THE INHERITANCE OF THE *DUFFY* BLOOD GROUPS : AN ANALYSIS OF 110 ENGLISH FAMILIES

R. R. RACE and RUTH SANGER

Medical Research Council Blood Group Research Unit, Lister Institute of Preventive Medicine, London, S.W. 1

Received 7.ix.51

THE *Duffy* blood group system was discovered by Cutbush, Mollison and Parkin in 1950. The anti-*Duffy* serum, anti-Fy^a, divides human beings into two phenotypes, those whose red cells are sensitised, $Fy(a+)$, and those whose red cells are not, $Fy(a-)$.

Cutbush and Mollison (1950) tested 27 families with 49 children and demonstrated that the antigen Fy^a is inherited by means of a gene which expresses itself in single as well as in double dose. This was confirmed by Race, Holt and Thompson (1951) who tested 58 families with 148 children.

Anti-Fy^a has not yet been shown to make a dosage distinction between bloods representing the genotype Fy^aFy^a and those representing the genotype Fy^aFy^b ; nor has an antiserum anti-Fy^b yet been identified. Fy^a is therefore in the phase, usually fleeting for blood group antigens, in which it may properly be called a dominant Mendelian character.*

The present paper reports the results of testing a further 52 families for the *Duffy* groups. These families are shown in table 1, together with the 58 families of Race, Holt and Thompson ; the total of 110 families with 274 children are analysed below.

Ascertainment.—The families of table 1 were quite unselected as far as blood groups are concerned. There was some selection for families with two or more children.

Twins.—When a pair of twins was not shown to be dizygous by sex or by any of the blood groups it has been scored as one child.

Other blood group systems.—The families have also been tested for the A_1 A_2 *B O*, *MNS*, *P*, *Rh*, *Lutheran*, *Kell* and *Lewis* groups. About half of them have also been tested for the *Kidd* groups.

Linkage.—In the previous paper (Race *et al.*, 1951) it was reported that no evidence could be found for linkage between the *Duffy* genes and the genes for other blood groups, or for phenyl thio-carbamide testing ; nor was there any evidence for partial sex linkage of the *Duffy* genes. The present addition of 52 families supports this independent segregation of the *Duffy* genes. Linkage calculations are not given for they form part of a paper on linkage and the blood groups at present in preparation (Holt, Thompson, Sanger and Race).

* Since the writing of this paper both an anti-Fy^a showing a dosage effect and an example of anti-Fy^b have been found.

GENE FREQUENCIES

The gene frequencies used in the calculations to follow are derived from the frequency of the *Duffy* phenotypes of unrelated English

TABLE 1

The Duffy *blood groups of 110 English families with 274 children*

Group number	Father	Mother	Number of families	Children	
				Fy(*a*+)	*Fy*(*a*−)
1	*Fy*(*a*+)	*Fy*(*a*+)	8	1	—
			18	2	—
			5	3	—
			1	4	—
2	*Fy*(*a*+)	*Fy*(*a*+)	5	1	1
			3	1	2
			1	2	1
			1	3	1
			1	3	2
			1	6	2
3	*Fy*(*a*+)	*Fy*(*a*−)	1	1	—
			3	2	—
			3	3	—
			2	6	—
	Fy(*a*−)	*Fy*(*a*+)	3	1	—
			4	2	—
			1	4	—
			1	5	—
4	*Fy*(*a*+)	*Fy*(*a*−)	1	—	1
			4	—	2
			4	1	1
			2	1	2
			1	2	2
			1	3	2
	Fy(*a*−)	*Fy*(*a*+)	3	—	1
			3	—	2
			5	1	1
			1	—	3
			3	1	2
			4	2	1
			2	1	3
			1	2	3
5	*Fy*(*a*−)	*Fy*(*a*−)	6	—	2
			5	—	3
			1	—	4
			1	—	5

people (table 2). The item " present series " represents tests done in this Unit on unrelated, unselected persons since the previous publication.

