Theoretical and Applied Genetics 38, 264–270 (1968)

The Role of Finite Population Size and Linkage in Response to Continued Truncation Selection

II. Dominance and Overdominance*

A. W. QURESHI**

Iowa State University, Ames, Iowa

Summary. In a direct Monte Carlo study, the effects and interactions of parental population size, close linkages between loci, degree of truncation selection and environmental variance are explored in the cases of complete dominance and overdominance with equilibrium gene frequency of .5. The base population simulated has a gene frequency of .5 at all 40 loci and is a state of gametic equilibrium. Response and fixation of genes are studied up to 30 generations.

In the case of complete dominance of desirable genes, long term response is described in terms of an initial phase of change and a later phase when there is a very slow approach to the limit and plateaus are established in the case of close linkages between loci. The change in the initial phase constitutes most of the total response. Increased selection in parental populations of 16 or less is ineffective when recombination value between adjacent loci is as low as .005. Linkage effects in reducing response appear to increase in geometric proportions with restricted population size.

Results of the overdominance model provide an interesting study of the effects of population size and linkage on heterozygosity while there is selection for the heterozygotes. These effects are partially explainable in terms of differences in heterozygosity between chromosomes and heterozygosity within chromosomes. The between chromosome component of heterozygosity appears to be rapidly reduced with finite population size and common heterozygosity within chromosomes is conserved due to linkages between loci. In general, however, the effects of selection pressure, linkage and population size interact strongly. With small populations and tight linkage, increased selection pressure causes a decrease in amount of fixation. With the largest population increased selection pressure tends to increase the amount of fixation. It is clear, however, that these conclusions must be regarded tentatively, and an exact picture of what is happening will be obtained only with very extensive simulation. It seems quite clear, also, that single locus finite population theory is essentially ineffective.

Introduction

In the first paper of this series of Monte Carlo studies (QURESHI et al., 1968), the process of continued selection was examined for additive genes. When selecting for a quantitative trait in plants or animals we are probably dealing with a large number of genes at many more loci than the number of chromosomes in the species. We are also confronted with gene action involving a variety of dominance and epistatic effects. In complex multi-locus situations with linkage and finite population size, a study of continued selection has to be conditional upon a given gene action and initial gene frequency.

When selection is based on the performance of a diploid individual and the zygotic phases in its genotype exhibit dominance, population size effects on response involve not only the change in gene frequency but also in the heterozygosity as discussed by KOJIMA (1961). A more interesting situation is encountered when the heterozygote is superior to both homozygotes. Population size and selection effects with respect to chance of fixation of a gene and maintenance of heterozygosity have been examined for this case by ROBERTSON (1962). HAYMAN and MATHER (1953), and REEVE (1955), have dealt with the problem of change in heterozygosity in the case of particular mating systems and selection for heterozygotes at a single locus. In multi-locus systems effects of linkage and gametic disequilibria also need to be considered.

This paper deals with a genetic model where gene effects at 40 loci are similar and additive but the genotypic value of a heterozygote at each locus is not equal to the mean of the genotypic values of the two homozygotes. Two special cases of dominance are considered, complete dominance of desirable genes, and pure overdominance with the genotypic values of the homozygotes equal. The matters of interest are the changes in the whole polygenic system arising from truncation selection and the dependence of these changes on population size, selection intensity, and the linkage of the several loci. The results presented here on these aspects are of exploratory nature and include data presented earlier in duplicated form (QURESHI, 1964).

Structure of Simulated Populations

The populations consist of diploid individuals of separate sexes and mating is random among selected parents. The number of parents in a population is constant over the generations, and the number of progeny generated from these depends upon the degree of truncation selection specified. The genotypes consist of 40 loci distributed equally on four chromosomes. Linkages are uniform between adjacent loci on a chromosome and there is no interference. Two alleles are simulated at each locus, the favorable allele is referred to as the plus allele. The initial population is generated by random mating of the parents that are heterozygous at all loci. The mean gene frequency is therefore .5 initially and the

* Journal Paper No. T-5905 Iowa Agriculture and Home Economics Experiment Station, Ames, Iowa. Supported by National Science Foundation Grant 19218 and National Institute of Health Grant GM 13827.

** On leave from West Pakistan Agricultural University, Lyallpur, West Pakistan.

expected numbers of coupling and repulsion phases are equal.

The genotypic value (g) of an individual is calculated as follows:

$$\text{Dominance model: } g = n_1 + n_2$$
$$\text{Overdominance model: } g = n_1$$

where n_1 is the number of $+-$ phases and n_2 is the number of $++$ phases in the genotype. The infinite population equilibrium for the dominance model has gene frequency equal to 1 and has gene frequency of .5 for the overdominance model. The environmental component of the phenotypic value is independent of the genotype and is constant over the generation. The simulation procedure of the above genetic model is similar to QURESHI et al. (1968).

