

COLOUR PATTERNS AND GENE FREQUENCIES IN NATURAL POPULATIONS OF A PLATYFISH *

HUGH GORDON and MYRON GORDON
Columbia University and New York Zoological Society

Received 20.iv.49

THE platyfish, *Platypoecilus maculatus* Günther, is one of the most highly polymorphic vertebrates in Central and North America, having over 150 recognisable colour patterns in its various natural populations. The black colour patterns are genetically controlled by two groups of genes. First, an autosomal series consisting of seven dominant alleles (Gordon, 1947) determines melanic patterns at or near the base of the tail fin ; these patterns are composed of micromelanophores. Second, a sex-linked series consisting of five dominant alleles (Gordon, 1948) produces another group of melanic patterns on the body proper ; these, however, are composed of macromelanophores, pigment cells which are much larger and more prominent than those which form the tail patterns.

According to a recent genetic and geographical analysis of the inheritance and distribution of the seven tail patterns, Gordon (1947) indicated that the four natural populations of the platyfish from the Rio Jamapa, Rio Papaloapan, Rio Coatzacoalcos and Rio Usumacinta are genetically distinct. In that analysis only single tail patterns were considered in comparing one population with another, but the frequencies of the dual tail pattern combinations were presented in the tabular matter.

The purpose of the present analysis is to determine whether the values observed for the seven single patterns are consistent with each other and with their various combinations. If they are, the observed frequency of a single pattern would correctly indicate the approximate frequency of the gene responsible for that pattern. Any conclusion drawn about the differences of the frequency of certain single patterns between two populations would also be true for the frequency of the genes corresponding to these patterns.

A careful analysis was made of the two largest populations, those of the Rio Jamapa and of Rio Papaloapan. No attempt was made at this time to calculate gene frequencies for the Rio Usumacinta and Rio Coatzacoalcos samples because the samples are small, considering

* From the Genetics Laboratory of the New York Aquarium at the American Museum of Natural History, New York 24, New York. This work was done with the support of a grant from the American Philosophical Society. The authors wish to express their thanks to the Departments of Birds and of Animal Behaviour for the generous use of their facilities at the American Museum of Natural History, and to Dr Sewall Wright, Dr Th. Dobzhansky, Dr Ernst Mayr and Mr James W. Atz for reading and criticising the manuscript.

the large number of possible patterns. Since there are no obvious inconsistencies in the observed numbers, however, it may be assumed that if the Rio Jamapa and Rio Papaloapan observations are consistent, the Rio Usumacinta and Rio Coatzacoalcos observations are also consistent.

METHOD

In the analysis of the Rio Jamapa platyfish population and in the first calculation in the study of the Rio Papaloapan population (see tables 1 and 2), the gene frequencies were found primarily by using the formula derived from Hardy (1908) :—

$$a = 1 - \sqrt{1-x}$$

where a is the gene frequency of the gene responsible for a pattern, and x the portion of the sample which shows that pattern both alone and in combination with another pattern. For each population, certain modifications of this method were needed, because some of the smaller tail patterns are concealed by the larger ones.

TABLE 1

Observed numbers, gene frequencies and calculated numbers based upon these gene frequencies for the various patterns in platyfish from the Rio Jamapa

Patterns	Observed nos.	Calculated nos.	Values for χ^2	Gene frequencies
+	57	59.9	0.14	+ 0.2641
O	277	305.6	2.68	O 0.3880
C	47	46.4	0.01	C 0.0875
Cc, CCc	89	68.5	6.13	Cc 0.0992
Co	44	44.1	0.00	Co 0.0838
T	54	40.3	5.06	T 0.0774
OC	62	58.3	0.23	...
OCc	69	66.2	0.11	...
OCo	72	55.9	4.63	...
OT	58	51.6	0.79	...
CCo	11	12.6	0.21	...
CT	9	11.6	0.59	...
CcCo	4	14.3	7.42	...
CcT	0	13.2	13.20	...
CoT	7	11.2	1.58	...
Total	860	859.7	42.84	1.0000

For example, CCc cannot be phenotypically distinguished from Cc alone, nor MMc from Mc (see fig. 1), so that the frequencies of C and M cannot be found by the formula. The frequency of C was found in each population by subtracting the frequency of Cc from that of C and Cc together. This latter frequency was first found by the formula indicated above.