Assuming that the phenotype *Fy*(*a*−) represents the genotype Fy^bFy^b then the gene frequencies may be derived in the following way :

the frequency of the gene $Fy^b = \sqrt{0{\cdot}3401} = 0{\cdot}5832$

and the frequency of the gene $Fy^a = 1 - 0{\cdot}5832 = 0{\cdot}4168$

TABLE 2

The frequency of the Duffy *phenotypes in England*

	Number tested	*Fy*(*a*+)	*Fy*(*a*−)
Cutbush and Mollison (1950) . .	205	133 *64·88%*	72 *35·12%*
Race, Holt and Thompson (1951) .	255	167 *65·49%*	88 *34·51%*
Present series	325	218 *67·08%*	107 *32·92%*
Total	785	518 *65·99%*	267 *34·01%*

The genotype frequencies are therefore :—

$$Fy^aFy^a \quad 0{\cdot}4168^2 = 0{\cdot}1737$$
$$Fy^aFy^b \quad 0{\cdot}4168 \times 0{\cdot}5832 \times 2 = 0{\cdot}4862$$
$$Fy^bFy^b \quad 0{\cdot}5832^2 = 0{\cdot}3401$$

These frequencies have been used to analyse the family material in two ways.

METHOD 1

Table 3 shows the expected incidence of the various matings and the relative frequency of *Fy*(*a*+) and *Fy*(*a*−) children of these matings. These expectations are applied in table 4 to the 110 families and their 274 children. It will be seen that there is good agreement between the calculated distribution and that actually found.

Professor Fisher has pointed out certain objections to this method which has almost become the standard way of verifying unifactorial inheritance. The mating *Fy*(*a*+) × *Fy*(*a*+) is genotypically of three kinds, and the mating *Fy*(*a*+) × *Fy*(*a*−) of two kinds. Children from say the phenotype mating *Fy*(*a*+) × *Fy*(*a*−) are not independent samples from a homogeneous population, with an expectation of 0·6316 *Fy*(*a*+) to 0·3684 *Fy*(*a*−), and should not be treated as independent in calculating goodness of fit. That this method does in fact usually give a very good fit is probably due to the shortage of large families in the data.

METHOD 2 (*a*)

In 1939 Professor Fisher suggested that account should be taken separately of the totals yielded by the group of families containing recessive children, and of the individual sizes of families containing none. This method was applied to the A_1 A_2 B O groups by Taylor and Prior (1939) and by Race, Ikin, Taylor and Prior (1942). Owing to an unfortunate lapse of memory the method was not referred to in *Blood Groups in Man* (Race and Sanger, 1950).

Two separate comparisons with expectation are made. In (*a*), we compare the number of families observed to contain no recessive children, irrespective of size, with the number expected, taking account

of the estimated gene frequency and of the sizes of all families observed. In (*b*), we compare the numbers of recessive children observed in

TABLE 3

The expected distribution of the Duffy *groups in English parents and offspring*

Genotypes				
Matings		Children		
Type	Frequency	Fy^aFy^a	Fy^aFy^b	Fy^bFy^b
$Fy^aFy^a \times Fy^aFy^a$	*0·0302*	0·0302	...	...
$Fy^aFy^a \times Fy^aFy^b$	*0·1689*	0·08445	0·08445	...
$Fy^aFy^b \times Fy^aFy^b$	*0·2364*	0·0591	0·1182	0·0591
$Fy^aFy^a \times Fy^bFy^b$	*0·1181*	...	0·1181	...
$Fy^aFy^b \times Fy^bFy^b$	*0·3307*	...	0·16535	0·16535
$Fy^bFy^b \times Fy^bFy^b$	*0·1157*	...	...	0·1157
	1·0000			

Phenotypes			
Matings		Children	
Type	Frequency	*Fy*(*a*+)	*Fy*(*a*−)
Fy(*a*+) × *Fy*(*a*+)	*0·4355*	0·8643	0·1357
Fy(*a*+) × *Fy*(*a*−)	*0·4488*	0·6316	0·3684
Fy(*a*−) × *Fy*(*a*−)	*0·1157*	...	1·0000
	1·0000		

TABLE 4

The Duffy *groups of 110 English families with 274 children analysed by method 1*

Matings			Children				
Type	Number		Total number	*Fy*(*a*+)		*Fy*(*a*−)	
	Observed	Expected		Observed	Expected	Observed	Expected
Fy(*a*+) × *Fy*(*a*+)	44	*47·9*	**102**	85	*88·2*	17	*13·8*
Fy(*a*+) × *Fy*(*a*−)	53	*49·4*	**136**	79	*85·9*	57	*50·1*
Fy(*a*−) × *Fy*(*a*−)	13	*12·7*	**36**	0	*0·0*	36	*36·0*
	110	*110·0*	**274**				

those families which contain any, with the numbers expected on Mendelian theory, when allowance has been made for the fact that

at this stage, each of these families is known to contain at least one recessive child.