The following levels of parental population size (N), recombination value between adjacent loci on a chromosome (r), degree of truncation selection (b), and the environmental variance (σ_e^2) are varied in different populations according to a 3^4 factorial design so that various effects and interactions are estimated.

$$N = 8, 16, 64$$
$$r = .005, .05, .5$$
$$b = \frac{1}{2}, \frac{1}{4}, \frac{1}{8}$$
$$\sigma_e^2 = 0,\ 2\,\sigma_{g0}^2,\ 8\,\sigma_{g0}^2$$

σ_{g0}^2 is the expected genotypic variance in the initial population. The selection process is continued for 30 generations, or less if fixation occurred earlier.

Selection pressure over the generations is determined by the degree of truncation and heritability. The heritability in the narrow sense, $h^2 = \sigma_A^2/(\sigma_g^2 + \sigma_e^2)$ which pertains to change in gene frequency, p, depends on the additive component of the genotypic variance, σ_A^2, and changes with changing p due to selection and random drift in a particular line. Following are the HARDY-WEINBERG expectations of h^2 for the three levels of σ_e^2 and a given p:

	σ_e^2	0	$2\,\sigma_{g0}^2$	$8\,\sigma_{g0}^2$
Dominance:				
	$p = .5$	.667	.222	.074
	.6	.571	.151	.047
	.7	.462	.083	.024
	.8	.333	.031	.008
	.9	.182	.005	.001
Overdominance:				
	.5	.000	.000	.000
	.6	.077	.026	.009
	.7	.276	.090	.030
	.8	.529	.161	.052
	.9	.780	.178	.053

The frequency of a dominant gene at a locus is expected to increase with selection, but in the case of overdominance a change towards equilibrium frequency is brought about with selection if p deviates from the equilibrium frequency due to random drift. Besides the changes in p, the effects of departure from HARDY-WEINBERG structure also need to be considered in order to determine the heritability and the effectiveness of selection over generations in the populations specified here. These effects include reduction in heterozygosity due to random drift and change in genotypic variance due to gametic disequilibria.

Quantitative Response

The genotypic mean in generation t, denoted by m_t ($t = 0, 1, \ldots$), is a function of the average number of various genotypic phases in the population. The population has a maximum genotypic mean when there is 100% heterozygosity at each locus. In the case of complete dominance the maximum genotypic mean is also attained when only plus genes are fixed in the population. For the purpose of standardization, response in generation t is defined in terms of percentage of maximum possible response as:

$$R_t = \frac{m_t - m_0}{m_{\max} - m_0} \times 100; \quad t = 0,1,\ldots, \tag{1}$$

where $m_{\max}$ denotes the maximum genotypic mean. If we denote the average frequency of the $+$ gene in generation t by p_t and the average frequency of $+-$ phases in generation t by $2\,Q_t$, R_t in the case of complete dominance may also be written as

$$R_t = \frac{(p_t - p_0) - [p_0\,(1 - p_0) - Q_t]}{(1 - p_0)^2} \times 100\,, \tag{2}$$

and in the case of overdominance as

$$R_t = \frac{2\,[Q_t - p_0\,(1 - p_0)]}{p_0^2 + (1 - p_0)^2} \times 100\,. \tag{3}$$

It may be seen from (2) and (3) that, with overdominance, R_t is a function only of change in the heterozygosity whereas with complete dominance it also depends upon the change in gene frequency.

Complete Dominance

Since the initial gene frequency, p_0, is simulated as .5, R_t is expected to vary between $+100$ and -100 depending upon the changes in population structure due to population size, linkage and selection. The positive values of R_t indicate net effectiveness of selection, whereas the negative values show that the depressing effects of population size and linkage are predominant.

Population size effects on response may be explained in terms of decrease in heterozygosity and increase in sampling error of the process. Random drift in mean gene frequency over 40 loci is expected to be zero but the drift at individual loci contributes to the variance. Degree of truncation, environmental variance, dominance variance and the additive variance may be pointed out as factors that determine the effectiveness of selection. Since degree of truncation and environmental variance are constant in each generation, any non-linearity of response over the generations is attributed to change in the additive and the dominance variance.

A considerable amount of sampling error is expected in the results since the data comprise a single replication. However, the graphs presented in Figures 1 to 3 indicate the general features of relationships. The main effects realized under the specified levels of N, r, b, and σ_e^2 are quite obvious from the figures. A major purpose of this presentation is to point out certain interactions that occur between these factors. Analyses of variance of the genotypic means in different generations were found to be inconsistent and inconclusive for this purpose be-

