a phenomenon which occurs more frequently in very dense populations.

CONCLUSION

Selection: Further Considerations

The discovery of direct evidence for selection from the phenotype frequencies of the *Pm* and *Ng* loci raises a number of questions. Among them: (a) How do the equilibrium allele frequencies predicted by the calculated selective forces compare with the observed allele frequencies in the populations? (b) Is there any indication of the nature of the selective forces? and (c) Is there any indication of heterogeneity of selective forces between woodlots, which might be responsible for the heterogeneity of allele frequencies between woodlots?

Comparison of the allele frequencies predicted by the estimates of selection with the observed allele frequencies is, in fact, a test of the validity of the selection estimates. Although the argument can be made that all of the selective forces have probably been underestimated, it might be expected that they are approximately correct in their relative magnitude and would therefore predict allele frequencies similar to those observed. Table 25 shows the comparison of the observed (the mean allele frequencies from the natural populations) and the predicted allele frequencies. These predictions were made, in the absence of adequate mathematical solutions, by means of a deterministic computer simulation program (Appendix I) similar to the simulation programs described previously. Except for PmA animals (for which the selection estimate is the sum of the prejuvenile and overwinter selection estimates), the estimates of selection (S) are the mean of the overwinter estimates. These selection estimates (together with segregation distortion at the *Pm* locus) predict equilibrium allele frequencies which are similar to those observed.

Table 25: Comparison of predicted allele frequencies based on segregation distortion (K) and selection (S) at the *Pm* locus, and selection at the *Ng* locus, with the allele frequencies observed in the natural populations.

Locus	Phenotype	K	S	Allele	Predicted Allele Frequency	Observed Allele Frequency
Pm	AB (♀♀)	0.56		*PmA*	0.73	0.60
	A		0.11	*PmB*	0.17	0.31
	B		0.19	*PmC*	0.10	0.09
	C		0.72			
Ng	J		0.22	*NgJ*	0.17	0.20
	M		0.09	*NgM*	0.68	0.58
	P		0.23	*NgP*	0.15	0.22
	JP		0.16			

It is not known whether the heterogeneity of allele frequencies between woodlots reflects environmental heterogeneity, or whether the differences result from genetic drift. No pattern of allele frequency which might indicate a selective basis to the heterogeneity was detected in the woodlots. However, Savage (pers. comm.) reports an apparent cline in the frequency of *PmC*, based on samples of *Peromyscus maniculatus* from several locations on the continent. The elucidation of the answers to the general problem of selection versus drift as it applies to regional heterogeneity will require many more years of study.

There has been no indication of the nature of any selective agents during the course of this study. A brief review of some of those red cell antigen and serum protein polymorphisms for which modes of selection have been postulated will serve to partially define the range of possible factors which may be acting on the polymorphisms described here.

Human red blood cell antigen polymorphisms, such as the ABO and Rh systems, have received considerable attention because of their medical importance. The antigens of the ABO system appear to confer resistance to a variety of diseases (Race and Sanger, 1968; Clarke, 1959) and have also been shown to interact with the Rh antigens, influencing the severity of hemolytic reactions in the case of maternal-foetal incompatibility (Levine, 1958; Cohen and Glass, 1959; Cohen, 1960). Other blood group systems, in man and other mammals, have not generally been found to affect fitness in such an obvious manner. Savage (1971) found no evidence for maternal-foetal incompatibility at the *Pm* locus in deer mice; in fact, Rasmussen (1964) assumed that there was little or no selection acting at the *Pm* locus, so as to use an observed heterozygote deficiency for the analysis of population structure. Cameron (1966) criticized this approach, after obtaining both direct and

indirect evidence for overdominance at the *Hg* (blood group) locus in an insular population of rabbits. Gilmour (1959) presented indirect evidence for overdominance at blood group loci in chickens, showing that segregation of variants continues after long periods of intense inbreeding. However, neither the study of Cameron nor that of Gilmour was productive in the identification of selective agents.

Disease appears to be a common selective force. For example, several human polymorphisms appear to affect resistance to malarial parasitism (review: Dubos, 1965). Manwell and Baker (1970) suggest that transferrin variation may confer differential disease resistance, because it has been found to have anti-viral and anti-bacterial activity *in vitro*, and appears to show more variation in gregarious species. Two independent studies (Canham, 1969; Tamarin and Krebs, 1969) report density-dependent selection on transferrin variants, which may or may not be related to disease resistance. Similarly, esterase allele frequencies have been found to correlate with population density (Semeonoff and Robertson, 1968).

Johnson *et al.* (1969) have shown that esterase allele frequencies in populations of harvester ants correlate with weather. Also, Koehn and Rasmussen (1967) found that esterase allele frequencies of *Catostomus clarki* populations in the Colorado River system correlate with latitude, suggesting that temperature, or perhaps photoperiod, is the selective agent. A similar study of esterases of another freshwater fish *(Notropis stramineus)* (Koehn, Perez and Merritt, 1971) in the Kansas River resulted in the proposal that heterozygotes are at an advantage at higher environmental temperatures and homozygotes at lower environmental temperatures.

