INHERITANCE OF BLACK HAIR PATTERNS IN CATTLE LACKING THE EXTENSION FACTOR FOR BLACK (E). III. A MULTIPLE ALLELIC HYPOTHESIS TO EXPLAIN THE INHERITANCE OF BLACKISH AND BLACKISH PATTERN[1]

LOUISE F. OVERTON,[2] R. LAIRD[3] AND N. S. FECHHEIMER

Department of Dairy Science and the Institute of Genetics, The Ohio State University, Columbus
and
Department of Dairy Science, The Ohio Agricultural Experiment Station, Wooster

INTRODUCTION

Baldwin *et al.* (1954, 1956) have defined the terms "blackish" and "blackish pattern" and have concluded that the gene for blackish is dominant to its allele, the nonblackish gene. With reference to blackish pattern, Baldwin *et al.* (1956) concluded that its expressivity depends on the presence of the blackish gene. Further, they confirmed that blackish pattern is sex-influenced, the gene for blackish pattern acting as a dominant to its allele for blackish (nonpattern) in the male and as a recessive in the female. From their study, Baldwin *et al.* (1954, 1956) suggested that there are two loci involved in the inheritance of blackish and blackish pattern.

In the present paper an alternative hypothesis involving only one locus will be tested. This locus has three alleles, *viz.*, the genes for blackish pattern, blackish, and nonblackish. The nonblackish gene is assumed to be recessive to the gene for blackish and also to the gene for blackish pattern. The blackish pattern gene is again assumed to be sex-influenced, being dominant to the gene for blackish in the male and recessive to it in the female. The symbols adopted for the genes are as follows: Bs_M, blackish pattern; Bs, blackish; and bs, nonblackish. The terminology is similar to that suggested by Baldwin *et al.* (1954, 1956). The lower case letter *b* in *bs* is used because this gene is completely recessive to the other two alleles; the subscript *M* in Bs_M is used because this gene is dominant to Bs in the male and recessive in the female. In the female it is assumed that the homozygous state of the blackish pattern gene is expressed phenotypically as blackish pattern while the heterozygous condition is expressed as blackish. The phenotypes of the various genotypes are shown in table 1.

EXPERIMENTAL PROCEDURE AND RESULTS

In order to test this hypothesis on a given sample, it is necessary to develop a method of statistical analysis. If a given population is panmictic with respect to the colors involved and if there are no gene interactions with other loci, we may use the method of maximum likelihood to obtain estimates of gene frequency from

a random sample and to test the hypothesis outlined above.   General descriptions of the maximum likelihood method are presented in Mather (1951) and Hogben (1946).

<div align="center">TABLE 1</div>

<div align="center">*The phenotypes of the genotypes of the blackish allelic series*</div>

| Genotype | Phenotype | |
	Male	Female
$Bs_M Bs_M$	blackish pattern	blackish pattern
$Bs_M Bs$	blackish pattern	blackish
$Bs_M bs$	blackish pattern	blackish
$Bs Bs$	blackish	blackish
$Bs bs$	blackish	blackish
$bs bs$	nonblackish	nonblackish

Development of Formulae

In order to apply the maximum likelihood method of estimation to gene frequencies of the present problem, let,

a_1 = the observed number of blackish pattern males,
a_2 = the observed number of blackish males,
a_3 = the observed number of nonblackish males,
a_4 = the observed number of blackish pattern females,
a_5 = the observed number of blackish females,
a_6 = the observed number of nonblackish females,
i = 1, 2, 3, 4, 5, 6 classes,
$n = a_1+a_2+a_3+a_4+a_5+a_6$, the total number of individuals,
$m = (a_1+a_2+a_3)/n$, the proportion of males in the sample,
$1-m = (a_4+a_5+a_6)/n$, the proportion of females in the sample,
p = the true frequency of Bs_M,
q = the true frequency of Bs,
r = the true frequency of bs,
        (Note that $p+q+r=1$.)
$\hat{p}$ = the maximum likelihood estimate of p,
$\hat{q}$ = the maximum likelihood estimate of q,
$\hat{r}$ = the maximum likelihood estimate of r,
f_i = the probability associated with the ith class,
nf_i = the expected number of individuals in the ith class.

Table 2 shows symbolically the observed numbers and probabilities for the six classes.