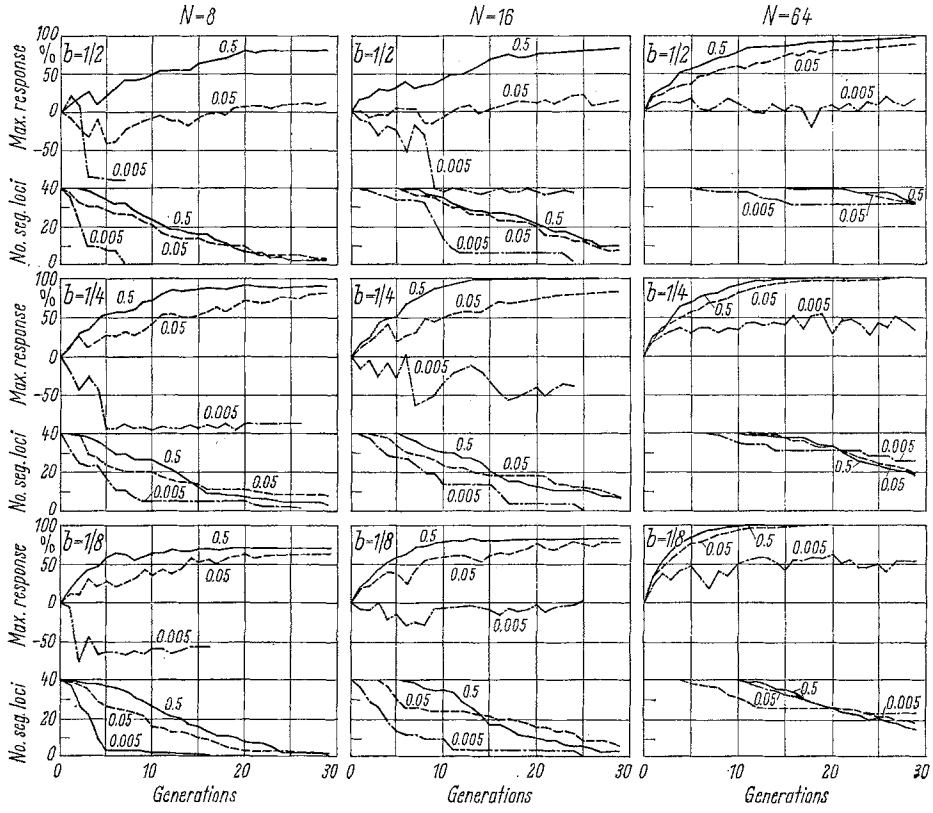

Fig. 1. Dominance model, $\sigma_e^2 = 0$: R_t and number of segregating loci in each generation for various conditions of N, b and r

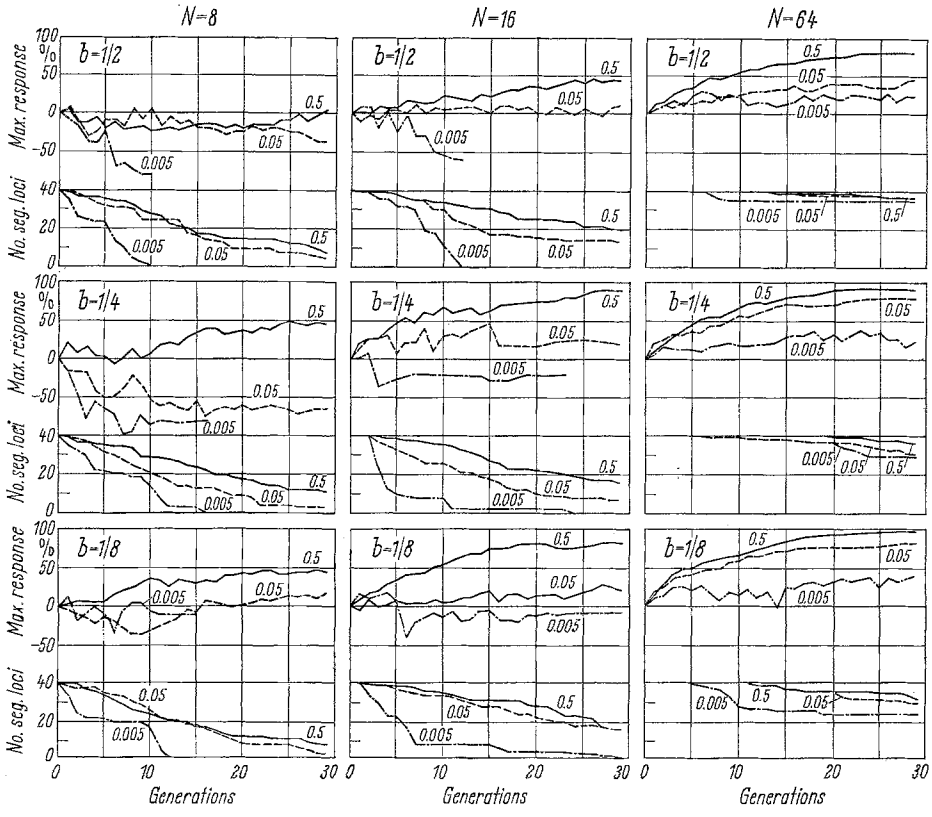

Fig. 2. Dominance model, $\sigma_e^2 = 2\sigma_{g0}^2$: R_t and number of segregating loci in each generation for various conditions of N, b and r

cause of sampling errors. A graphic comparison of the response curves presented in the figures is more informative.