The formulæ given by Taylor and Prior (1939) for the matings $B \times O$ and $B \times B$ can be directly applied to matings $Fy(a+) \times Fy(a-)$ and $Fy(a+) \times Fy(a+)$. These formulæ are also applicable to family studies of the *P*, *Lutheran* and *Kidd* blood groups where

a represents the gene frequency of Fy^a (or *P*, Lu^a or Jk^a)
and b represents the gene frequency of Fy^b (or *p*, Lu^b or Jk^b)

The explanation given below is almost a direct transcription from Taylor and Prior's lucid account of the method, with the necessary change of symbols.

Mating Fy(a+)×Fy(a−)

The $Fy(a+)$ parent can be one of two genotypes, Fy^aFy^a, the frequency of which is a^2, or Fy^aFy^b with the frequency of $2ab$. The frequency of the phenotype $Fy(a+)$ in the population is therefore a^2+2ab ; hence the probability of an $Fy(a+)$ person being Fy^aFy^a is $\frac{a^2}{a^2+2ab}$ or $\frac{a}{a+2b}$, and of being Fy^aFy^b is $1-\frac{a}{a+2b}$. From the mating $Fy^aFy^a \times Fy^bFy^b$ all children will be $Fy(a+)$, from the mating $Fy^aFy^b \times Fy^bFy^b$ some children may be $Fy(a-)$. The probability of a child of the mating $Fy^aFy^b \times Fy^bFy^b$ being of the phenotype $Fy(a+)$ is $\frac{1}{2}$, of two children being both $Fy(a+)$ is $(\frac{1}{2})^2$ and of n children all $Fy(a+)$ is $(\frac{1}{2})^n$. The occurrence of an $Fy(a-)$ child is at present the only means of knowing that the $Fy(a+)$ parent is heterozygous, hence in the mating $Fy(a+) \times Fy(a-)$ the families are divided into two groups, those with all children $Fy(a+)$, and those with some children $Fy(a-)$.

All children Fy(a+)

Families of 1 child.—Probability of this child being $Fy(a+)$ = probability of the $Fy(a+)$ parent being $Fy^aFy^a + \frac{1}{2} \times$ (probability of the $Fy(a+)$ parent being Fy^aFy^b) $= \frac{a}{a+2b} + \frac{1}{2}\left(1 - \frac{a}{a+2b}\right)$

Therefore the expected number of families with one child, this child being $Fy(a+)$

$$= \begin{bmatrix}\text{observed number of } Fy(a+) \times Fy(a-) \\ \text{families with one child.}\end{bmatrix} \times \left[\frac{a}{a+2b} + \frac{1}{2}\left(1 - \frac{a}{a+2b}\right)\right]$$

Since a = the frequency of the gene $Fy^a = 0{\cdot}4168$, and b = the frequency of the gene $Fy^b = 0{\cdot}5832$ and the total number of 1 child families of the mating type $Fy(a+) \times Fy(a-)$ is 8, then the expected number of such families with 1 child, this child being $Fy(a+)$ is $8 \times 0{\cdot}6316$ or $5{\cdot}0528$; 4 were observed.

Families of n *children.*—Probability of all n children being

$$Fy(a+) = \frac{a}{a+2b} + \left(\frac{1}{2}\right)^n \left(1 - \frac{a}{a+2b}\right)$$

As above, to get the expected number of such families this figure must be multiplied by the observed number of families of n children with parents $Fy(a+) \times Fy(a-)$.