These studies demonstrate that a wide variety of selective agents can influence the distribution of genetic variation. Description of the particular

selective agents affecting the distribution of the variation described in this study will probably require an equally wide variety of physiological, ecological and genetical approaches.

Genetic Load

Because the traditional idea that polymorphisms are maintained by selection has been questioned on the grounds that no organism could cope with the "cost" (Haldane, 1957) of selection acting on many variable loci, it is pertinent to consider the estimates of selection obtained for the three loci studied here with respect to the total "cost," or genetic load. Table 26 shows the calculations of the mean fitness ($\bar{W}$) for each locus, which is simply the proportion of animals remaining after selection has occurred. In every case, it can be argued that the values of S are underestimates, because of the difficulty of estimating the magnitude of selection which occurred prior to the collection of the juvenile samples.

It is a common procedure to estimate the average fitness of a population as a result of variation at more than one locus by obtaining the product of the values of $\bar{W}$ for each locus (Kimura and Crow, 1964). The biological reality of this procedure has been questioned (Sved, Reed and Bodmer, 1967; King, 1967; Milkman, 1967). One of the bases for the questioning is that the model of Kimura and Crow ignores the possibility of gene interaction. In order to look for gene interaction between the *Tf*, *Pm* and *Ng* loci, contingency χ^2 tests were done on all possible combinations of phenotypes. The expected values were calculated by assuming that if there is no gene interaction or linkage disequilibrium, then the expected frequency of a phenotype with respect to more than one locus is simply the product of the frequencies of the phenotypes attributable to the individual loci. Since the results

Table 26: Calculation of $\bar{W}$, the proportion of the population remaining after selection (S) on phenotypes of each of the loci.

Locus	Phenotype	Frequency	S	Proportion Eliminated	$\bar{W}$
Pm	A	$p^2 = 0.36$	.11	0.04	
	B	$q^2 = 0.10$	.19*	0.02	
				1 - 0.06	= 0.94
Ng	J	$p^2 = 0.04$	.22*	0.01	
	M	$q^2 = 0.34$	.09*	0.03	
	P	$r^2 = 0.05$	.23*	0.01	
	JP	$2pr = 0.09$	.16*	0.01	
				1 - 0.06	= 0.94
Tf	J	$p^2 = 0.27$	.04	0.01	
	M	$q^2 = 0.23$	.04	0.01	
				1 - 0.02	= 0.98

*These values of S are the mean values of the selection calculated from the two intragenerational comparisons.

from the laboratory crosses indicate that there is a very low probability that the loci are tightly linked, linkage disequilibrium is unlikely to be a reason for any deviations from the expected frequencies. Thus the presence of significant deviations from the expected frequencies would be evidence for gene interaction. No significant χ^2 values resulted from these tests. The possibility that there is some interaction between the loci is not ruled out, however.

The cumulative value of $\bar{W}$ from the *Pm*, *Tf* and *Ng* loci is 0.94 X 0.94 x 0.98 = 0.87. Using Kimura's model, the conclusions would be that even though selection has probably been underestimated, variation in fitness at these three loci would necessitate the removal of 13% of the population. It is obvious that the ability of a population to replace such a large proportion through overreproduction would be overwhelmed if many more loci showed the same degree of variation of fitness. Alternative models, incorporating soft selection and fluctuating selection pressures (for which there is some evidence in this study, particularly with respect to the *Ng* locus), have been proposed (Wallace, 1970) and need mathematical definition.

An Evaluation of the Method of Study

The principle conclusion of this thesis is that selection is a factor affecting the frequencies of the phenotypes of all three loci studied. This conclusion was reached by direct observation of phenotype frequencies, rather than from inferential evidence such as the observation of temporal or spatial constancy of allele frequencies, or the detection of correlations of variation with environmental factors. In the opinion of Kimura and Ohta (1971), the observation of stable allele frequencies, and allele frequency patterns which have been commonly thought to occur because of selection, such as clines,

can be explained in terms of neutral allele theory. However, in this study the direct observation of what appears to be effects of selection, particularly the intragenerational changes, are not so easily explained if the different phenotypes of the loci studied here confer no differences in fitness on the individual animals.

Studies of particular genetic variation in natural populations have rarely included the combination of multiple populations and multiple loci together with an extensive breeding program. These aspects of the study will be discussed in terms of their usefulness for the formulation of reliable conclusions.

(a) *Multiple populations:* The examination of samples from several populations frequently served as a control for the evaluation of hypotheses based on the phenotype distributions in one woodlot, or a small number of woodlots. For example, if the distribution of transferrin phenotypes in the 1969 sample from IL2, in which there was a significant excess of heterozygotes, were examined by itself, the hypothesis that there is selection against transferrin homozygotes would probably be formulated. In contrast, the significant deficiency of transferrin heterozygotes in the 1970 sample from RR would cause the formulation of a much different hypothesis. Although it is quite possible that different non-random processes could have caused these deviations, these processes are unlikely to be defined with any reliability. Statistically significant deviations among a group of samples can only be considered as part of the normal stochastic variation and probably do not merit special interest unless the frequency of their occurrence is much greater than expected.

