THE RHESUS FACTOR
A STUDY IN SCIENTIFIC METHOD

By R. A. FISHER
University of Cambridge

RECENT progress in the understanding of the human blood group factors has been so rapid that an account of the story as it unfolded itself to English workers during the war years may be of interest to American scientists in diverse fields. Neither I, nor, I suppose, anyone else could give a full account of the present condition of the researches which are being prosecuted, and such an account, if it were possible, would in any case be very complex, and difficult to those not concerned as specialists with genetics or with serology. The limitation of my subject to the particular line of advance of the group of English workers centred at the Galton Laboratory Serum Unit, operating at Cambridge from the outbreak of war in 1939 to the Summer of 1946, will, I hope, enable me to tell a clearer tale, and to lay stress upon the exact line of reasoning by which theoretical views were from time to time expanded and fitted to assimilate the next group of experimental facts that were ascertained.

A number of genetic factors recognisable by serological tests are now known. Typically the red cells of some genotypes are agglutinated, and those of others not agglutinated, by a testing fluid, usually a serum containing a chemically active substance, known as an antibody. In the older systems it has been customary to designate the *antibody* by a Greek letter, such as α, corresponding with the *antigen,* designated by the Latin letter, A. Genotypes developing the antigen A may then be recognised by a test fluid containing α in sufficient concentration, or potency, provided no other active antibody is present in the reagent. Moreover, family studies of the genetics of the blood groups have shown it to be generally, though perhaps not invariably, true that the genotypes which develop the antigen A are those which contain a particular hereditary particle of the germ plasm, or gene, which Ford has in this case designated by the symbol G^A.

Thus the antigens responsible for the spontaneous antibodies of the system discovered by Landsteiner about 1900 [1], are known as A_1, A_2, B, and O, and are inherited as if determined by a set of four allelomorphs of the same Mendelian factor. The antigens of the second important blood group factor in Man, known as M and N, and discovered by Landsteiner and Levine [2], are apparently determined by a pair of allelomorphic genes L^M and L^N, and react with antibodies known by the Greek letters μ and ν.

Other blood group factors are known, and more will certainly be discovered. The two I have mentioned, together with the *Rhesus* factor, perhaps the most interesting and important of all, touch our lives at many points, and have already done much to clarify our ideas about

the human race in general. I cannot here do justice to the contributions which have been and are being made to blood group studies, but it may be worth while very briefly to list the points of view from which they are of importance:

(i) They constitute stepping stones in human genetics, by providing markers on several chromosomes and facilitating linkage studies and the genetic mapping of the human germ plasm.

(ii) They are of great medical importance, primarily in making blood transfusion possible, but also in guarding against its possible dangers. The *Rhesus* factor also has elucidated the nature of haemolytic disease, which has, until now, been a serious cause of infantile and foetal mortality.

(iii) They are of forensic importance in the recognition of individuals, in the recognition of parenthood as in the mistaken interchange of children, and in the recognition of disputed paternity.

(iv) All three factors show important differences of frequency among different human races, and cannot in the future be ignored in ethnographic studies.

Rhesus Factor

The Rhesus factor was discovered by Levine in 1939 [3], in a case of stillbirth in which the mother's serum was found to contain an antibody capable of agglutinating the red cells of about 85% of white American donors. Shortly afterwards Landsteiner and Wiener [4], showed that a similar antibody could be induced in rabbits and guinea-pigs by the injection of the blood of the Rhesus Monkey, *Macacus rhesus*. This antibody I shall for reference call Δ or anti-*D*. It was quickly recognised to be causally associated with foetal injury and haemolytic disease, in cases in which the father and the child possessed some antigen absent in the mother; in other words, the father and child were Rhesus positive (+ve), the mother Rhesus negative (—ve), and the mother has reacted to the presence of the foreign antigen in the foetus by generating the corresponding antibody.

The train of thought suggested by the simple facts stated above to a serologist familiar with gene frequencies, and, so to speak, genotype-conscious, is as follows:

(i) The serum distinguishes two phenotypes in known proportions:
The Rhesus positive (+ve), including 85%.
The Rhesus negative (—ve), including 15%.

(ii) The Rhesus negative phenotype lacks a hypothetical gene *R*, responsible for the reaction, and must consist wholly of the genotype *rr*. The frequency of *r* genes must, therefore, be $\sqrt{.15}$, or 39%, and the frequency of *R* genes must be 1 — .39, or 61%.

(iii) With mating at random with respect to the Rhesus factor the frequencies of the genotypes are calculable:
RR must be $(.61)^2$, or 37%; that of
rr is $(.39)^2$, or 15%; and, therefore, that of the heterozygote
Rr must be 1 — .37 — .15, or 48%.

These inferences will control our deductions from the next group of experimental facts that were discovered.