In the Rio Papaloapan platyfish several patterns involving C , Co , or T , when combined with M are indistinguishable phenotypically from similar combinations with Mc (see fig. 1). The gene frequencies of

M and Mc in the population were found by first considering the fish with the doubtful patterns as neither M nor Mc and then making corrections for these doubtful individuals. The formulæ for these corrections may be derived as follows. Suppose that a portion d of

TABLE 2

Observed numbers, gene frequencies, and calculated numbers based upon these gene frequencies for the various patterns in platyfish from the Rio Papaloapan

Patterns	Obs. nos.	Calculated I	Additional	Obs. regrouped	Calculated II	Values for χ^2
+	643	675.0	...	643	639.4	0.02
O	639	694.2	...	639	678.8	2.33
M	368	306.6	...	368	356.5	0.37
Mc, MMc	40	31.1	MT, McT	42	41.2	0.02
C	158	150.9	...	158	146.7	0.87
Cc, CCc	428	390.4	MCc	434	458.8	1.34
Co	276	298.2	...	276	291.6	0.83
T	138	118.0	...	138	115.9	4.21
OM	132	117.2	...	132	138.3	0.29
OMc	5	7.7	...	5	4.5	0.06
OC	63	60.8	...	63	60.8	0.08
OCc	163	135.5	...	163	135.5	5.58
OCo	125	115.2	...	125	115.2	0.83
OT	54	48.2	...	54	48.5	0.62
MC, McC	23	31.4	...	23	35.6	4.46
MCc	6	65.7
$MCo, McCo$	73	60.1	...	73	67.7	0.41
MT, McT	2	25.8
$McCc$	4	5.6	...	4	2.4	1.07
CCo	30	28.6	...	30	28.6	0.07
CT	11	11.9	...	11	12.2	0.12
$CcCo$	63	64.3	...	63	64.3	0.03
CcT	24	26.5	...	24	26.9	0.31
CoT	24	23.1	...	24	23.0	0.04
Totals	3492	3492.0	...	3492	3492.4	23.96
Gene frequencies						
+		0.4397			0.4279	
O		0.1865			0.1865	
M		0.0906			0.1061	
Mc		0.0073			0.0034	
C		0.0467			0.0466	
Cc		0.1039			0.1039	
Co		0.0884			0.0884	
T		0.0369			0.0372	
Totals		1.0000			1.0000	

all fish fall in one of the groups, MC, McC ; $MCo, McCo$; and MT, McT . Let v and w be the frequencies of M and Mc respectively, and v' and w' be the frequencies these genes would have if the fish in the combinations just mentioned were neither M nor Mc . v' cannot be determined by the formula but w' and $v' + w'$ can. Let d'_1 and d'_2 be