(b) *Multiple loci:* The use of three loci as genetic markers has also served as a control. The best example of how this was an advantage is the manner in which selection against transferrin homozygotes was detected. The

observation that the O/E ratios of transferrin phenotypes are usually very close to one (see Table 13) might have resulted in the conclusion that there was no selection acting, except that it was determined, from the distributions of phenotypes of the *Pm* and *Ng* loci, that the expected heterozygote O/E ratio was less than one.

(c) *Inheritance data:* Without the extensive breeding program, neither of the segregation distortions which have been described would have been discovered. Although the distortion at the *Tf* locus probably has little effect on allele or phenotype frequencies, the distortion at the *Pm* locus appears to be important. This can be demonstrated by determining the expected frequencies of the *Pm* locus alleles, based on the calculated estimates of selection but not the segregation distortion. The resulting expected allele frequencies are more different from the observed than those predicted when the effect of the segregation is included. (When the segregation distortion is not included, the predicted frequencies of *PmA*, *PmB* and *PmC* are 0.46, 0.42 and 0.13, respectively. These can be compared with the observed allele frequencies and the predicted allele frequencies [including segregation distortion] given in Table 25).

Rasmussen (1969) has strongly criticized typological thinking with respect to the use of biochemical differences as taxonomic indicators. In general, I would extend his criticism to apply to the use of a single population and a single locus as the basis for drawing conclusions about a population or a species. Examples of this are common in the literature. In fact, Richmond and Powell (1970), in reporting a significant excess of heterozygotes at a single enzyme locus in one population of *Drosophila paulistorum*, state that "This is probably the first reported case of heterosis associated with an enzyme locus in a natural population of *Drosophila*." It would seem

hazardous to conclude that heterosis is the cause of a single significant heterozygote excess without supportive data, particularly when it can be shown that factors other than selection can cause heterozygote excesses.

During this study, approximately 2600 mice from the natural populations were captured and phenotyped and approximately 2500 offspring were obtained from the breeding colony. Although the data on the inheritance of the genetic systems are probably extensive enough so that the hereditary basis for the variation has been well established, the samples from the natural populations have been found to be generally inadequate for the definition of the basis of the regional heterogeneity. However, in my opinion, this information is most likely to be obtained simply by extending the scope of this study, rather than by shifting the general approach. The most important phase of such an extension would be an intense effort to define the relevant demography of the populations with mark-release programs. Data from such mark-release studies can be combined with information acquired from phenotype distributions (such as the information with respect to polyandrous litters) to describe the processes of reproduction and dispersal as they are important in affecting the distribution of variation. The other extensions I would plan are simply to study more populations over a wider geographical range and to continue such studies over a number of years. Enough samples are needed so that it can be determined if populations with similar allele frequencies are being affected by the same deterministic factors. Hypotheses based on such an analysis could be tested by manipulating allele and phenotype proportions in populations.

SUMMARY

1. Over 2000 deer mice *(Peromyscus maniculatus)* were live trapped in woodlots in the Edmonton, Alberta area during the years 1968-71. The phenotypes of these mice were determined for three loci: the *Tf* locus, which controls the synthesis of transferrin; the *Pm* locus, which controls the synthesis of some red blood cell antigens; and the *Ng* locus, which controls the appearance of certain "white bands" on starch gels stained for esterase activity.

2. Nine different transferrin bands were described following starch gel electrophoresis of serum. These were named G, H, I, J, K, L, M, P and Q, in order of decreasing electrophoretic mobility. The antigens controlled by the *Pm* locus (A, B and C) were detected by testing red blood cells for agglutination in each of three antisera (anti-A, anti-B and anti-C). Six "white bands" were detected on starch gels following electrophoresis of serum. These were named H, J, K, M, P and R, in order of decreasing electrophoretic mobility.

3. The results of laboratory breeding showed that the *Pm* locus antigens and the bands on starch gels of the *Tf* and *Ng* loci are all inherited as if controlled by codominant alleles of single autosomal loci. However, PmAB females segregated a significant excess of *PmA* alleles to their offspring. Also, the TfJM X TfJM cross produced a significant excess of TfM offspring. Neither of these segregation distortions appear to be caused by zygotic selection.

4. In every woodlot the common transferrin alleles were *TfJ* and *TfM*. The three alleles of the *Pm* locus were present in all the woodlots. The common alleles of the *Ng* locus were *NgJ*, *NgM* and *NgP*. There was a large amount of

heterogeneity of allele frequencies between woodlots with respect to all three loci.

5. Deterministic computer simulation of populations was used to determine the type and magnitude of selection necessary to balance the segregation distortions. It was concluded that the distortion at the *Tf* locus is unlikely to be a major factor in the maintenance of the polymorphism. However, distortion of the type at the *Pm* locus would appear to be an important factor modifying allele frequencies and would be expected to be opposed by selection against the homozygote favored by the distortion.

6. Taking into account theoretical considerations, several different methods for the detection of natural selection were applied to the data collected from wild caught mice.

7. It appears that there is selection of approximately 11% against PmA animals. Also, PmB animals appear to be subject to overwinter selection of approximately 19%. Although there was a large reduction in the proportion of PmC animals during the overwinter period, the numbers were too small to allow meaningful estimation of selection.

