

17th. April 1946.

My dear Henry,

In talking about your Dominula data to a Part II class last term it occurred to me that although the seven years recorded show clearly significant variation in the gene ratio, yet that it might be considered, contrary to my own opinion, that this variation was due to what Wright and others have called drift or random survival, and that if a means could be found of testing homogeneity subject to such a disturbance, your data provided a priceless opportunity and probably the first sound test available of seeing whether a small isolated population did show variations in gene ratio of the order of magnitude supposed by Wright, or whether, on the contrary, owing to annual fluctuations in the intensity and direction of selection, the random survival component was, in fact, of no consequence.

The result of these reflections was the enclosed sheet, where I begin with a straight  $\chi^2$  test of homogeneity of the gene ratio in the seven years, and then make a more refined test using what, I have called the angular transformation, which allows for the fact that a low estimate like 5% has higher precision in the percentage scale than would an estimate of 10% based on the same number of individuals, whereas on the angular scale, for which I have used ordinary degrees of angle, the precision depends only on the size of the sample. Since the widest deviations were those of the first two years in the positive direction,  $\chi^2$  is reduced by using the angular transformation

that is still large, about 19.9 for six degrees of freedom.

The question now is, could so large a measure of heterogeneity have been produced by variations in the random survival? Here, of course, a good deal depends on the supposed size of the breeding population, for with an infinitely large population there would be no variation due to random survival, and the smaller the population is made the larger are the effects which could easily be ascribed to it.

Now, owing to your marking and recapture scheme, we have in this case good data as to the numbers flying for most of the years, and you will probably agree with me that I have fairly grossly overestimated the random survival effect, or underestimated the breeding population by taking it to be only 1,000 insects in each year. The real effect will probably be only one third, or some small fraction of that which I have allowed in my calculations.

On the basis of 1,000 insects, which is only just more than double the numbers enumerated in 1941 and 1944, and taking the insects enumerated to belong to the breeding population in each year, which is true, I think, except for some larvae reared in captivity in 1941, I find that  $\chi^2$  is reduced only to about 14.8, which is still very significant for six degrees of freedom. In fact, with the low population number assumed, less than half of the full discrepancy is accounted for, and it would appear that in the Cothill population selective forces must have five or six times as much effect as can be ascribed to random survival.

In view of the immensely exaggerated stress which Wright and others have been tempted to lay on this factor, I believe this aspect of your *Dominula* work at Cothill will deserve some emphasis.

Remember you are due to make me a visit fairly soon. I am hoping to get Mather to stay with me for the nights of 8th. and 9th. of May. You might care to come at the same time, or at any other which would suit you better.

Yours sincerely,

Seven years records of medionigra variant gene in Cothill population  
of Collinornha dominula.

Year	Variant genes	Total n	p%	$\theta^\circ$	$\pi\theta$	
1939	41	446	9.192825	17.6408	7471.8198	Angular transformation  $\sin^2 \theta = p.$  $\left\{ \sum N_i \theta_i^2 - \frac{\sum^2 N_i \theta_i}{\sum N_i} \right\}$ $\div 820.7016$
1940	26	234	11.111111	19.4711	1556.2374	
1941	63	922	6.832972	15.1533	13971.3126	
1942	22	410	5.365854	13.3937	5491.4170	
1943	30	538	5.576208	13.6588	7348.4344	
1944	45	992	4.536290	12.2974	12199.0208	
1945	48	744	6.451613	14.7142	10947.3648	
	—	—			—	
	275	4286			62385.6278	
Mean			6.416239	14.55568		1946 84 1972
		$\chi^2_6$	21.8087	19.8814		

For this angular transformation  $\chi^2$  is somewhat reduced, since the largest deviations are in the positive direction, and these have the lower actual precision in measure per cent.

Conceivably the variation from year to year might be due to random survival rather than annual selection. If the breeding population were constant, equal to  $\frac{1}{2}N$ , and samples of  $\frac{1}{2}N_1, \dots, \frac{1}{2}N_7$  were taken in successive years, the covariance matrix in angular measure, with the constant factor  $820.7016 = 8100/\pi^2$ , would be

$1/N_1 + 2/N$	$2/N$	$1/N$	0	0	0	0
$2/N$	$1/N_2 + 1/N$	$1/N$	0	0	0	0
$1/N$	$1/N$	$1/N_3$	0	0	0	0

0	0	0	$1/N_4 - 1/N$	0	0	0
0	0	0	0	$1/N_5$	$1/N$	$1/N$
0	0	0	0	$1/N$	$1/N_6 + 1/N$	$2/N$
0	0	0	0	$1/N$	$2/N$	$1/N_7 + 2/N$

Taking  $N = 2000$ , an <sup>a</sup> ~~extravagantly~~ low value, intended to over-estimate the effect of random survival, this, multiplied by 10,000 gives

32.4215	10.0000	5.0000	0	0	0	0
10.0000	47.7350	5.0000	0	0	0	0
5.0000	5.0000	10.8460	0	0	0	0
0	0	0	19.3902	0	0	0
0	0	0	0	18.5874	5.0000	5.0000
0	0	0	0	5.0000	15.0806	10.0000
0	0	0	0	5.0000	10.0000	23.4409

Inverting the matrix, and removing the factor 10000, we find

347.0642	- 58.7863	- 132.8960	0	0	0	0	155.3819
- 58.7863	230.0762	- 78.9645	0	0	0	0	92.3254
- 132.8960	- 78.9645	1019.5065	0	0	0	0	807.8060
0	0	0	515.7244	0	0	0	515.7244
0	0	0	0	596.6287	- 158.1684	-59.7369	378.6734
0	0	0	0	- 158.1684	966.6110	-378.6232	429.3194
0	0	0	0	- 59.7369	- 378.6232	600.8799	162.4698
							2542.2003

2967.1668    2245.6975    11568.1989    6907.4579    5324.4556    4155.2741    3368.7688    36537.0196

$$\sum_j A_{ij} = A_i \quad \left\{ \sum T_i \theta_i - \frac{\sum^2 T_i}{\sum A_i} \right\} \times \frac{\pi^2}{8100}$$

14°.37220332

$$\sum_j A_{ij} H_j = T_i$$

$\chi^2_6$  14.8469

$$\sum T_i / \sum A_i = \bar{\theta}$$

25 15.033

In this lower margin are the sums of the entries in each column multiplied by the corresponding value of  $\bar{\theta}$ . The right hand column shows the simple sums of the entries in the rows. The ratio of the two grand totals

$$36537.0196 \div 2542.2003 = 14.3722$$

gives the weighted mean value of  $\theta$ .

To find  $\chi^2$  the column totals are multiplied by the corresponding observed value of  $\theta$ , and added, the corresponding grand total is multiplied by the average  $\bar{\theta}$  and subtracted. The difference dividing by 820.7016 is  $\chi^2$  for six degrees of freedom.

The value obtained

$$\chi^2 = 14.8469$$

is near the 2% point for six degrees of freedom, and clearly significant. Even if the annual number in the breeding population were no more than 1,000 individuals, certainly <sup>as low as</sup> ~~less than~~ the least number in any of these years, random survival would not account for the variations in frequency observed, which must be ascribed to selective action varying in intensity and direction from year to year. It seems extremely probable that this is true with all isolated populations, even though so small as to be not very long enduring. Random survival is probably in no real case an effective factor in determining variability, or the distribution of gene ratios.