

University of Montana

ScholarWorks at University of Montana

Biological Sciences Faculty Publications

Biological Sciences

1983

Linkage Disequilibrium Generated by Selection Against Null Alleles at Duplicate Loci

Fred W. Allendorf

University of Montana - Missoula, Fred.Allendorf@umontana.edu

Follow this and additional works at: https://scholarworks.umt.edu/biosci_pubs



Part of the [Biology Commons](#)

Let us know how access to this document benefits you.

Recommended Citation

Allendorf, Fred W., "Linkage Disequilibrium Generated by Selection Against Null Alleles at Duplicate Loci" (1983). *Biological Sciences Faculty Publications*. 299.
https://scholarworks.umt.edu/biosci_pubs/299

This Response or Comment is brought to you for free and open access by the Biological Sciences at ScholarWorks at University of Montana. It has been accepted for inclusion in Biological Sciences Faculty Publications by an authorized administrator of ScholarWorks at University of Montana. For more information, please contact scholarworks@mso.umt.edu.

NOTES AND COMMENTS

LINKAGE DISEQUILIBRIUM GENERATED BY SELECTION AGAINST
NULL ALLELES AT DUPLICATE LOCI

Baker and co-workers (Roberts and Baker 1973; Baker 1975; Baker and Kaeding 1981) have described a highly interesting case of linkage disequilibrium among active and null alleles at four alpha-esterase loci in *Drosophila montana*. The four tightly linked loci (less than 1% crossing over) apparently arose from two duplications (Roberts and Baker 1973). The ancestral locus was duplicated and the two resulting loci diverged in function. A second tandem duplication event produced the present four loci linked in the order 1-3-2-4 with pairs 1-2 and 3-4 being functionally equivalent.

Baker et al. have found strong linkage disequilibrium between loci 1 and 2 and between 3 and 4. Considering each of these pairs of loci by themselves, there is a far greater number of chromosomes having only one active allele than having either two active or two null alleles. This pattern of disequilibrium is apparently stable in both time and space. Recently Baker has argued persuasively that these relationships are maintained by natural selection (Baker and Kaeding 1981).

These results have been interpreted to show that "there is an optimum number of active loci per chromosome" (Baker 1975, p. 4,098) so that the products of both functional pairs of loci are beneficial but "too much enzyme is detrimental" (Roberts and Baker 1973, p. 722). Such a finding would, as suggested, be "of some theoretical importance relative to the neutrality hypothesis for electrophoretic variants" (Baker 1975, p. 4,099). Such "fine-tuning" of enzyme levels by natural selection would support the view that natural selection is of primary importance in determining the amount and pattern of allozymic variation in natural populations.

The observed linkage disequilibria are best explained by the action of natural selection. However, these results can also be explained by a different pattern of natural selection in which only individuals that are homozygous for null alleles at both loci within a functional pair are selected against.

Consider individually each functional pair of loci showing strong linkage disequilibrium. There are two alternative models of natural selection for active and null alleles at a duplicated locus to consider. In the model suggested by Baker one assumes the highest fitness for individuals having an intermediate amount of enzyme activity. I have previously considered the evolutionary dynamics for this model of selection (Allendorf 1979). In the second model individuals lacking enzyme activity are selected against and all other phenotypes have equal fitness. This model has been considered by many authors (Li 1980 and references therein).