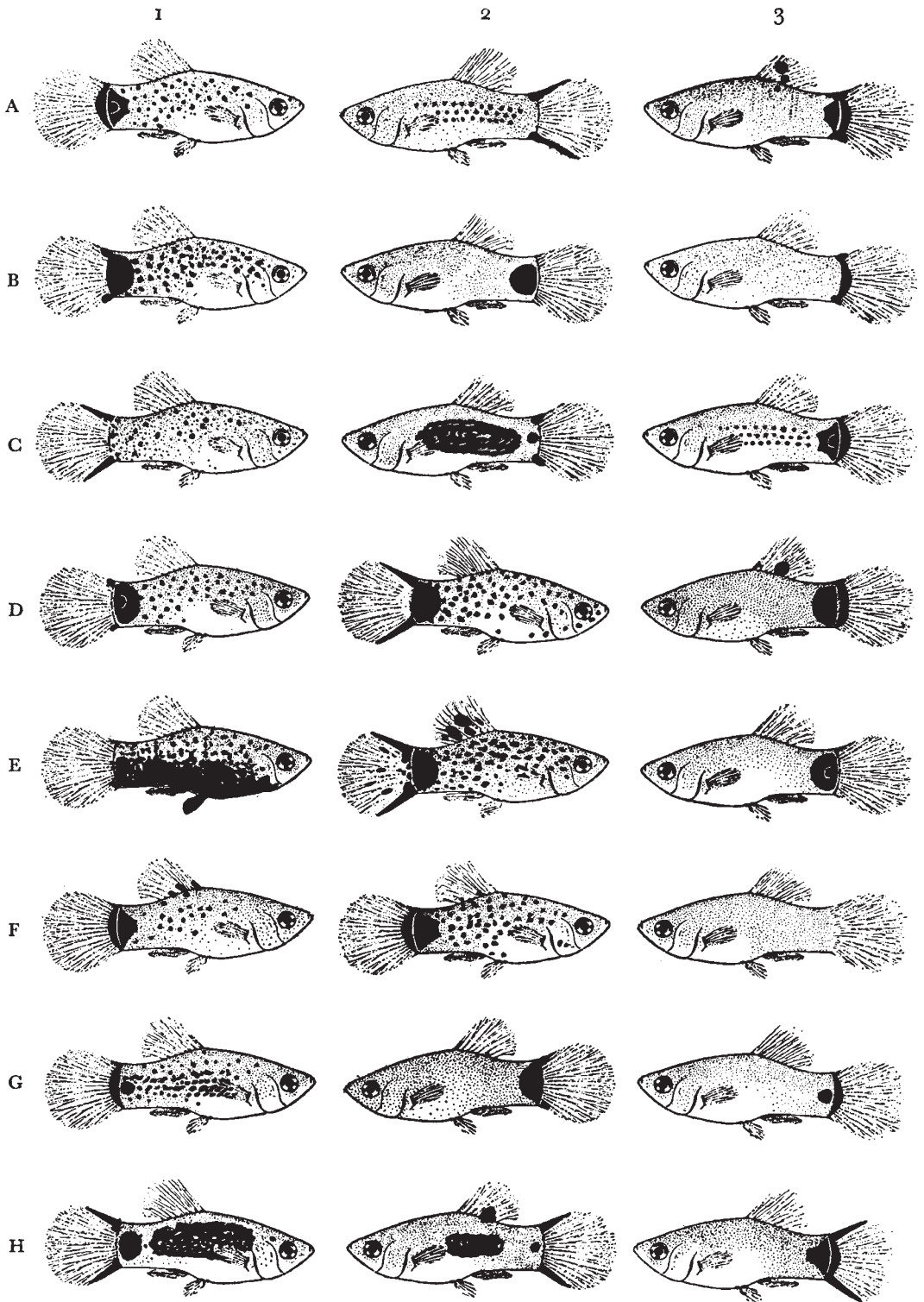


FIG. 1.—A random assortment of platyfish (*Platypoecilus maculatus*) from Mexico and Guatemala showing the great polymorphism in colour patterns. The tail patterns are referable to autosomal genes which form a multiple allelic series. The heavily spotted patterns on the body proper and on the dorsal fin are referable to sex-linked genes which form another multiple allelic series. (For identification of colour genes, see Key to Figures on opposite page.)

the portion of the fish in these combinations which are genotypically *M* and *Mc* respectively, then :—

$$d = d'_1 + d'_2$$

$$\frac{d'_1}{d'_2} = \frac{v}{w} \quad (\text{under theoretical conditions})$$

Let :

$$d_1 = v - v'$$

$$d_2 = w - w'$$

Since only one of the two alleles of each of the fish in the combinations being considered is either *M* or *Mc* :—

$$2d_1 = d'_1$$

$$2d_2 = d'_2$$

$$d'_1 + d'_2 = d = 2(d_1 + d_2)$$

$$\frac{d_1}{d_2} = \frac{v}{w}$$

$$\frac{d_1 + d_2}{d_2} = \frac{v + w}{w}$$

$$\frac{d}{2d_2} = \frac{v + w}{d_2 + w'}$$

$$dd_2 + dw' = 2d_2v + 2d_2w$$

Key to Figures on opposite page.