8. There appeared to be great variation in the type and magnitude of selection acting on phenotypes controlled by the *Ng* locus, both before and after the collection of the juvenile samples. During the overwinter periods the average apparent selection pressures were approximately 22% against NgJ, 9% against NgM, 23% against NgP and 16% against NgJP.

9. The heterozygote deficiencies in the juvenile samples at both of the *Pm* and *Ng* loci were used to predict the preselection proportion of heterozygotes at the *Tf* locus. Using this method, evidence for selection against TfJ and TfM animals of approximately 4% was obtained.

10. An example of how genetic studies can help to describe the organization of a population was obtained. The distribution of phenotypes in 11 of 107 litters from females which were pregnant when captured indicated that there was more than one male parent. Given that only a portion of litters having more than one male parent will be detected, the actual frequency of occurrence of the phenomenon was calculated to be between 19 and 43%.

11. Evidence was obtained for selection acting on the phenotypes of all of the three loci studied. Thus it seems unlikely that cryptic variation in natural populations is frequently maintained by stochastic phenomena. Also, because the estimates of selection are great enough so as to not be compatible with simplistic calculations of genetic load, it is suggested that theories of genetic load involving soft selection and fluctuating selection pressures are more applicable.

BIBLIOGRAPHY

Allison, A.C. 1954. Protection afforded by sickle-cell trait against subtertian malarial infection. Brit. Med. J., 1: 290.

Arnheim, N. and C.E. Taylor. 1969. Non-Darwinian evolution: Consequences for neutral allelic variation. Nature, 223: 900-903.

Ashton, G.C., D.G. Gilmour, C.A. Kiddy, and F.K. Kristjansson. 1967. Proposals on nomenclature of protein polymorphisms in farm livestock. Genetics, 56: 353-362.

Bateman, N. 1960. Selective fertilization at the T-locus of the mouse. Genet. Res. (Camb.), 1: 226-238.

Bearn, A.G. and W.C. Parker. 1966. Transferrin. *In* A. Gottschalk (ed.), Glycoproteins, their composition, structure, and function. Elsevier Publishing Co., Amsterdam.

Birdsall, D.A., J.A. Redfield and D.G. Cameron. 1970. White bands on starch gels stained for esterase activity: a new polymorphism. Biochem. Genet., 4: 655-658.

Braend, M. 1965. Nomenclature of polymorphic protein systems. Nature, 206: 1067.

Brewer, G.J. 1967. Achromatic regions of the tetrazolium stained starch gels: inherited electrophoretic variation. Am. J. Human Genet., 19: 674-680.

Bruck, D. 1957. Male segregation ratio advantage as a factor in maintaining lethal alleles in wild populations of house mice. Nat. Acad. Sci. (U.S.), Proc. 43: 152-158.

Cain, A.J., J.M. King and P.M. Sheppard. 1960. New data on the genetics of polymorphism in the snail *Cepaea nemoralis* L. Genetics, 45: 393-411.

Cameron, D.G. 1966. Blood group polymorphism in an insular population of rabbits, *Oryctolagus cuniculus*. Ph.D. Thesis, Stanford University.

Cameron, D.R. and R. Moav. 1957. Inheritance in *Nicotiana tobacoum*. XXVII. Pollen killer, an alien genetic locus inducing abortion of microspores not carrying it. Genetics, 42: 326-335.

Canham, R.P. 1969. Serum protein variation and selection in fluctuating populations of cricetid rodents. Ph.D. Thesis, University of Alberta.

Canham, R.P., D.A. Birdsall and D.G. Cameron. 1970. Disturbed segregation at the transferrin locus of the deer mouse. Genet. Res. (Camb.), 16: 355-357.

Castle, W.E. 1903. The laws of heredity of Galton and Mendel, and some laws governing race improvement by selection. Proc. Am. Acad. Arts Sci., 39: 223-242.

Clarke, C.A. 1959. Correlations of ABO blood groups with peptic ulcer, cancer, and other diseases. Am. J. Human Genet., 11: 400-404.

Clarke, C.A. and P.M. Sheppard. 1959. The genetics of *Papilio dardanus*, brown. I. Race Cenea from South Africa. Genetics, 44: 1347-1358.

Cohen, B.H. 1960. ABO-Rh interaction in an Rh-incompatibly mated population. Am. J. Human Genet., 12: 180-209.

Cohen, B.H. and B. Glass. 1959. The relation of the ABO and Rh blood groups to differential reproduction. Am. J. Human Genet., 11: 414-417.

Cooper, D.W. 1968. The use of incomplete family data in the study of selection and population structure in marsupials and domestic animals. Genetics, 60: 147-156.

Creed, E.R., W.H. Dowdeswell, E.B. Ford and K.G. McWhirter. 1959. Evolutionary studies on *Maniola jurtina*. Heredity, 16: 363-391.

Crow, F.J. and M. Kimura. 1970. An introduction to population genetics theory. Harper and Row, New York.

Dessauer, H.C. and E. Nevo. 1969. Geographic variation of blood and liver proteins in cricket frogs. Biochem. Genet., 3: 171-188.