TABLE 5

Analysis of families of the mating type Fy(a+) × Fy(a−)

Family size	Total families	Number of families having only $Fy(a+)$ children	
		Expected	Observed
1	8	5·05	4
2	23	10·29	7
3	13	4·62	3
4	4	1·24	1
5	3	0·86	1
6	2	0·55	2
Total . .	53	22·61	18

Applying these formulæ to the matings $Fy(a+) \times Fy(a-)$ shown in table 1, the expectations for the number of families in which all children are $Fy(a+)$ are as shown in table 5.

Some children Fy(a−)

The expected number of families with $Fy(a-)$ children is obtained by subtracting the expected number of families whose children are all $Fy(a+)$ from the total number of observed families of the type $Fy(a+) \times Fy(a-)$. Applying this to the present investigation we find that the expected number of such families is 53—22·61 or 30·39, while 35 were observed.

Mating Fy(a+) × Fy(a+)

In this mating also the parents can be of two genotypes Fy^aFy^a or Fy^aFy^b and the families are again divided into two groups, those with all children $Fy(a+)$, and those with some children $Fy(a-)$.

Probability of one parent being $Fy^aFy^b = \frac{2b}{a+2b}$.

Probability of both parents being $Fy^aFy^b = \left(\frac{2b}{a+2b}\right)^2$.

Thus the probability of at least one parent being

$$Fy^aFy^a = 1 - \left(\frac{2b}{a+2b}\right)^2.$$

All children Fy(a+)

Families of 1 child.—Probability of this child being $Fy(a+)$ = the probability of at least one parent being $Fy^aFy^a + \frac{3}{4} \times$ (the probability of both parents being Fy^aFy^b) $= 1 - \left(\frac{2b}{a+2b}\right)^2 + \frac{3}{4}\left(\frac{2b}{a+2b}\right)^2$

Therefore the expected number of families with 1 child, this child being $Fy(a+)$

$$= \begin{bmatrix}\text{observed number of}\\ \text{families with 1 child.}\end{bmatrix} \times \left[1 - \left(\frac{2b}{a+2b}\right)^2 + \frac{3}{4}\left(\frac{2b}{a+2b}\right)^2\right]$$

Families of n *children.*—Probability of all children being

$$Fy(a+) = 1 - \left(\frac{2b}{a+2b}\right)^2 + \left(\frac{3}{4}\right)^n \left(\frac{2b}{a+2b}\right)^2$$

Applying these formulæ to the matings $Fy(a+) \times Fy(a+)$ shown in table 1 the expectations for the number of families in which all children are $Fy(a+)$ are as given in table 6.

TABLE 6

Analysis of families of the mating type Fy(a+)×Fy(a+)

Family size	Total families	Number of families having only $Fy(a+)$ children	
		Expected	Observed
1	8	*6·91*	8
2	23	*17·54*	18
3	9	*6·17*	5
4	2	*1·26*	1
5	1	*0·59*	0
8	1	*0·51*	0
Total . .	44	*32·98*	32

Some children Fy(a−)

The expected number of families with $Fy(a-)$ children will be the difference between the total number of observed families of the type $Fy(a+) \times Fy(a+)$ and the expected number of such families whose children are all $Fy(a+)$. In the present example the expected number of such families is 44—32·98 or 11·02, while 12 were observed.

A summary of the results of this method of family analysis is given in table 7.

METHOD 2 (*b*)

In the families $Fy(a+) \times Fy(a+)$ with a child $Fy(a-)$ it is possible to calculate the additional number of $Fy(a-)$ children expected.

Table 1 shows that there are 12 such matings with 22 *Fy*(*a*+) children and 5 additional *Fy*(*a*−) children (after the 12 diagnostic children have been excluded). The expected number of additional *Fy*(*a*−) children is simply $\frac{1}{4}(22+5)$ or 6·75.