Position	Autosomal alleles								Sex-linked alleles				
	+	O	M	Mc	C	Cc	Co	T	Sp	Sr	Sd	Sb	N
1. A B C D E F G H	+	O	+	+	C	Cc	+	+	Sp	+	+	+	+
	+	+	+	Mc	+	+	+	+	Sp	+	+	+	+
	+	+	+	+	+	+	Co	+	Sp	+	+	+	+
	+	O	+	Mc	+	+	+	+	Sp	+	+	+	+
	+	+	+	+	+	+	+	+	Sp	+	+	Sb	+
	+	+	+	+	+	Cc	+	+	Sp	+	Sd	+	+
	+	O	+	+	C	+	+	+	Sp	Sr	+	+	+
	+	+	+	Mc	+	+	Co	+	+	+	+	+	N
2. A B C D E F G H	+	+	+	+	+	+	Co	T	+	Sr	+	+	+
	+	+	M	+	+	+	+	+	+	+	+	+	+
	+	O	+	+	+	+	+	T	+	+	+	+	N
	+	+	M	+	+	+	Co	+	Sp	+	+	+	+
	+	+	+	Mc	+	+	Co	+	Sp	+	Sd	+	+
	+	+	+	+	+	Cc	+	+	Sp	+	+	+	+
	+	O	+	Mc	+	+	+	+	+	+	+	+	+
	+	O	+	+	+	+	Co	+	+	+	Sd	+	N
3. A B C D E F G H	+	+	+	+	+	Cc	+	T	+	+	Sd	+	+
	+	+	+	+	C	+	+	T	+	+	+	+	+
	+	O	+	+	+	Cc	+	+	+	Sr	+	+	+
	+	+	M	+	+	+	+	+	+	+	Sd	+	+
	+	O	M	+	+	+	+	+	+	+	+	+	+
	+	+	+	+	+	+	+	+	+	+	+	+	+
	+	O	+	+	C	+	+	+	+	+	+	+	+
	+	+	+	+	+	Cc	Co	+	+	+	+	+	+

$$\begin{aligned}
 d_2[d-2(v+w)] &= -dw' \\
 d_2 &= \frac{dw'}{2(v+w)-d} \\
 d_2 &= \frac{dw'}{2\left(v+w-\frac{d}{2}\right)} \\
 d_2 &= \frac{dw'}{2(v'+w')} \\
 w &= w' + \frac{dw'}{2(v'+w')} \\
 w &= w' \left[1 + \frac{d}{2(v'+w')} \right] \quad . \quad . \quad (1) \\
 v &= (v+w) - w \\
 v &= v' + w' + \frac{d}{2} - w \quad . \quad . \quad (2)
 \end{aligned}$$

(1) and (2) were used to find first w and then v .

For both the Rio Jamapa and the Rio Papaloapan populations the frequency of + was found by subtracting the total of all the other gene frequencies from 1.

In the second calculation of the gene frequencies in the Rio Papaloapan population, some of the phenotypes contained so many genotypes that a different procedure was used to find the gene frequencies. This new method is a generalisation of the use of the formula $a = 1 - \sqrt{1-x}$; this accounts for the values of the frequencies of several genes being the same in both calculations.