Dobzhansky, T. and H. Levene. 1948. Genetics of natural populations. XVII. Proof of operation of natural selection in wild populations of *Drosophila pseudoobscura*. Genetics, 33: 537-547.

Dowdeswell, W.H. 1961. Experimental studies on natural selection in the butterfly *Maniola jurtina*. Heredity, 16: 39-52.

Dubos, R. 1965. Man adapting. Yale University Press.

Dunn, L.C. 1957. Evidence of evolutionary forces leading to the spread of lethal genes in wild populations of house mice. Nat. Acad. Sci. (U.S.), Proc. 43: 158-163.

Dunn, L.C. 1965. Analysis of a complex gene in the house mouse. Cold Spring Harbor Symp. Quant. Biol., 21: 187-195.

Dunn, L.C. and J. Suckling. 1956. Studies of genetic variability in populations of wild house mice. I. Analysis of seven alleles at locus *T*. Genetics, 41: 344-352.

Dunsgord, I. and C.C. Bowley. 1967. Techniques in blood grouping. Oliver and Boyd, Edinburgh.

Fisher, R.A. 1930. The genetical theory of natural selection. Clarendon Press, Oxford.

Ford, E.B. 1940. Polymorphism and taxonomy. *In* J. Huxley (ed.), The new systematics. Clarendon Press, Oxford.

Ford, E.B. 1964. Ecological genetics, Methuen, London.

Fuller, W.A. 1969. Changes in numbers of three species of small rodent near Great Slave Lake, N.W.T., Canada, 1964-1967, and their significance for general population theory. Zool. Fennici, Ann. 6: 113-144.

Giblett, E.R. 1962. The plasma transferrins. Progr. Med. Genet., 2: 34-63.

Gilmour, D.G. 1959. Segregation of genes determining red cell antigens at high levels of inbreeding in chickens. Genetics, 44: 14-33.

Haldane, J.B.S. 1932. The causes of evolution. Harper, London.

Haldane, J.B.S. 1937. The effect of variation on fitness. Am. Naturalist, 71: 337-349.

Haldane, J.B.S. 1949. Disease and evolution. Ricerca sci, 19, Supplement 68.

Haldane, J.B.S. 1957. The cost of natural selection. J. Genet., 55: 511-524.

Hardy, G.H. 1908. Mendelian proportions in a mixed population. Science, 28: 49-50.

Harris, H. 1959. Human biochemical genetics. Cambridge University Press.

Harris, H. 1969. Enzyme and protein polymorphism in human populations. Brit. Med. Bull., 25: 5-13.

Hartl, D.L. 1970a. Population consequences of non-Mendelian segregation among multiple alleles. Evolution, 24: 415-423.

Hartl, D.L. 1970b. Analysis of a general population genetic model of meiotic drive. Evolution, 24: 538-545.

Heilmeyer, L., W. Keller, O. Vivell, W. Keiderling, K. Betke, F. Wöhler and H.E. Schultze. 1961. Kongenitale atransferrinämie bei einem sieben Jahre alten kind. Dt. med. Wschr., 86: 1745.

Hiraizumi, Y., L. Sandler and J.F. Crow. 1960. Meiotic drive in natural populations of *Drosophila melanogaster*. III. Populational implications of the segregation-distorter locus. Evolution, 14: 433-444.

Ingram, V.M. 1957. Gene mutations in human hemoglobin: the chemical difference between normal and sickle-cell hemoglobin. Nature, 180: 326.

Johnson, F.M., H.E. Schaffer, J.E. Gillaspy and E.S. Rockwood. 1969. Isozyme genotype-environment relationships in natural populations of the harvester ant, *Pogonomymex barbatus*, from Texas. Biochem. Genet., 3: 429-450.

Kettlewell, H.B.D. 1957. The contribution of industrial melanism to our knowledge of evolution. Adv. Sci., 52: 245-252.

Kimura, M. and J.F. Crow. 1964. The number of alleles that can be maintained in a finite population. Genetics, 49: 725-738.

Kimura, M. and T. Ohta. 1971. Protein polymorphism as a phase of molecular evolution. Nature, 229: 467-469.

King, J.A. 1967. Behavioral modification of the gene pool. *In* J. Hirsch (ed.), Behavior-genetic analysis. McGraw-Hill, Toronto.

King, J.A. (ed.) 1968. Biology of *Peromyscus* (Rodentia). Special publication No. 2, The American Society of Mammalogists.

King, J.L. 1967. Continuously distributed factors affecting fitness. Genetics, 55: 483-492.

King, J.L. and T.H. Jukes. 1969. Non-Darwinian evolution. Science, 164: 788-798.

Koehn, R.K., J.E. Perez and R.B. Merritt. 1971. Esterase enzyme function and genetical structure of populations of the freshwater fish *(Notropis stramineus)*. Am. Naturalist, 105: 51-69.

Koehn, R.K. and D.I. Rasmussen. 1967. Polymorphic and monomorphic serum esterase heterogeneity in catostomid fish populations. Biochem. Genet., 1: 131-144.

Levin, B.R., M.L. Petras and D.I. Rasmussen. 1969. The effect of migration on the maintenance of a lethal polymorphism in the house mouse. Am. Naturalist, 103: 647-661.