When there is free recombination between loci, response curves can be described in terms of an initial phase of change and a later phase when there is slow approach to the limit. The rate of response in the initial phase determines the total response to a large extent. The decrease due to low N in the rate of response is smaller when selection is more intense. Population size effects are not as pronounced in the initial phase of response as they are in later generations. With increased truncation selection and no environmental variance, there is a large increase in R_t in the first few generations and a continued plateau is seen in later generations until all loci are fixed. With large environmental variance or low heritability, there is an increasing but slower response till fixation. The fluctuations in response curves in small parental populations are increased with environmental variance. Increased level of truncation appears to result in a smoother response over generations even when the environmental variance is very large and population size is as low as 8. When parental population size is as large as 64, maximum response appears to be attainable with very low selection pressure. With free recombination, only in the case of an initial heritability of .07, N equal to 8 and b equal to 1/2 is a negative value of R_t obtained after 30 generations of selection.

Close linkages between loci, as specified in this study, have a marked effect on the rate of response even when parental population is as large as 64. Two obvious patterns of response are seen in this regard: (1) increased rate of fixation of genes due to closer linkage resulting in a decline in R_t; (2) when heterozygosity or the number of segregating loci is maintained in the population due to selection, a phase of plateaued response is noticed. With r equal to .05, maximum response appears to be attainable in large populations in spite of a slower rate of response. In

small populations and low selection pressure a definite ceiling is imposed. Fixation of clusters of genes is seen when r as low as .005 is simulated. Thus, a negative R_t is seen in all cases except when N is 64. In parental populations of 8 there is a very rapid fixation of all loci making selection almost ineffective.

In larger populations with close linkage, a plateau in response is seen in later generations after an initial change. During the plateaued phase there is no fixation of genes. It appears that in these situations selection is effective only in maintaining heterozygosity without causing any change in response. COMSTOCK (1961) has reported a similar situation where linked dominant genes show "operational overdominance". The data in this study are not adequate to discuss these plateaus in terms of the genetic structure of the population. In a detailed MonteCarlo study, QURESHI and KEMPTHORNE (1968) have shown that in a multilocus situation, gametic disequilibria, arising due to selection and finiteness of population size, increase the dominant component of the variance and decrease the additive component. Hence, selection is ineffective in bringing about response although variability in the population is large. Linkages between loci tend to conserve the gametic disequilibria and thereby cause an "operational plateau" in the population.

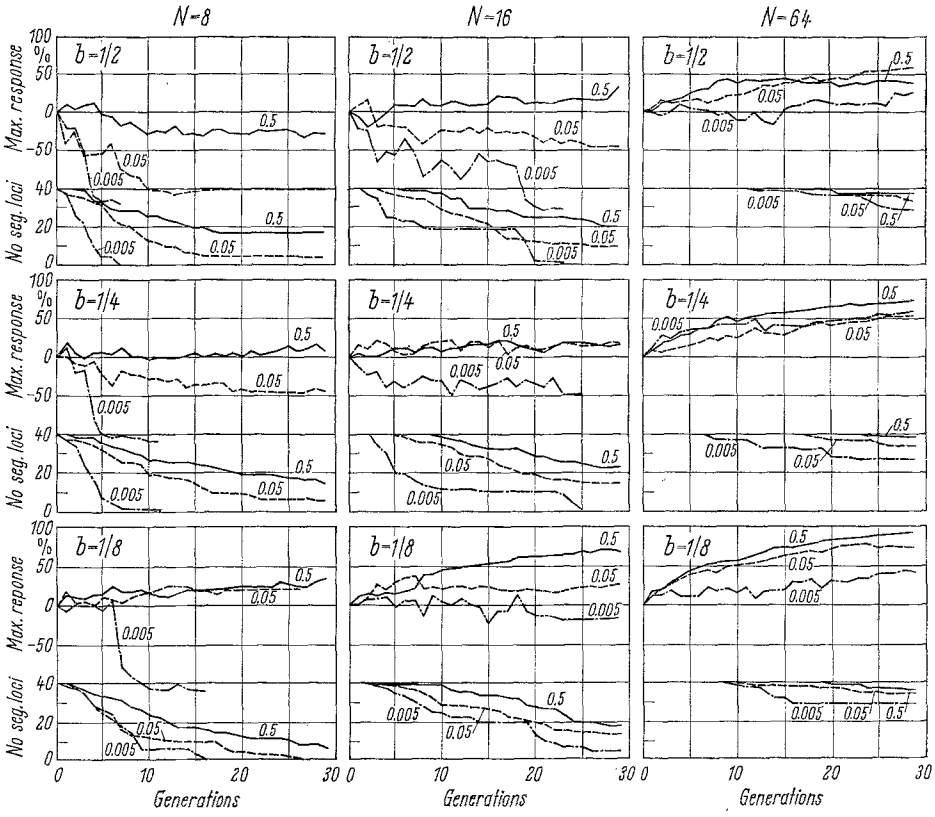

Fig. 3. Dominance model, $\sigma_z^2 = 8\sigma_{g0}^2$: R_t and number of segregating loci in each generation for various conditions of N, b and r