Each genotype should occur in the proportion a^2 if it is homozygous or $2ab$ if heterozygous, a and b being the gene frequencies involved. Since:—

$$\begin{aligned}
 (a_1 + a_2 + \dots + a_n)^2 &= a_1^2 + a_2^2 + \dots + a_n^2 \\
 &\quad + 2a_1a_2 + 2a_1a_3 + \dots + 2a_1a_n \\
 &\quad + 2a_2a_3 + 2a_2a_4 + \dots + 2a_2a_n \\
 &\quad + \\
 &\quad \cdot \\
 &\quad \cdot \\
 &\quad \cdot \\
 &\quad + 2a_{n-1}a_n \\
 \text{i.e. } \left(\sum_{i=1}^n a_i \right)^2 &= \sum_{i=1}^n \sum_{j=1}^n a_i a_j
 \end{aligned}$$

the squares of the sums of certain gene frequencies can be found by adding together the proportion observed for each of certain genotypes. The sum of these gene frequencies is found by taking the square root. Because, in some cases, more than one genotype has the same phenotype, not all such sums of gene frequencies can be found. By correctly choosing eight (the number of genes in the allelic series) sums,

equations which can be easily solved for the gene frequencies may be obtained. The eight equations should be chosen so that each contains as many gene frequencies as possible and the equations have one and only one solution.

Let the following symbols represent the gene frequencies of the eight allelic genes :—

Frequency of +	=	p
„ O	=	q
„ M	=	v
„ Mc	=	w
„ C	=	t
„ Cc	=	u
„ Co	=	s
„ T	=	r

The equations used are :—

$$p+q+v+w+t+u+s+r = 1.0000 . \quad . \quad . \quad (1)$$

$$p \quad +v+w+t+u+s+r = 0.8135 . \quad . \quad . \quad (2)$$

$$p+q+v+w+t \quad +s+r = 0.8961 . \quad . \quad . \quad (3)$$

$$p+q+v+w+t+u \quad +r = 0.9116 . \quad . \quad . \quad (4)$$

$$p+q+v+w \quad +s+r = 0.8494 . \quad . \quad . \quad (5)$$

$$p+q \quad +t \quad +s+r = 0.7866 . \quad . \quad . \quad (6)$$

$$p \quad +t \quad +s \quad = 0.5629 . \quad . \quad . \quad (7)$$

In each of these seven equations either both v and w or neither v nor w appears. The eighth equation must give the value of $p+q+v$ or $p+v$, the only such sums which can be obtained and which contain only one of the two frequencies, v and w . There is a deviation of the observed number for O from one that can be calculated on the basis of the gene frequencies of $+$ and O which can be obtained without determining the frequencies of M and Mc . In addition, in finding $p+v$ only two observed values are used. Therefore, another method was used to determine v and w .

The gene frequencies were calculated from the seven equations with the sum of the frequencies of M and Mc being calculated instead of these frequencies. Then values of each phenotype were calculated using these gene frequencies. Genotypes were grouped together as though M and Mc were alike phenotypically so that the sum of their gene frequencies was sufficient for these calculations. The biggest difference between the calculated and observed values where M and Mc are involved was found to occur in the phenotypes M , Mc , MT , McT , MMc . The difference in Cc , CCc , MCc is ignored because the M fish form only a small part of this group. In order to make the goodness-of-fit of the observed values to the calculated numbers as nearly perfect as possible, the difference between them in the case of M , Mc , MT , McT , MMc was divided between M and Mc , MT , McT , MMc in proportion to the observed values for these groups of patterns. This made sure that the total contribution to χ^2 from these

groups was minimal.* Dividing the difference in this way, the calculated value of the M phenotype is obtained. From this value, V , and the frequency of $+$, p , the frequency of M , v , can be obtained because :—

$$2pv + v^2 = V$$

$$v = -p + \sqrt{V + p^2}$$

Then the frequency of M can be subtracted from the sum of the frequencies of M and Mc to find the frequency of Mc .