TABLE 7

Analysis of certain of the families by method 2 (a)

Class of mating	Class of family	Expected number of families	Observed number of families	χ^2	d.f.
Fy(*a*+) × *Fy*(*a*+)	All children *Fy*(*a*+)	*32·98*	32	0·1167	1
	Some children *Fy*(*a*−)	*11·02*	12		
Fy(*a*+) × *Fy*(*a*−)	All children *Fy*(*a*+)	*22·61*	18	1·6373	1
	Some children *Fy*(*a*−)	*30·39*	35		
				1·7540	2
				probability = 0·40	

Similarly there are 35 families *Fy*(*a*+) × *Fy*(*a*−) with a child *Fy*(*a*−). In these families there are 31 *Fy*(*a*+) and 22 additional *Fy*(*a*−) children (again after the exclusion of the 35 diagnostic children). The expected number of additional *Fy*(*a*−) children is $\frac{1}{2}(31+22)$ or 26·50.

The analysis by Method 2 (*b*) is summarised in table 8.

TABLE 8

Analysis by method 2 (b) of children from families containing Fy(a−) *children*

Matings	Children				
		Fy(*a*+)	Additional *Fy*(*a*−)	χ^2	d.f.
Fy(*a*+) × *Fy*(*a*+)	Expected	20·25	6·75	0·6049	1
	Observed	22	5		
Fy(*a*+) × *Fy*(*a*−)	Expected	26·50	26·50	1·5283	1
	Observed	31	22		
				2·1332	2
				probability = 0·34	

The purpose of such analyses as those described in this paper is in the first place to disclose how a newly discovered blood group system is inherited. Conversely, once the manner of inheritance is established beyond doubt, good agreement between calculated and observed frequencies encourages confidence that the tests are being properly done. Furthermore, though the way of inheritance of a blood group system may be known with certainty, statistical analysis may reveal unexpected complications, such as for example, a differential survival rate.

SUMMARY

The results are given of testing a further 325 unrelated English persons with anti-Fy^a. These bring the published number of such persons tested to 785, of which 518 or 65·99 per cent. were positive. This corresponds to the gene frequencies $Fy^a = 0{\cdot}4168$ and $Fy^b = 0{\cdot}5832$.

The *Duffy* blood groups of a further 52 families are presented; these families together with 58 previously reported from this Unit are submitted to statistical analysis. A method of analysis for the A_1 A_2 B O groups, suggested by Fisher in 1939, is applied to the *Duffy* system. The family results are in good agreement with those expected on the assumption that the antigen Fy^a is controlled by a gene which expresses itself in single as well as in double dose.

Acknowledgments.—The authors have to thank Professor Fisher for any originality which this paper may contain.

For gifts of the serum anti-Fy^a we are indebted to Dr P. L. Mollison and Miss Marie Cutbush of the Medical Research Council Blood Transfusion Research Unit, to Dr P. Vogel and Dr R. E. Rosenfield of New York and to Dr J. J. van Loghem, Jr. of Amsterdam. For most of the other anti-sera used in the investigation we are indebted to Dr A. E. Mourant of the Medical Research Council Blood Group Reference Laboratory.

We are particularly grateful to the members of the families for their blood, and for their encouraging interest and cheerful tolerance of our intrusion.

REFERENCES

CUTBUSH, MARIE, AND MOLLISON, P. L. 1950. The *Duffy* blood group system. *Heredity*, *4*, 383-389.

CUTBUSH, MARIE, MOLLISON, P. L., AND PARKIN, DOROTHY M. 1950. A new human blood group. *Nature*, *165*, 188.

HOLT, HELENE A., THOMPSON, JOAN S., SANGER, RUTH, AND RACE, R. R. 1952. Linkage relations of the blood group genes of man : an analysis of 487 families with two or more children. *Heredity*, in the Press.

RACE, R. R., HOLT, HELENE A., AND THOMPSON, JOAN S. 1951. The inheritance and distribution of the *Duffy* blood groups. *Heredity*, *5*, 103-110.

RACE, R. R., IKIN, ELIZABETH W., TAYLOR, G. L., AND PRIOR, AILEEN M. 1942. A second series of families examined in England for the A_1 A_2 B O and *MN* blood group factors. *Ann. Eugen.*, Lond., *11*, 385-394.

RACE, R. R., AND SANGER, RUTH. 1950. *Blood Groups in Man.* Oxford : Blackwell Scientific Publications.

TAYLOR, G. L., AND PRIOR, AILEEN M. 1939. Blood groups in England. III. Discussion of the family material. *Ann. Eugen.*, Lond., *9*, 18-44.