TABLE 1
SIMULATION RESULTS OF THE AMOUNT OF LINKAGE DISEQUILIBRIUM GENERATED

Repeat	$s = .00$			$s = .02$			$s = .10$		
	t^*	D		t	D		t	D	
		Min.†	Final‡		Min.	Final		Min.	Final
1	2,565	-.180	. . .	359	-.125	-.125	155	-.025	. . .
2	1,508	-.082	. . .	461	-.128	-.072	114	-.143	-.118
3	1,060	-.024	. . .	722	-.070	. . .	131	-.126	-.045
4	3,446	-.112	. . .	761	-.030	. . .	237	-.007	. . .
5	2,275	-.180	-.130	520	-.075	. . .	175	-.210	-.209
6	347	-.115	-.047	612	-.180	-.180	131	-.047	-.015
7	1,434	-.097	. . .	459	-.074	. . .	160	-.022	. . .
8	1,507	-.039	. . .	353	-.164	-.127	129	-.003	. . .
9	3,768	-.071	. . .	763	-.029	. . .	150	-.003	. . .
10	2,960	-.228	-.228	521	-.089	. . .	143	-.065	-.052
11	1,345	-.145	-.050	722	-.083	. . .	136	-.016	. . .
12	1,498	-.013	. . .	857	-.039	. . .	149	-.015	. . .
13	1,304	-.161	. . .	525	-.026	. . .	171	-.115	-.103
14	1,884	-.073	. . .	312	-.011	. . .	209	-.053	-.030
15	964	-.090	. . .	964	-.038	. . .	143	-.057	-.012
16	887	-.087	. . .	435	-.078	-.047	199	-.203	-.203
17	628	-.153	. . .	706	-.029	. . .	126	-.199	-.160
18	1,699	-.032	. . .	321	-.019	. . .	187	-.216	-.179
19	1,425	-.128	. . .	704	-.062	. . .	281	-.169	-.117
20	683	-.011	. . .	974	-.183	. . .	269	-.226	-.226
21	4,043	-.041	. . .	961	-.106	. . .	207	-.130	-.091
22	2,514	-.240	. . .	680	-.039	. . .	145	-.212	-.208
23	1,113	-.112	-.022	361	-.053	-.024	270	-.026	. . .
24	1,692	-.075	. . .	858	-.015	. . .	233	-.219	-.219
25	936	-.075	. . .	650	-.013	. . .	139	-.003	. . .
Median	1,498	-.090		650	-.062		155	-.065	
Mean	1,725	-.103		662	-.070		176	-.100	

* Number of generations until loss of duplicate gene expression (see text).

† Minimum value of linkage disequilibrium (D) generated.

‡ Columns lacking values represent repeats in which only one locus was polymorphic at the end of the simulation.

Consider a random mating population with two loci, each with two alleles, an active allele (A,B), and a null allele (a,b). Assume a recombination rate between loci of $r = 0.005$ and unidirectional mutation from the active to the null allele with frequency $u = 0.0005$. In a diploid population, there are five phenotypic classes with the following fitnesses:

	BB	Bb	bb
AA	$1 - s$	$1 - \frac{s}{2}$	1
Aa	$1 - \frac{s}{2}$	1	1
aa	1	1	0.

Thus, s is a measure of the selection against phenotypes having too much enzyme activity. In the second model of selection, all phenotypes having at least one active allele are equally fit and therefore $s = 0$.

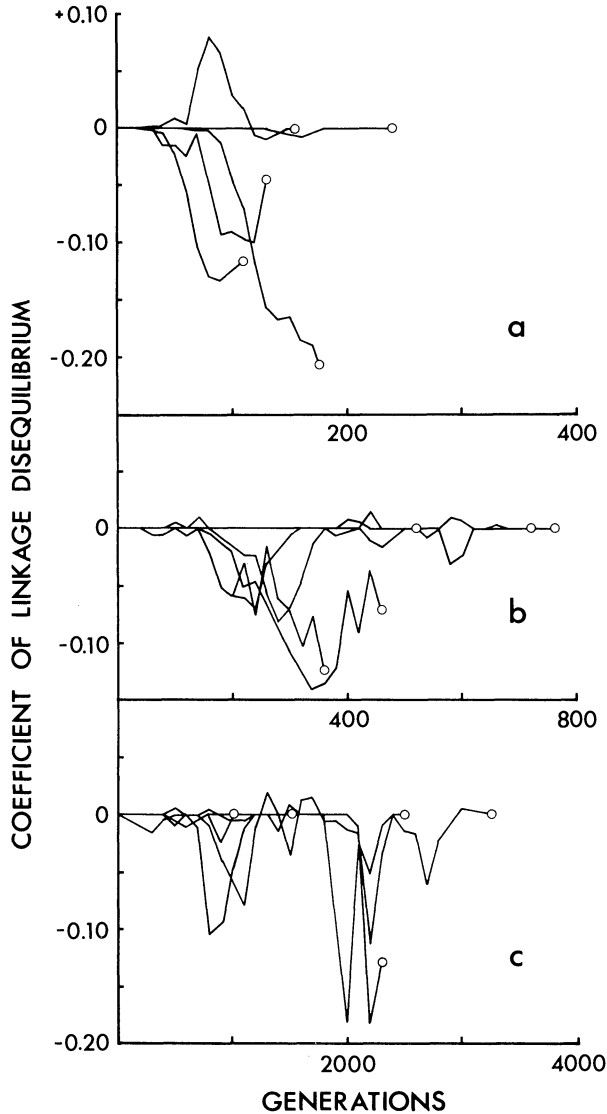


FIG. 1.—Amount of linkage disequilibrium (D) generated in first five simulation repeats with each value of s : a, 0.10; b, 0.02; and c, 0.