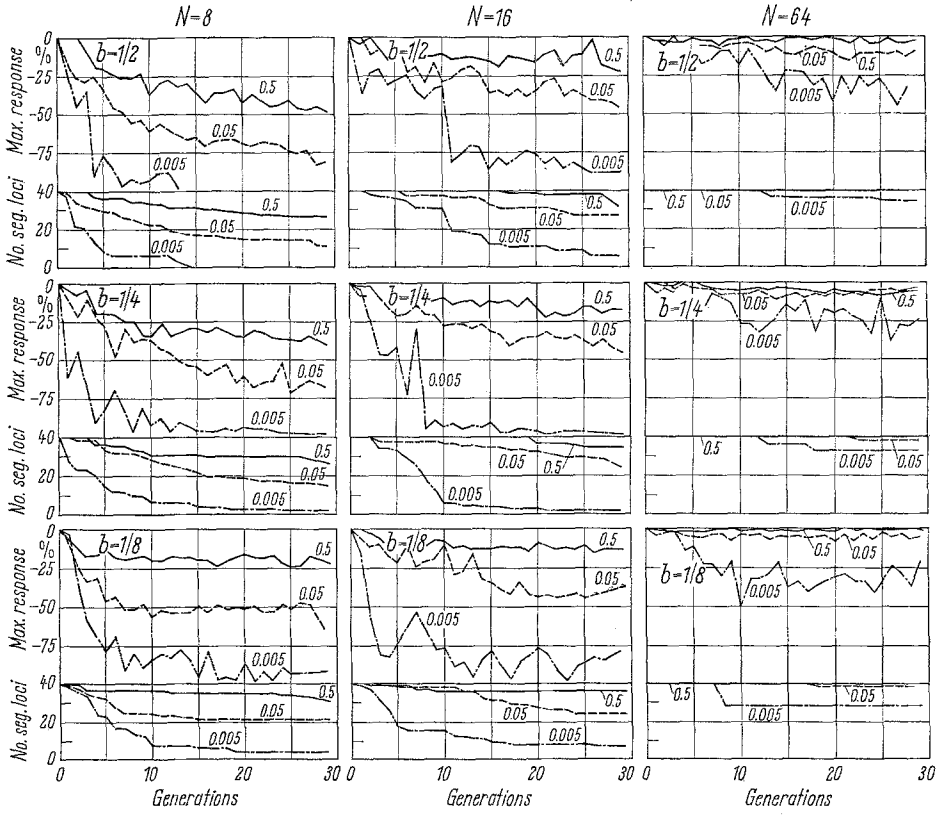

Fig. 4. Overdominance model, $\sigma_z^2 = 0$: R_t and number of segregating loci in each generation for various conditions of N, b, and r

Overdominance

Response curves for this model, presented in Figures 4 to 6, indicate the changes in heterozygosity as a net result of operative forces. Since truncation selection of 1/2 or more is simulated in this study, a stable gene frequency equilibrium as well as stable gametic equilibria would be expected in infinite populations irrespective of linkages between loci. In finite populations reduction in heterozygosity is expected over the generations depending upon its size. Also, drift in gene frequency from its equilibrium value and gametic disequilibria in a particular line will result because of sampling nature of the

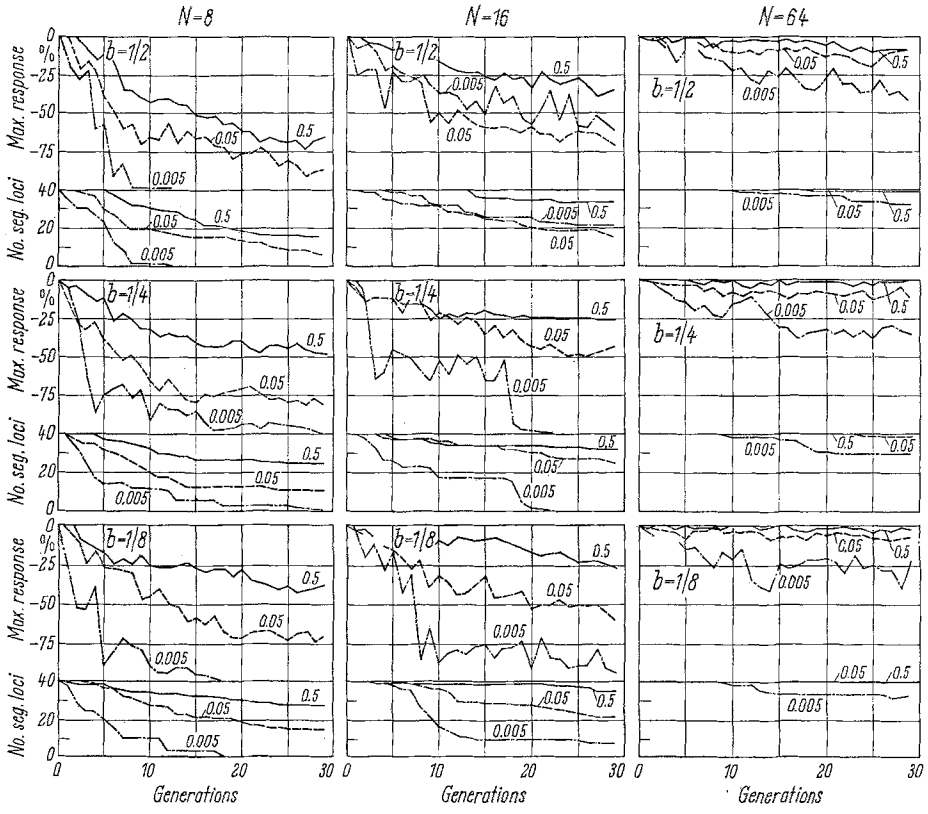

Fig. 5. Overdominance model, $\sigma_e^2 = 2\sigma_{g0}^2$: R_t and number of segregating loci in each generation for various conditions of N, b, and r