Levine, P. 1958. The influence of the ABO system on Rh hemolytic disease. Human Biology, 30: 14-28.

Lewontin, R.C. 1967a. Population genetics. Ann. Rev. Genet., 1: 37-70.

Lewontin, R.C. 1967b. An estimate of average heterozygosity in man. Am. J. Human Genet., 19: 681-685.

Lewontin, R.C. 1968. The effect of differential viability on the population dynamics of *t* alleles in the house mouse. Evolution, 22: 262-273.

Lewontin, R.C. and L.C. Dunn. 1960. The evolutionary dynamics of a polymorphism in the house mouse. Genetics, 45: 705-722.

Lewontin, R.C. and J.L. Hubby. 1966. 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.

Li, C.C. 1955a. Population genetics. University of Chicago Press.

Li, C.C. 1955b. The stability of an equilibrium and the average fitness of a population. Am. Naturalist, 89: 281-296.

Li, C.C. 1969. Population subdivision with respect to multiple alleles. Ann. Human Genet., 33: 23-29.

Loegering, W.Q. and E.R. Sears. 1963. Distorted inheritance of stem-rust inheritance of timstein wheat caused by a pollen-killing gene. Can. J. Genet. Cytol., 5: 65-72.

Lush, I.E. 1966. The biochemical genetics of vertebrates except man. North-Holland Publishing Co., Amsterdam.

Malecha, S.R. and R.H. Tamarin. 1969. Plasma transferrins in three species of *Rattus* on Pacific islands. Am. Naturalist, 103: 664-669.

Manwell, C. and C.M.A. Baker. 1970. Molecular biology and the origin of species: Heterosis, protein polymorphism, and animal breeding. University of Washington Press, Seattle.

Milkman, R.D. 1967. Heterosis as a major cause of heterozygosity in nature. Genetics, 55: 493-495.

Neel, J.V., F.M. Salzano, P. Jungueira, F. Keitev and D. Maybury-Lewis. 1964. Studies on the Xavante Indians of the Brazilian Mate Grosso. Am. J. Human Genet., 16: 52-140.

Pauling, L., H.A. Itano, S.J. Singer and I.C. Wells. 1949. Sickle-cell anemia, a molecular disease. Science, 110: 543.

Petras, M.L. 1967. Studies of natural populations of *Mus*. I. Biochemical polymorphisms and their bearing on breeding structure. Evolution, 21: 259-274.

Poulik, M.D. 1957. Starch-gel electrophoresis in a discontinuous system of buffers. Nature, 180: 1477-1479.

Prout, T. 1953. Some effects of variations in the segregation ratio and of selection upon the frequency of alleles. Acta Genet. et Statist. Med., 4: 148-151.

Race, R.R. and R. Sanger. 1968. Blood groups in man. 5th edition. Blackwell, Oxford.

Rasmussen, D.I. 1961. Erythrocytic antigenic differences between individuals of the deer mouse, *Peromyscus maniculatus*. Genet. Res. (Camb.), 2: 449-455.

Rasmussen, D.I. 1964. Blood group polymorphism and inbreeding in natural populations of the deer mouse *Peromyscus maniculatus*. Evolution, 18: 219-229.

Rasmussen, D.I. 1968. Genetics. *In* J.A. King (ed.), Biology of *Peromyscus* (Rodentia). Special Publication No. 2, The American Society of Mammalogists.

Rasmussen, D.I. 1969. Molecular taxonomy and typology. Bioscience, 19: 418-420.

Rasmussen, D.I. and R.K. Koehn. 1966. Serum transferrin polymorphism in the deer mouse. Genetics, 54: 1353-1357.

Richmond, R.C. and J.R. Powell. 1970. Evidence of heterosis associated with an enzyme locus in a natural population of *Drosophila*. Nat. Acad. Sci. (U.S.), Proc. 67: 1264-1267.

Robertson, A. 1965. The interpretation of genotypic ratios in domestic animal populations. Anim. Prod., 7: 319-324.

Savage, E.A. 1971. The description of a blood group system in deer mice *(Peromyscus maniculatus)*. M.Sc. Thesis, University of Alberta.

Savage, E.A. and D.G. Cameron. 1971. Blood group complexity: the *Pm* locus in *Peromyscus maniculatus*. Anim. Blood Grps biochem. Genet., 2: 23-29.

Selander, R.K. 1970. Behavior and genetic variation in natural populations. Am. Zoologist, 10: 53-66.

Selander, R.K., W.G. Hunt and S.Y. Yang. 1969. Protein polymorphism and genic heterozygosity in two European subspecies of the house mouse. Evolution, 23: 379-390.

Selander, R.K. and S.Y. Yang. 1969. Protein polymorphism and genic heterozygosity in a wild population of the house mouse *(Mus musculus)*. Genetics, 63: 653-667.

Selander, R.K., S.Y. Yang, R.C. Lewontin and W.E. Johnson. 1970. Genetic variation in the horseshoe crab *(Limulus polyphemus)*, a phylogenetic "relic". Evolution, 24: 402-414.

Semeonoff, R. and F.W. Robertson. 1968. A biochemical and ecological study of plasma esterase polymorphism in natural populations of the field vole, *Microtus agrestis* L. Biochem. Genet., 1: 205-227.