<div align="center">TABLE 2</div>

<div align="center">*Algebraic symbols for observed numbers and probabilities*</div>

	Class i	Observed number a_i	Probability f_i
1	(blackish pattern males)	a_1	$m\ (p^2+2pq+2pr)$
2	(blackish males)	a_2	$m\ (q^2+2qr)$
3	(nonblackish males)	a_3	$m\ (r^2)$
4	(blackish pattern females)	a_4	$(1-m)\ (p^2)$
5	(blackish females)	a_5	$(1-m)\ (2pq+2pr+q^2+2qr)$
6	(nonblackish females)	a_6	$(1-m)\ (r^2)$

Then, the exact probability π of obtaining the observed sample is,

$$\pi = \frac{n!}{a_1!a_2!a_3!a_4!a_5!a_6!}(p^2+2pq+2pr)^{a_1}(q^2+2qr)^{a_2}(r^2)^{a_3}(p^2)^{a_4}$$
$$(2pq+2pr+q^2+2qr)^{a_5}(r^2)^{a_6}m^{(a_1+a_2+a_3)}$$
$$(1-m)^{(a_4+a_5+a_6)}.$$

A simplified logarithmic likelihood expression L may be derived from π by factoring, by using the relationship $(p+q+r)^2=1$, and by omitting constants:

$$L = a_1 \log p + a_1 \log (p+2q+2r) + a_2 \log q + a_2 \log (q+2r)$$
$$+ 2a_3 \log r + 2a_4 \log p + a_5 \log (1-p^2-r^2) + 2a_6 \log r.$$

The maximum likelihood method requires that L be maximized, by taking the partial derivatives of L with respect to p, q, and r, setting the resulting three equations equal to zero, and solving them simultaneously for $\hat{p}$, $\hat{q}$, and $\hat{r}$. In the present problem the variables are mutually dependent under the restriction

$$p+q+r=1,$$

such that the partial derivatives of L with respect to each of the three variables become complicated. The method of Lagrange multipliers provides a tool for solving such a system. A discussion of this method may be found in Widder (1947), page 113.

By the method of Lagrange, we introduce the Lagrange multiplier λ and form a new function V, where,

$$V = L + \lambda(p+q+r-1).$$

Then we may treat p, q, and r as if they were independent variables and we set

$$\frac{\partial V}{\partial p} = \frac{\partial L}{\partial p} + \lambda = 0$$

$$\frac{\partial V}{\partial q} = \frac{\partial L}{\partial q} + \lambda = 0$$

$$\frac{\partial V}{\partial r} = \frac{\partial L}{\partial r} + \lambda = 0.$$

We may solve this system of equations simultaneously with

$$p+q+r-1=0.$$

We will thus obtain values of λ, p, q, and r. The estimates of p, q, and r will be the maximum likelihood estimates $\hat{p}$, $\hat{q}$, and $\hat{r}$, which we seek. Since we are interested in λ only as a tool, we may discard it.

Following the plan outlined above, we obtain the equations,

$$\frac{\partial V}{\partial p} = \frac{a_1}{p} + \frac{a_1}{p+2q+2r} + \frac{2a_4}{p} - \frac{2a_5 p}{1-p^2-r^2} + \lambda = 0 \qquad (1)$$

$$\frac{\partial V}{\partial q} = \frac{2a_1}{p+2q+2r} + \frac{a_2}{q} + \frac{a_2}{q+2r} + \lambda = 0 \qquad (2)$$

$$\frac{\partial V}{\partial r} = \frac{2a_1}{p+2q+2r} + \frac{2a_2}{q+2r} + \frac{2a_3}{r} - \frac{2ra_5}{1-p^2-r^2} + \frac{2a_6}{r} + \lambda = 0 \qquad (3)$$

$$p+q+r=1. \qquad (4)$$

Eliminating q by substituting $q=1-p-r$, we obtain

$$\frac{a_1}{2-p} + \frac{a_1+2a_4}{p} - \frac{2a_5 p}{1-p^2-r^2} + \lambda = 0 \qquad (5)$$

$$\frac{2a_1}{2-p} + \frac{a_2}{1-p-r} + \frac{a_2}{1-p+r} + \lambda = 0 \qquad (6)$$

$$\frac{2a_1}{2-p}+\frac{2a_3+2a_6}{r}-\frac{2a_5r}{1-p^2-r^2}+\frac{2a_2}{1-p+r}+\lambda=0 \tag{7}$$

(7) minus (6) produces the equation,

$$\frac{2a_3+2a_6}{r}-\frac{2a_5r}{1-p^2-r^2}-\frac{a_2}{1-p-r}+\frac{a_2}{1-p+r}=0 \tag{8}$$

and (5) minus (6) gives,

$$\frac{a_1+2a_4}{p}-\frac{a_1}{2-p}-\frac{2a_5p}{1-p^2-r^2}-\frac{a_2}{1-p-r}-\frac{a_2}{1-p+r}=0. \tag{9}$$