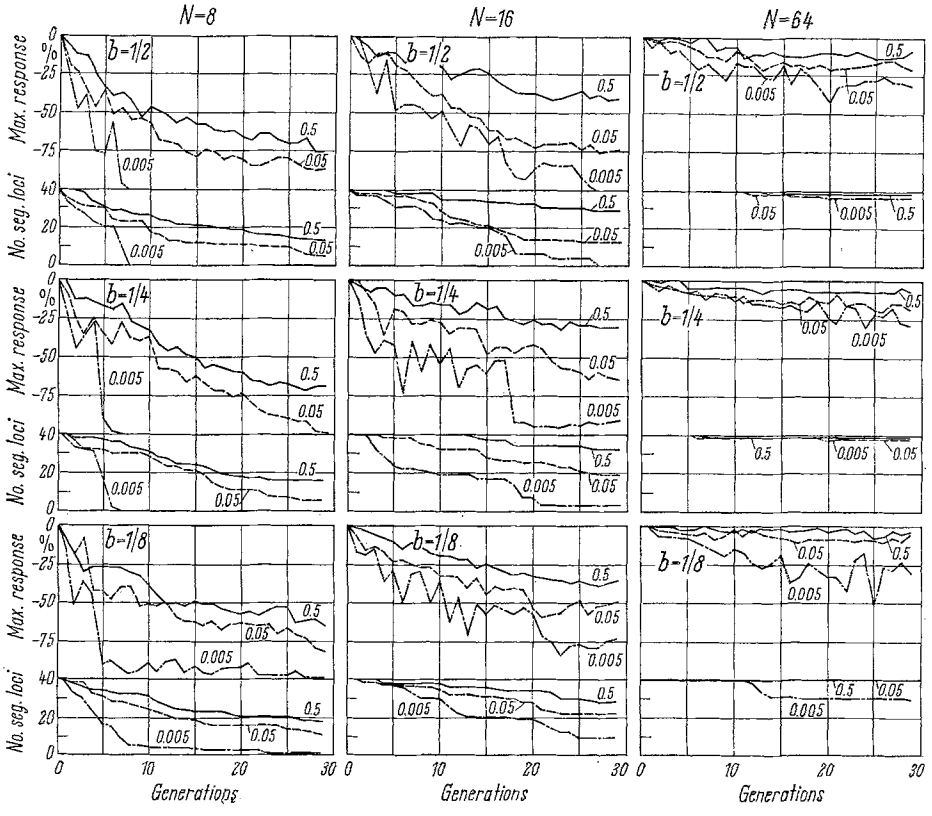

Fig. 6. Overdominance model, $\sigma_e^2 = 8\sigma_{g0}^2$: R_t and number of segregating loci in each generation for various conditions of N, b, and r

process. Selection and linkage effects are thereby involved in altering R_t, although the population is initially at equilibrium.

An examination of the response curves in Figures 4 to 6 shows that the simulated levels of truncation selection have similar effects in very large populations. In parental populations of 8 or 16, increased selection pressure slows down the rate of decrease in heterozygosity when r is equal to .05 or more. Since additive variance is initially zero, selection is operative only when there is a drift in gene frequency and when sufficient variability is released in the population through recombinations. With tight linkages between loci ($r = .005$), increased selection is seen to be ineffective in maintaining heterozygosity. It is apparent that in the extreme case of gene clusters practically no variability is released in small populations even after a number of generations of random mating. Increased selection pressure in this case appears to decrease further the heterozygosity in small populations.

The pattern of response curves in the case of close linkage is that of a rapid drop in R_t in the initial phase of response and a very slow rate of decrease or plateau in the later phase. Close linkage increases the rate of decrease in heterozygosity in small populations since blocks of genes or whole chromosomes are segregated in each generation rather than individual genes. A strong interaction between population size and linkage is also apparent in the graphs. Linkage effects in reducing heterozygosity are very much magnified in parental populations of 16 or less. The decrease in recombination between loci due to population size has been studied by WRIGHT (1933). The results of the present study suggest that the problem of decrease in heterozygosity due to linkage in finite populations deserves more theoretical treatment than is reported in literature.