Sheppard, P.M. 1959. Natural selection and heredity. Hutchinson, London.

Smithies, O. 1959. An improved procedure for starch-gel electrophoresis: further variation in the serum proteins of normal individuals. Biochem. J., 71: 585-587.

Sokal, R.R. and F.J. Rohlf. 1969. Biometry. Freeman, San Francisco.

Soper, J.D. 1964. The mammals of Alberta. Hamly Press Ltd., Edmonton.

Stickel, L.F. 1968. Home range and travels. *In* J.A. King (ed.), Biology of *Peromyscus* (Rodentia). Special Publication No. 2, The American Society of Mammalogists.

Sved, J.A., T.E. Reed and W.F. Bodmer. 1967. The number of balanced polymorphisms that can be maintained in a natural population. Genetics, 55: 469-481.

Tamarin, R.H. and C.J. Krebs. 1969. *Microtus* population biology. II. Genetic changes at the transferrin locus in fluctuating populations of two vole species. Evolution, 23: 183-211.

Terman, C.R. 1968. Population dynamics. *In* J.A. King (ed.), Biology of *Peromyscus* (Rodentia). Special Publication No. 2, The American Society of Mammalogists.

Van Valen, L. 1963. Haldane's dilemma, evolutionary rates, and heterosis. Am. Naturalist, 97: 185-190.

Wallace, B. 1970. Genetic load. Prentice-Hall, Inc., Englewood Cliffs, N.J.

Weinberg, W. 1908. Über den Nachweis der Vererbung beim Menschen. Jh. Ver. vaterl. Naturk. Württemb., 64: 369-382.

Wright, S. 1931. Evolution in Mendelian populations. Genetics, 16: 97-159.

Wright, S. 1943. Isolation by distance. Genetics, 28: 114-138.

Wright, S. 1946. Isolation by distance under diverse systems of mating. Genetics, 31: 39-59.

Wright, S. 1951. The genetical structure of populations. Ann. Eugenics, 15: 323-354.

APPENDIX I

COMPUTER PROGRAMS

The following programs are written in FOCAL, for use in a DEC PDP-8.

Program A (page 109). G-test (a rows X columns test of heterogeneity). This test is described in detail by Sokal and Rohlf (1970), who give a more extensive program written in FORTRAN. In order to show how it was adapted for use in a small computer, this program is given here.

Enter: Numbers of rows, followed by each row. Note that this program is for three columns, but can easily be modified.

Printout: G (total), followed by the value of G for each subset formed by sequentially removing each row.

Program B (page 110). Population simulation: 2 alleles; segregation distortion in one sex; selection against both homozygotes.

Enter: Initial genotypic proportions (P, Q, R); segregation distortion constant (K); values of selection coefficients against the homozygotes (J, against the homozygote favored by the distortion, and H).

Printout: Each generation, the generation number and new values of the allele frequencies are printed. When the change in allele frequency in one generation is less than 0.00001, the simulation ceases and "EQUILIBRIUM REACHED" is printed.

Program C (page 111). Population simulation; 2 alleles; segregation distortion in the heterozygote-heterozygote cross; selection against both homozygotes.

Enter: Initial allele frequencies (P,Q); segregation distortion constant (K); values of selection coefficients against the homozygotes (SO, against the homozygote favored by the distortion, and ST).

Printout: Same as Program B.

Program D (page 112). Population simulation; 3 alleles; segregation distortion in one heterozygote and one sex; selection against all homozygotes and the heterozygote for the two alleles which are not at a disadvantage because of the distortion.

Enter: Initial allele frequencies (P,Q,R); segregation distortion constant (K); selection against the homozygote favored by the distortion (SO), the homozygote at a disadvantage because of the distortion (ST), the homozygote not affected by the distortion (SH), and the heterozygote for the two alleles which are not at a disadvantage because of the distortion (SJ).

Printout: Each generation, the generation number and the allele frequencies are printed. When the change in all allele frequencies is less than 0.0001, the simulation ceases and "EQUILIBRIUM REACHED" is printed. If the generation number reaches 200, the simulation stops.

Program A

```
C-FOCAL,1969

01.10 E
01.20 A ,SS,!F X=1,1,SS!D 3
01.50 S LN=(SA+SB+SC)*FLOG(SA+SB+SC)
01.60 S GH=2*(SL+SM+SN+LN-SP-SA*FLOG(SA)-SB*FLOG(SB)-SC*FLOG(SC))
01.70 T %,GH,!!F R=1,1,SS!D 4
01.80 Q

03.10 A A(X),B(X),C(X)!S SL=SL+A(X)*FLOG(A(X))
03.20 S SM=SM+B(X)*FLOG(B(X))!S SN=SN+C(X)*FLOG(C(X))
03.50 S SP=SP+(A(X)+B(X)+C(X))*FLOG(A(X)+B(X)+C(X))
03.60 S SA=SA+A(X)!S SB=SB+B(X)!S SC=SC+C(X)!T !

04.10 S SR=SL-A(R)*FLOG(A(R))+SM-B(R)*FLOG(B(R))
04.20 S Y=A(R)+B(R)+C(R)!S LP=(SA+SB+SC-Y)*FLOG(SA+SB+SC-Y)
04.30 S LD=(SA-A(R))*FLOG(SA-A(R))+(SB-B(R))*FLOG(SB-B(R))
04.36 S LG=(SC-C(R))*FLOG(SC-C(R))
04.40 S GG=2*((SR+SN-C(R)*FLOG(C(R)))+LP-SP+Y*FLOG(Y)-LD-LG)!T ,GG,!
*
```