By eliminating fractions in (8) and (9), we may obtain equations of estimation of p and r:

$$(a_2+a_3+a_5+a_6)r^4-(a_2+2a_3+a_5+2a_6)r^2+(a_2-a_5)r^2p^2+2(a_3+a_5+a_6)r^2p$$
$$-(a_3+a_6)p^4+2(a_3+a_6)p^3-2(a_3+a_6)p+(a_3+a_6)=0 \tag{10}$$

and,

$$(a_1+2a_4)+(a_1+2a_4)r^4-2(a_1+2a_4)r^2-(a_1+a_4)r^4p+2(2a_1+a_2+3a_4)r^2p$$
$$-(2a_1+3a_2+2a_4-2a_5)r^2p^2+(a_2-a_5)r^2p^3-(3a_1+2a_2+5a_4)p$$
$$+(2a_1+3a_2+2a_4-2a_5)p^2+(2a_1+a_2+4a_4+5a_5)p^3-(3a_1+3a_2+4a_4+4a_5)p^4$$
$$+(a_1+a_2+a_4+a_5)p^5=0. \tag{11}$$

Since there exists no easy algebraic approach to the simultaneous solution of these two equations, they may be solved for any specific example by iteration. A discussion of this iterative method will appear in the examples of the subsequent section.

Test of Hypothesis

The data used to test the hypothesis in this study are those of Baldwin *et al.* (1954, 1956). Only males of 1 year of age and above and females of 2 years of age and above were used. Animals carrying much white (*lw lw*) were omitted because the presence or absence of the blackish pattern on such animals could not be determined. The animals used in this study were 229 Jerseys (61 males and 168 females) and 372 Ayrshires (66 males and 306 females). The numbers in each of the 6 phenotypic classes are indicated in table 3.

TABLE 3

Observed phenotypic frequencies of Jerseys and Ayrshires

Phenotype	Jerseys	Ayrshires
Blackish pattern males	46	36
Blackish males	14	23
Nonblackish males	1	7
Blackish pattern females	50	51
Blackish females	112	200
Nonblackish females	6	55

Jersey data.—By formulae (10) and (11) the two equations of estimation of p and r for the Jersey case are,

$$f(p,r)=133\ r^4-140\ r^2-98\ r^2p^2+238\ r^2p-7p^4+14p^3-14p+7=0 \tag{12}$$

and,

$$g(p,r)=146+146\ r^4-292\ r^2-96\ r^4p+512\ r^2p-10\ r^2p^2-98\ r^2p^3-416\ p$$
$$+10\ p^2+866\ p^3-828\ p^4+222\ p^5=0. \tag{13}$$

Since there is no easy algebraic method of solution, these equations must be solved simultaneously by iteration. This method involves successive approximations of p and r, until the values of $f(p,r)$ and $g(p,r)$ become as close to zero as the desired degree of accuracy. Three decimal accuracy was chosen in the present problem.

In the Jersey case the value of r was narrowed to the range .171 to .173; and p, to the range .531 to .533. Table 4 represents the values of $f(p,r)$ and $g(p,r)$ under the various combinations of p and r in the above ranges.

The best simultaneous solution of the two equations is $p = .532$ and $r = .172$; therefore, these values are designated as the maximum likelihood estimates $\hat{p}$ and $\hat{r}$. Through the relationship $\hat{q} = 1 - \hat{p} - \hat{r}$, we obtain $\hat{q} = .296$. It will be noted that $\hat{r}$ does not differ appreciably from the value of .19 found by the method of Baldwin et al. (1954).

From $\hat{p}$, $\hat{q}$, and $\hat{r}$ theoretically expected phenotypic frequencies may be computed by means of the f_i defined in table 2. Table 5 presents observed and calculated numbers for the six phenotypes.

TABLE 4

Values of f(p,r) *and* g(p,r) *corresponding to approximations for* p *and* r

p	r	$f(p,r)$	$g(p,r)$
.531	.171	.013	.107
.532	.171	.011	.039
.533	.171	.008	− .029
.531	.172	.001	.096
.532	.172	− .001	.029
.533	.172	− .003	− .039
.531	.173	− .010	.085
.532	.173	− .012	.018
.533	.173	− .015	− .050

TABLE 5

Quantities from Jersey data necessary for calculation of χ^2

Class i	Observed number a_i	Calculated number nf_i
1 (blackish pattern males)	46	47.572
2 (blackish males)	14	11.540
3 (nonblackish males)	1	1.802
4 (blackish pattern females)	50	47.572
5 (blackish females)	112	115.541
6 (nonblackish females)	6	4.973

The calculated frequencies were compared with the observed frequencies by the chi-square test. In the Jersey case the calculated numbers in classes 3 and 6 are so small that the two classes were combined for purposes of χ^2 calculation; therefore, we must consider that we have only five classes. Since two of the gene frequencies were calculated independently (the third by difference from 1), the degrees of freedom appropriate to the chi-square test are the number of classes minus the number of independent parameters estimated from the data minus

1, *i.e.*, $5-2-1=2$. The chi-square value for the Jerseys is .816 for 2 degrees of freedom with a probability of about 96%, indicating that the suggested hypothesis fits the data very well.