Fixation of Genes

The trend of fixation of loci that pertains to the rate of response has been discussed in the previous

Table 1. *Dominance model: Mean squares obtained from analysis of variance of the number of inferior genes fixed*

Source	df	Generation 8			Generation 30		
	σ_e^2	0	$2\sigma_{g0}^2$	$8\sigma_{g0}^2$	0	$2\sigma_{g0}^2$	$8\sigma_{g0}^2$
N	2	129.00**	75.4**	181.0**	160.0**	264.0**	435.0**
r	2	131.0**	66.3**	116.0**	268.0**	125.0**	199.0**
b	2	1.4	2.3	18.9	21.4**	14.9*	43.0*
$N \times r$	4	51.2**	16.9	39.1**	54.9**	28.4**	47.1*
$N \times b$	4	4.6*	4.1	6.8	5.2	9.3	6.6
$b \times r$	4	.2	3.3	14.3	7.6	3.4	6.2
$N \times r \times b$	8	.9	5.8	5.2	2.8	2.9	7.5

*, ** → significant at 5% and 1% level when compared with 3-factor interaction m.s.

Table 2. *Dominance model: Number of loci fixed for inferior allele in 30 generations*

		$r = .005$			$r = .05$			$r = .5$		
σ_e^2	b	$N = 8$	16	64	8	16	64	8	16	64
0	$\frac{1}{2}$	19	21	2	9	7	0	2	2	0
	$\frac{1}{4}$	19	14	1	2	2	0	1	0	0
	$\frac{1}{8}$	16	10	0	4	2	0	3	2	0
2	$\frac{1}{2}$	18	16	0	12	6	1	9	3	0
	$\frac{1}{4}$	18	12	0	16	6	0	4	0	0
	$\frac{1}{8}$	11	11	2	8	5	0	4	1	0
8	$\frac{1}{2}$	22	23	2	20	12	0	9	2	1
	$\frac{1}{4}$	21	15	0	14	6	0	6	5	0
	$\frac{1}{8}$	21	11	0	8	4	0	6	2	0

Table 3. *Overdominance model: Mean squares obtained from A.O.V. of the number of fixed loci*

Source	df	Generation 15			Generation 30		
	σ_e^2	0	$2\sigma_{g0}^2$	$8\sigma_{g0}^2$	0	$2\sigma_{g0}^2$	$8\sigma_{g0}^2$
N	2	824.0**	1180.0**	1331.0**	1170.0**	1564.0**	2011.0**
r	2	1039.0**	577.0**	451.0**	1055.0**	779.0**	579.0**
b	2	0.5	0.0	0.0	12.7	10.1	7.0
$N \times r$	4	132.0**	95.0*	88.0**	117.0**	60.9**	77.7**
$N \times b$	4	25.7	20.6	6.6	22.4*	12.9	9.2
$b \times r$	4	4.1	23.1	12.4	3.0	23.7	4.7
$N \times r \times b$	8	10.8	16.0	8.7	4.3	9.7	7.3

*, ** → significant at 5% and 1% level when compared with 3-factor interaction m.s.

Table 4. *Overdominance model: Number of loci fixed within 30 generations*

		$r = .005$			$r = .05$			$r = .5$		
σ_e^2	b	$N = 8$	16	64	8	16	64	8	16	64
0	$\frac{1}{2}$	40	25	6	31	14	0	16	6	0
	$\frac{1}{4}$	39	39	7	26	15	8	13	6	2
	$\frac{1}{8}$	36	33	12	19	16	1	8	4	0
2	$\frac{1}{2}$	40	25	8	34	24	1	25	7	1
	$\frac{1}{4}$	40	40	11	30	16	2	17	8	0
	$\frac{1}{8}$	40	33	9	26	19	0	15	6	0
8	$\frac{1}{2}$	40	39	3	34	36	1	25	10	1
	$\frac{1}{4}$	40	37	7	35	22	1	24	8	2
	$\frac{1}{8}$	40	31	11	30	18	0	22	12	0

section. The number of segregating loci in each generation is given in Figures 1 to 6 in order to indicate the rate of approach to complete fixation. In this section we shall discuss the effects and interaction of selection, population size and linkage from the point of view of chance of fixation of an allele. It was not possible to follow the simulated populations until all loci were fixed as the approach to complete fixation was very slow especially with parental populations of more than 8 and with free recombinations. However, the proportion of fixed loci after 30 generations of selection and random mating provides some information in this regard.

In the case of complete dominance the trend of response indicates also the fixation of plus genes in successive generations. Analyses of the number of plus genes fixed, give results similar to those presented under quantitative response. The number of inferior genes fixed in each generation, under the specified conditions, indicates the process of fixation more clearly. Analyses of variance, given for 8th and 30th generations in Table 1, show that population size and linkage effects are cumulative over the generations and that differences due to various levels of truncation are small in early generations but become considerably larger in later generations. Figures given in Table 2 indicate that increased selection intensity is important in reducing the proportion of inferior genes fixed when parental population is 16 or less and when genes are closely linked. A strong interaction between population size and linkage as indicated by the response curves is also seen with respect to the fixation of inferior genes. Table 2 illustrates that population size and linkage effects are definitely non-additive.