Program B

```
C-FOCAL,1969

01.10 E
01.20 A P,Q,R,K,J,H,!
01.30 S L=1-K; S Z=0
01.35 S T=X
01.40 S AA=(P+Q*K)*(P+Q*0.5)*(1-J)
01.50 S AB=(P+Q*K)*(R+Q*0.5)+(P+Q*0.5)*(R+Q*L)
01.60 S BB=(R+Q*L)*(Q*0.5+R)*(1-H)
01.70 S N=AA+AB+BB
01.80 S P=AA/N; S Q=AB/N; S R=BB/N
01.90 S X=P+Q/2; S Y=R+Q/2; S Z=Z+1

02.30 T %3.00,"GEN #",Z,
02.50 T %4.03,"   P",X,"   Q",Y,!
02.70 I ((FABS(X-T))-0.00001) 4.1,4.1,1.35

04.10 T "EQUILIBRIUM REACHED"
*
```

Program C

```
C-FOCAL,1969

01.10 E
01.30 A P,Q,K,SO,ST,!; S Z=0
01.35 S JJ=P↑2*Q↑2+P↑4*SO+2*P↑3*Q*SO+4*P↑2*Q↑2*K↑2*SO
01.36 S X=P
01.40 S AA=P↑2+4*P↑2*Q↑2*K↑2-JJ
01.50 S KK=Q↑4*ST+4*P↑2*Q↑2*(1-K)↑2*ST
01.51 S KK=KK+2*P*Q↑3*ST
01.60 S AB=2*P*Q-8*P↑2*Q↑2*K*(1-K)+2*P↑2*Q↑2
01.70 S BB=Q↑2+P↑2*Q↑2-4*P↑2*Q↑2*(1-K)↑2-KK
01.80 S N=AA+AB+BB
01.90 S XX=AA/N; S YY=AB/N; S ZZ=BB/N

02.10 S Z=Z+1; T %3.00, "GEN #",Z,
02.20 S P=XX+YY/2; T %4.03, "  P",P,
02.30 S Q=ZZ+YY/2; T %4.03, "  Q",Q,!
02.31 I ((FABS(X-P))-0.00001) 4.1,4.1,1.35

04.10 T "EQUILIBRIUM REACHED"!; Q
*
```

Program D

```
C-FOCAL,1969

01.10 E
01.20 A X,Y,Z,K,SO,ST,SH,SJ,!
01.25 S U=0
01.30 S J=X;S L=Y; S M=Z;
01.40 S PP=X↑2; S QQ=Y↑2; S RR=Z↑2
01.50 S PQ=2*X*Y; S QR=2*Y*Z; S PR=2*X*Z
01.60 S NP=(PP+PQ*K+PR/2)*(PP+PQ/2+PR/2)*(1-SO)
01.70 S NQ=(QQ+QR/2+PQ*(1-K))*(QQ+QR/2+PQ/2)*(1-ST)
01.80 S NR=((RR+QR/2+PR/2)↑2)*(1-SH)
01.90 S XQ=(PP+PQ/2+PR/2)*(QQ+QR/2+PQ*(1-K))
01.93 S HQ=XQ+(PP+PQ*K+PR/2)*(QQ+QR/2+PQ/2)
01.95 S XR=(QQ+QR/2+PQ/2)*(RR+QR/2+PR/2)
01.97 S HR=XR+(QQ+QR/2+PQ*(1-K))*(RR+QR/2+PR/2)

02.10 S YR=(PP+PQ/2+PR/2)*(RR+QR/2+PR/2)
02.13 S UR=(YR+(PP+PQ*K+PR/2)*(RR+QR/2+PR/2))*(1-SJ)
02.20 S N=NP+NQ+NR+HQ+HR+UR
02.30 S X=(NP+HQ/2+UR/2)/N; S Y=(NQ+HR/2+HQ/2)/N
02.40 S Z=(NR+HR/2+UR/2)/N; S U=U+1
02.50 T %3.00,"GEN #",U,
02.55 T %4.03,"   P",X,"   Q",Y,"   R",Z,!
02.60 I ((FABS(J-X))-0.0001) 2.7,2.7,1.3
02.70 I ((FABS(L-Y))-0.0001) 2.8,2.8,1.3
02.80 I ((FABS(M-Z))-0.0001) 4.1,4.1,1.3
02.90 I (U-200) 1.3,4.3,4.3

04.10 T "EQUILIBRIUM REACHED"
04.12 T %3.00,"   GEN #",U,!
04.13 T %4.03,"P",X,!,"Q",Y,!,"R",Z,!
04.15 Q
04.30 T "  GEN #=200"; Q
*
```