Ayrshire data.—The Ayrshire data in table 3 lead to equations of estimation (14) and (15):

$$f(p,r) = 285\ r^4 - 347\ r^2 - 177\ r^2p^2 + 524\ r^2p - 62\ p^4 + 124\ p^3 - 124\ p + 62 = 0 \quad (14)$$

and,

$$g(p,r) = 138 + 138\ r^4 - 276\ r^2 - 87\ r^4p + 496\ r^2p + 157\ r^2p^2 - 177\ r^2p^3 - 409\ p$$
$$- 157\ p^2 + 1299\ p^3 - 1181\ p^4 + 310\ p^5 = 0. \quad (15)$$

Solution of these equations gives $\hat{p}=.378$ and $\hat{r}=.400$; therefore, $\hat{q}=.222$. It will be noted that the value for $\hat{r}$ is similar to the corresponding value of .42 found by Baldwin *et al.* (1954). Observed and calculated frequencies necessary for calculation of χ^2 are presented in table 6.

TABLE 6

Quantities necessary for calculation of χ^2
for Ayrshire case

Class i	Observed numbers a_i	Calculated numbers nf_i
1	36	40.370
2	23	14.939
3	7	10.535
4	51	43.745
5	200	213.426
6	55	48.985

In this case $\chi^2=8.795$ with 3 degrees of freedom. The suggested hypothesis is therefore rejected at the 5 percent level of significance but accepted at the 2 percent level.

DISCUSSION

The above analysis has shown that the proposed hypothesis of gene expression is accepted in the Jersey case but is rejected at the 5 percent level in the Ayrshire case. If the genes are the same in the two breeds, an assumption which seems reasonable, the results are inconclusive regarding the validity of the hypothesis. This may be due to the fact that either or both of the sexes in either or both of the breeds are not random samples of the population in question. Further samples might help to elucidate this point. There is also the possibility of non-random mating, but as far as is known, there is no discrimination for or against blackish pattern or blackish.

The results of matings could also distinguish between the present hypothesis and that of Baldwin *et al.* (1954, 1956). If a mating between a blackish bull and a nonblackish cow produced a male blackish pattern offspring, then the multiple allelic hypothesis would be rejected; since, under this hypothesis, the bull would be either *BsBs* or *Bsbs* and the cow *bsbs*, no blackish pattern offspring could result from this mating. Under the hypothesis of Baldwin and his associates the cow could be *bsbs Bp_M*−; when mated with a blackish bull *Bs*− *BpBp* she could produce a blackish pattern male (*BsBs Bp_MBp*). Another distinguishing mating could be a blackish pattern cow with a nonblackish bull. By the multiple allelic hypothesis only blackish pattern male offspring and blackish female offspring

can result. On the other hand, under the hypothesis of Baldwin *et al.* (1954) males may be blackish pattern or nonblackish, while females may be of all three phenotypes. Such matings might be undertaken to test the multiple allelic hypothesis.

Progeny resulting from matings of the two kinds suggested have not been observed for two reasons: 1) The data used for the calculations in this paper were collected by Baldwin *et al.* (1954, 1956) for a different purpose and did not include information on mating structure; 2) these matings occur relatively infrequently and therefore resulting progeny are quite scarce.

SUMMARY AND CONCLUSIONS

A multiple allelic hypothesis is suggested to explain the inheritance of blackish and blackish pattern in Jersey and Ayrshire cattle. The three alleles are blackish pattern (Bs_M), blackish (Bs) and nonblackish (bs). Blackish pattern is sex-influenced.

Sample data were used to test this hypothesis. The hypothesis was accepted—with a high degree of probability—in the Jersey data but was rejected at the 5 percent level of significance in the Ayrshire data. The results for the two breeds are conflicting and are inconclusive regarding the validity of the hypothesis. Matings are suggested for further testing this hypothesis.

LITERATURE CITED

Baldwin, C. S., L. O. Gilmore, and N. S. Fechheimer. 1954. Inheritance of black hair patterns in cattle lacking the extension factor for black (E). I. Blackish. Jour. Dairy Sci. 37: 1368–1375.

——, ——, ——, and D. C. Rife. 1956. Inheritance of black hair patterns in cattle lacking the extension factor for black (E). II. Blackish pattern. Jour. Dairy Sci. 39: 435–439.

Hogben, L. 1946. An Introduction to Mathematical Genetics. W. W. Norton and Co., New York. 260 pp.

Mather, K. 1951. The Measurement of Linkage in Heredity, 2nd edition. John Wiley and Sons, New York. 149 pp.

Widder, D. V. 1947. Advanced Calculus. Prentice-Hall, Inc., New York. 432 pp.