The results for the overdominance model, presented in Table 3 and 4, indicate population size and linkage effects in spite of selection against fixation of genes. It was mentioned in the previous section that with close linkage the rate of decrease in heterozygosity appeared to be increased with increasing intensity of selection. It is interesting to note in Table 4 that the proportion of fixed loci within 30 generations is also increased under these conditions. ROBERTSON (1962) has shown that when the equilibrium gene frequency is not within a range of .2 to .8, selection for heterozygosity may accelerate fixation in finite populations. The genetic nature of a similar situation with linkage needs further study.

Acknowledgement

The help and criticism extended by Professors Oscar KEMPTHORNE and L. N. HAZEL in designing and conducting the study is gratefully acknowledged.

Zusammenfassung

Die Wirkung und Wechselwirkung der elterlichen Populationsgröße, der engen Koppelung zwischen Loci, dem Grad der Verteilungsstutzung bei Selektion und der Umweltvarianz wurden in einer direkten Monte-Carlo-Simulation für den Fall vollständiger Dominanz und Superdominanz für eine Gleichgewicht-Genfrequenz von 0,5 untersucht. Die Ausgangspopulation besaß eine Genfrequenz von 0,5 an allen 40 Loci, sie befand sich im Zustand des Gametengleichgewichts (gleiche Häufigkeit von Repulsions- und Attraktions-Typen). Die Reaktion der Populationen und die Fixierung der Gene wurden über dreißig Generationen hinweg verfolgt.

Im Falle vollständiger Dominanz erwünschter Gene kann die Langzeitreaktion in eine Initialphase mit schnellen Änderungen und eine späte Phase mit sehr langsamer Näherung an einen Grenzwert unterteilt werden, wobei im Falle enger Koppelung zwischen den Loci Plateaus gebildet werden.

Die Änderungen während der Initialphase machen den größten Teil der Gesamtreaktion aus. Besteht die elterliche Population aus sechzehn oder weniger Individuen, so ist eine zunehmende Selektionsintensität unwirksam, wenn die Rekombinationswerte zwischen benachbarten Loci 0,005 oder niedriger sind. Hinsichtlich der Reduktion der Reaktionsgeschwindigkeit scheinen Koppelungseffekte geometrisch proportional der begrenzten Populationsgröße zuzunehmen.

Die Ergebnisse des Superdominanzmodells erlauben eine interessante Untersuchung der Wirkung der Populationsgröße und der Koppelung auf die Heterozygotie bei selektiver Begünstigung der Heterozygoten. Die Wirkungen sind teilweise erklärbar durch Unterschiede der Heterozygotie zwischen den Chromosomen und innerhalb der Chromosomen. Die Heterozygotiekomponente „zwischen den Chromosomen" wird mit begrenzter Populationsgröße stark reduziert, während die allgemeine Heterozygotie innerhalb der Chromosomen infolge der Koppelung zwischen den Loci erhalten bleibt. Im allgemeinen gibt es strenge Interaktion zwischen den Wirkungen des Selektionsdruckes, der Koppelung und der Populationsgröße. Bei enger Koppelung in kleinen Populationen vermindert ein erhöhter Selektionsdruck das Ausmaß der Genfixierungen. In der größten Population dagegen führt ein erhöhter Selektionsdruck zu einer Zunahme der Fixierungsrate.

Es ist jedoch kar, daß Folgerungen dieser Art noch vorsichtig zu bewerten sind und ein genaues Bild der tatsächlichen Vorgänge erst mit sehr viel ausführlicheren Simulationen erhalten werden kann. Auf jeden Fall wird klar, daß eine Theorie über begrenzte Populationen bei Betrachtung einzelner Loci im Grunde genommen unwirksam ist.

References

1. COMSTOCK, R. C.: Linkage in relation to evidence concerning level of dominance. Proc. X National Poultry Breeders Roundtable. U.S.A. Mimeo. 94–125 (1961). — 2. HAYMAN, B. I., and K. MATHER: The progress of inbreeding when homozygotes are at a disadvantage. Heredity **7**, 165–184 (1953). — 3. KOJIMA, K.: Effects of dominance and size of population on responses to mass selection. Genetical Research **2**, 177–188 (1961). — 4. QURESHI, A. W.: A monte carlo evaluation of the role of finite population size and linkage in response to continuous mass selection. II. Dominance and overdominance. Tech. Report MC 9. Stat. Lab. Iowa State University, Ames, Iowa (1964). — 5. QURESHI, A. W., and O. KEMPTHORNE: Effects of population size, selection intensity, and linkage on the genotypic variance within and between lines under continued truncation selection. (Submitted). — 6. QURESHI, A. W., O. KEMPTHORNE, and L. N. HAZEL: The role of finite population size and linkage in response to continued truncation selection. I. Additive gene action. Theor. and Appl. Gen. **38**, 256–263 1968. — 7. REEVE, E. C. R.: Inbreeding with homozygotes at a disadvantage. Ann. Hum. Genet., Lond., **19**, 332–346 (1955). — 8. ROBERTSON, A.: Selection for heterozygotes in small populations. Genetics **47**, 1291 to 1300 (1962). — 9. WRIGHT, S.: Inbreeding and recombination. Proc. Nat. Acad. Sci. **19**, 420–433 (1933).