I have considered three different values of s : 0.10, 0.02, and 0. When $s > 0$, changes in gametic frequencies between generations can be predicted in the deterministic case using recursion equations (Karlin and McGregor 1971). Changes in gametic frequencies were also modeled for all three values of s using a Monte Carlo simulation with a population size of $N = 200$ (Allendorf 1979).

In table 1 I show the results of 25 simulation repeats for each value of s , with initial frequencies of the active alleles at each locus of 1.0. A simulation ended

when loss of duplicate gene expression occurred, i.e., when the frequency of the phenotypes having more than two active alleles was less than 1%. As expected, loss of duplicate gene expression occurred more rapidly with larger values of s . More importantly, however, all three modes of selection resulted in substantial amounts of linkage disequilibrium (also see fig. 1).

In fact, there was more linkage disequilibrium for the case in which all phenotypes with an active allele have equal fitness than for the case with weak selection ($s = 0.02$) against phenotypes having too much enzyme. This effect is not simply the result of the more rapid loss of duplicate gene expression when $s = 0.02$. There is no significant correlation between the number of generations until loss and the minimum value of D generated when $s = 0$. Rather, for the case in which $s = 0$, allele frequencies at the two loci drift along a ridge of equal fitnesses where the frequency of the phenotype lacking enzyme activity is equal to the mutation rate (Christiansen and Frydenberg 1977). When $s > 0$, selection acts to eliminate "excess" copies of active alleles. The trajectory of allele frequencies with $s > 0$ is closer to the one-locus polymorphism boundary than is the ridge of equal maximum fitness when $s = 0$. Substantial linkage equilibrium can only be present when both loci are polymorphic.

The specific dynamics of these models will be affected by the values of N , r , u , and s . However, the equilibrium for both models of selection is the eventual loss of duplicate gene expression. It is, therefore, difficult to explain the similar gametic frequencies in isolated populations with either model. Probably the strongest argument for $s = 0$ is the apparent persistence of the duplication. Even with weak selection against having too much enzyme activity, duplicated gene expression is expected to be lost rather rapidly (table 1). The important point, however, is that selection for an intermediate amount of enzyme activity is not necessary to produce the linkage disequilibrium detected in the alpha-esterase system of *D. montana*.

ACKNOWLEDGMENTS

I thank W. Baker, T. Prout, and an anonymous reviewer for their comments on this note, K. Knudsen for preparing the figure, and D. Parkin and the Cripps Computer Centre of the University of Nottingham for providing the facilities and time for the simulations presented here that were completed while I was supported by a NATO Fellowship. During preparation of this manuscript, I was supported by National Science Foundation grants DEB-8004681 and ISP-8011449.

LITERATURE CITED

- Allendorf, F. W. 1979. Rapid loss of duplicate gene expression by natural selection. *Heredity* 43:247-258.
- Baker, W. K. 1975. Linkage disequilibrium over space and time in natural populations of *Drosophila montana*. *Proc. Natl. Acad. Sci. USA* 72:4095-4099.
- Baker, W. K., and E. A. Kaeding. 1981. Linkage disequilibrium at the alpha-esterase loci in a population of *Drosophila montana* from Utah. *Am. Nat.* 117:804-809.

- Christiansen, F. B., and O. Frydenberg. 1977. Selection-mutation balance for two nonallelic recessives producing an inferior double homozygote. *Am. J. Hum. Genet.* 29:195-207.
- Karlin, S., and J. McGregor. 1971. On mutation selection balance for two-locus haploid and diploid populations. *Theor. Popul. Biol.* 2:60-70.
- Li, W.-H. 1980. Rate of gene silencing at duplicate loci: a theoretical study and interpretation of data from tetraploid fishes. *Genetics* 95:237-258.
- Roberts, R. M., and W. K. Baker. 1973. Frequency distribution and linkage disequilibrium of active and null esterase isozymes in natural populations of *Drosophila montana*. *Am. Nat.* 107:709-726.

FRED W. ALLENDORF

DEPARTMENT OF ZOOLOGY
UNIVERSITY OF MONTANA
MISSOULA, MONTANA 59812

Submitted July 30, 1981; Accepted July 9, 1982