In each of the two calculations from the Rio Papaloapan population and the one for the Rio Jamapa, after the gene frequencies were found, the number of fish of each pattern which were expected on the basis of these gene frequencies was found where this had not already been done. Except for the first calculation of the Rio Papaloapan population, the goodness-of-fit of the observed values of these expected values was then found by means of the χ^2 test.†

ANALYSIS

If the gene frequencies are obtained for the Rio Papaloapan sample and the number of fish that would be expected to show each pattern on the basis of these gene frequencies is calculated, certain discrepancies between these calculated and the observed numbers

* If C is the total contribution to χ^2 , e_1 and e_2 the calculated values which are to be found, k their known sum, and o_1 and o_2 the respective observed values :—

$$C = \frac{(e_1 - o_1)^2}{e_1} + \frac{(e_2 - o_2)^2}{e_2}$$

$$e_1 + e_2 = k$$

$$C = \frac{(e_1 - o_1)^2}{e_1} + \frac{(k - e_1 - o_2)^2}{k - e_1}$$

$$C = e_1 - 2o_1 + \frac{o_1^2}{e_1} + k - e_1 - 2o_2 + \frac{o_2^2}{k - e_1}$$

$$\frac{dC}{de_1} = 1 - \frac{o_1^2}{e_1^2} - 1 + \frac{o_2^2}{(k - e_1)^2} = 0$$

$$\frac{o_2^2}{(k - e_1)^2} = \frac{o_1^2}{e_1^2}$$

$$\frac{o_2}{e_2} = \pm \frac{o_1}{e_1}$$

It is obvious that when the sign is $+$, this condition gives C its minimum value with positive e_1 and e_2 .

† Dr Sewall Wright, in a letter, suggested that the gene frequencies for the Rio Papaloapan population be calculated by a different method. This method involved grouping $McCc$ with Cc , CCc and MCc . By this method the frequencies of O , Cc and Co were calculated by the formula $a = 1 - \sqrt{1 - x}$. The frequency of C in platyfish without Cc (Cc conceals the presence of C) was found by the formula and multiplied by the total frequency of genes other than Cc to find the frequency of C in the entire population. In the same way, the frequency of T and the sum of the frequencies of M and Mc were determined. The frequency of $+$ was found by subtracting the total of the other frequencies from 1. The total frequency of M and Mc was divided between M and Mc in proportion to the observed frequencies of OM and OMc respectively. The results of this method were, as Dr Wright calculated, approximately the same as those obtained by the method described above.

occur. The most striking of these are between the numbers for the phenotypes MCc and MT , McT . The calculations indicate that more $+$, O , Co , MC and McC , MCc , McT and MT fish should have been observed. They also show that too many M , MMc and Mc , CCc and Cc , T , OM , OC , OCo , $McCo$ and MCo fish were observed as compared to the calculated values. The observed values for C , OMc , OC , OT , $McCc$, CCo , CT , $CcCo$, CcT and CoT are about correct. Here, as in the following similar statements, the deviations are listed without regard to significance; most of the deviations can be considered to be caused by chance. An examination of the appearance (see fig. 1) of the patterns MCc and MT , McT reveals a strong probability that, if observed, most MCc fish would be considered to be Cc alone and most of the MT and McT fish would be taken for Mc alone.

Assuming that this was what occurred, some of the discrepancies can be corrected and others are lessened. The value of χ^2 is now 23.96; the corresponding probability is about 0.07. The observed frequencies of O , Cc and CCc , MCc , Co , MC and McC are too small. Those of M , C , T , OCc and OCo are too large. The observed numbers for $+$, Mc , MMc , McT and MT , OM , OMc , OC , OT , $McCo$ and MCo , $McCc$, CCo , CT , $CcCo$, CcT and CoT are about equal to the calculated numbers.

In the Rio Jamapa population the situation is simplified by the absence of M and Mc . The value of χ^2 is 42.84. The probability is less than 0.001. The deviation between expected and observed values for O is similar to that in the Rio Papaloapan; the other deviations in the Rio Jamapa are dissimilar to the corresponding Rio Papaloapan deviations. The observed values larger than their expected values are CCc , Cc , T and OCo . The observed values smaller than their expected values are O , $CcCo$, CcT , CoT . The nearly equal values occur for $+$, C , Co , OC , OCc , OT , CCo , CT .

DISCUSSION

The calculation for Gordon's (1947) observed values of MCc and MT , McT in the Rio Papaloapan population of platyfish give such a poor fit as to require some correction. If the gene frequency of M was assumed to be low enough for a reasonable fit for MCc , then the calculated values for the other patterns containing M would be so greatly reduced as to make the fit of observed to calculated values extremely bad for these patterns. Therefore, unless a different method of inheritance from that assumed in the calculations is involved, the fish cannot all actually have the patterns they appear to have.

A critical examination of the appearance of the patterns MCc and MT , McT , reveals that MCc would almost be the same as Cc alone and MT and McT would appear like Mc . Combining the MCc fish with the Cc fish, *i.e.* not attempting to separate MCc fish

from Cc fish in either the observations or calculations, and the MT and McT with Mc fish, allows new gene frequencies to be calculated. The new expected values and the regrouped observed values have a goodness-of-fit which is satisfactory. This goodness-of-fit indicates that the hypothesis that MCc fish were counted as Cc and that MT and McT fish were counted as Mc is satisfactory with reference to the consistency of the regrouped observed values with each other as well as the appearance of the patterns. It is therefore likely to be correct.

There are other possibilities. If random values totaling one are chosen for the gene frequencies, and expected values are calculated on the basis of these gene frequencies, some of the patterns will have been observed too frequently and others too infrequently as compared with these calculated values. If it is assumed that some of the fish in the groups with too high observed values actually should be in the groups with too low values, a satisfactory fit can be obtained. Two points must be kept in mind, however. First, except in the case noted above, patterns which seem to be phenotypically distinguishable from each other would have to be considered to be indistinguishable at least some of the time. Second, except in a few cases, the number of patterns which would have to be assumed indistinguishable is large.

There is the possibility that the MCc fish were counted as M , or that some were counted as M and others as Cc , and, as was assumed above, the McT and MT fish as Mc . However, the combination MCc appears like Cc rather than M .

In the Rio Jamapa the deviations of observed from calculated values are clearly significant. No simple explanation of the deviations is apparent. Any explanation must take into account that deviations similar to those in the Rio Jamapa sample do not occur in the Rio Papaloapan sample. Since this is so, the appearance of the single patterns and combinations, which are the same in both rivers, cannot be used to explain the situation in the Rio Jamapa sample as they were used for the Rio Papaloapan sample. (The particular patterns so used for the Rio Papaloapan calculations: MCc , McT and MT , do not occur in the Rio Jamapa population.) If it is assumed that incorrectly calculating the gene frequencies causes the deviations in the cases of $CcCo$ and CcT , correcting the gene frequencies would produce even more significant deviations for at least some of the other patterns containing Cc , Co or T .

The combinations of $CcCo$ and CcT cannot be less viable than the other patterns, unless other patterns of Cc , Co or T are more viable to an extent which just maintains the gene frequency at equilibrium. The fish which should have $CcCo$ and CcT may exist but have patterns other than $CcCo$ and CcT . This shift from one pattern to another would then be assumed to be produced by something which exists in the Rio Papaloapan and not in the Rio Jamapa. This factor might be a difference in the viability of the patterns which causes

deficiency of $CcCo$ and CcT fish while producing an excess of other patterns containing Cc , Co or T . Another possibility is that this factor changes the patterns of genotypically $CcCo$ and CcT fish so that these patterns resemble other patterns. Still other patterns may also be affected by this factor.

In commenting on the manuscript of this paper, Dr Wright said : "The mechanism by which so many alleles are kept in circulation is the most interesting question here. The simplest is, of course, a selective advantage of heterozygotes over both homozygotes. There doesn't seem to be much evidence for this but the amount of advantage needed is so slight that it might be difficult to detect. It can only be done in any case where there is no overlap of classes. Another possible mechanism is a selective advantage of any gene when rare as compared with when common."

Dr Ernst Mayr, in a letter, pointed out that in the case of O in the Rio Jamapa sample the number of fish with O alone was too small compared with the number with O in combination with another pattern. Since all fish homozygous for O are part of the former group, he stated that this discrepancy, although not significant, suggested a selective advantage of heterozygotes over homozygotes. However, other discrepancies occur which can not be explained in this way, e.g. the excess of single C and T in the Rio Jamapa sample. In this sample, 511 fish with single patterns were observed while only 504.9 were expected. Thus no selective advantage of heterozygotes over homozygotes is indicated by the data. It must be pointed out, however, that these data, owing to pattern complexities, are not particularly suitable for the determination of selective advantages of the various genotypes.

CONCLUSIONS

Using the data of Gordon (1947) on phenotypic frequencies of populations of the polymorphic Mexican platyfish (*Platypoecilus maculatus*) from the Rio Jamapa and Rio Papaloapan, the gene frequencies were calculated by several formulæ derived from Hardy's Formula. These gene frequencies were then used to calculate the ratios of all the phenotypes that would be expected from them and the goodness-of-fit, of the observed phenotypic ratios to the calculated ones, determined by the χ^2 test.

The seven autosomal, dominant, allelomorphic genes for tail patterns found in the Rio Papaloapan platyfish can be recognised when occurring singly, but some of their combinations are indistinguishable. That the patterns MCc , McT and MT were not observed in sufficient numbers is indicated. Fish with these patterns were probably considered to have the pattern Cc when they had MCc , and Mc when they had McT and MT . If the observed numbers of fish are regrouped on the supposition that MCc fish cannot be distinguished phenotypically from Cc fish in most cases, nor McT and MT fish from Mc ones, the

various combinations and single patterns are found to occur in frequencies consistent with each other.

In the Rio Jamapa population the complicating patterns M and Mc are missing. This simplified its analysis somewhat. However, the dual CCc pattern which occurs in the Rio Jamapa population is indistinguishable from the single Cc pattern. Allowing for this difficulty, discrepancies still exist among the observed frequencies of single patterns and the various combinations of those patterns. An explanation for this is lacking at the present time.

The Rio Papaloapan and Rio Jamapa populations differ with respect to whether the tail patterns $CcCo$ and CcT occur with a frequency consistent with that predicated from the gene frequencies of Cc , Co and T .

In both populations the frequency of single patterns, as employed by Gordon (1947), is an adequate indication of the gene frequency of the gene responsible for that pattern.

SUMMARY

Seven autosomal, dominant multiple alleles produce a series of seven pigment patterns in the tail region of *Platypoecilus maculatus*, the common aquarium platyfish. In most cases, but not in all, different genotypes for tail patterns may be phenotypically distinguished. The observed frequencies of the various patterns were compared with each other by calculating the frequencies of the corresponding genes. In the Rio Papaloapan population the frequencies of the patterns are consistent with each other if certain similar appearing patterns are grouped and evaluated together. In the Rio Jamapa population, on the other hand, the frequencies are not consistent with each other.

In the two platyfish populations analysed, from the Rio Jamapa and the Rio Papaloapan (and presumably also from two other rivers, the Rio Coatzacoalcos and the Rio Usumacinta) the determination of the frequency of an individual pattern, uncombined with other patterns, is an adequate indication of its corresponding gene frequency.

The authors are indebted to Dr Howard Levene, Columbia University, for the following comment: "It should be noted that the fit would have been somewhat better, and χ^2 would have been reduced, if the gene frequencies had been estimated by the method of maximum likelihood. However, the χ^2 for the Rio Jamapa population would almost certainly still be significant. Since the conclusions reached would be essentially unaffected, and the uncertainties of classification make extreme precision impossible, the considerable labor of computing maximum likelihood estimates does not seem necessary."

REFERENCES

- GORDON, MYRON. 1947. Speciation in fishes. Distribution in time and space of seven dominant multiple alleles in *Platypoecilus maculatus*. *Advances in Genetics*, 1, 95-132.

GORDON, MYRON. 1948.

Effects of five primary genes on the site of melanomas in fishes and the influence of two color genes on their pigmentation. In "The Biology of Melanomas." *Spec. Pub. N.Y. Acad. Sci.*, 4, 216-268.

HARDY, G. K. 1908.

Mendelian proportions in a mixed population. *Science*, 28, 49-50.