In May 1942 Taylor and Mollison [5], issued an appeal in the *British*

Medical Journal for sera from the mothers of children suffering from haemolytic disease. The response was good, and a large number of sera became available for testing. Of those capable of easy classification, the great majority, as was to be expected, contained antibody Δ (reacting with 85% of the English population) in various concentrations, and with various admixtures with which I need not trouble you. One serum in particular, however, gave strikingly different results, and evidently contained a different antibody, which I shall denote by γ or anti-*c*. The γ serum reacted positively with about 81% of English donors. The difference in percentage is unimportant, and would, indeed, need a very large sample to be significant. The difference appeared clearly, however, when the same series of donors was classified simultaneously with both sera. It then appeared that all Rhesus negative donors, or, as I should now say, Δ —ve donors, were γ +ve. Expressed as a two by two table, we find

TABLE I—SIMULTANEOUS REACTIONS TO Δ AND γ

	Δ +*ve*	Δ —*ve*	*Total*
γ +ve	66	15	81
γ —ve	19	0	19
Total	85	15	100

The interpretation of this table was made possible by family studies, which led to the important generalisation that not only were all the Δ —ve persons γ +ve, but also the children and the parents of Δ —ve persons were also γ +ve. This could only be interpreted as proving that the 48% designated by the genotype *Rr* was included in the 66% reacting positively with both sera, and that 18%, or about half of the 37% designated by *RR*, were γ +ve while the remainder, or 19%, were γ —ve. There must, therefore, be a genetic distinction among those supposedly *RR* persons. There must be at least two genes, one, let us say R_1, negative to γ, and the second, R_2, reacting positively with it.

This conclusion was decisive for the further course of the research. Six distinct genotypes came into view, with frequencies calculable on the principle explained above. There were now conceived:

TABLE II—FREQUENCIES OF SIX HYPOTHETICAL GENOTYPES

Genes		*Genotypes*					
R_1	44%	R_1R_1	19%	γ —ve	R_2R_2	3%	
R_2	17%	R_1R_2	15%		R_2r	13%	
r	39%	R_1r	35%		rr	15%	Δ —ve

The first effect of the conception of these six genotypes, all except R_2R_2 being fairly common, was the recognition of two new kinds of

antibody. The reactions postulated for the three genes were those shown in the first two columns of the following table:

TABLE III—REACTIONS OF FOUR ANTIBODIES WITH THE THREE COMMON ALLELES

	Δ	γ	H	Γ
R_1	+	—	—	+
R_2	+	+	+	—
r	—	+	—	—

The new antibodies showed the reactions of the 3rd and 4th columns. H or anti-*E* reacted only with R_2, alone or in combination, Γ or anti-*C* reacted only with R_1. The recognition of what these new sera were doing was easy in the light of the genotypes already postulated. Thus H reacted with about 30% of all donors, but the rough percentages written above show a total of 31% for the three genotypes containing R_2; moreover all of these ought to be positive with both the earlier tests, which could scarcely happen if the new antibody was unrelated to the genes already recognised. Similarly, Γ reacted with about 70%, and these include all the γ —ve and none of the Δ —ve. The finding of these new sera thus strongly confirmed the view that six main genotypes were present, and indeed enabled all to be recognized by objective tests, with the exception that the rare genotype R_2R_2 was still indistinguishable from the commoner R_2r.

The new sera not only confirmed the position won so far, but in conjunction with those previously in use provided a network of tests sufficiently fine to catch and identify certain rare types of donor blood which did not fit into the five big classes already distinguishable. About 4% of those donors tested with all four sera gave results in one way or another anomalous, and with increased experience in the reliability of the different sera available it was possible to recognize four new genes, all rare, constituting together only about 5% of the total gene frequency, yet with reactions as distinctive as the three more abundant genes first recognized. These are shown in the enclosed portion of the following table:

TABLE IV—REACTIONS OF SEVEN ALLELES TO THE FOUR ANTIBODIES FIRST KNOWN, WITH EXTENSIONS SUGGESTED

	Δ	γ	H	Γ	δ	η
R_1	+	—	—	+	—	+
R_2	+	+	+	—	—	—
r	—	+	—	—	+	+
R_0	+	+	—	—	—	+
R''	—	+	+	—	+	—
R'	—	—	—	+	+	+
R_z	+	—	+	+	—	—
R_y	—	—	+	+	+	—

In the cases of R_0, R'' and R' the new genes were found in conjunction with r, for each of these gives in the presence of r a distinctive series of reactions. Since r reacts positively only with γ, it is only the reaction with γ of the new genes which could be uncertain. The reagent γ was, however, shown to possess the valuable property of giving distinctly stronger reactions to the double dose of antigen compared with the single dose. From this it was at once judged that R_0 and R'' were γ +ve and R' alone γ —ve. These judgements were, of course, capable of confirmation when the rare gene appeared in more than one member of the same family.

The case of R_z was more difficult, and, as it turned out, more instructive; it was found in conjunction with R_1; the reactions γ —ve, H +ve were distinctive, and showed that a new gene had been found; the reactions with Δ and Γ were, however, uncertain. A consideration of the whole system, as shown below, led, however, to the expectation that two genes might be expected to exist having respectively the reactions assigned in the table to R_z and R_y. Only later, when a family was found in which the new gene was exhibited combined with R_1 in one member and with r in another, was it clear which had been found. It proved to be that to which the symbol R_z had been assigned. R_y has not so far been discovered; it is certainly extremely rare, though its gene frequency may be estimated as about 50 per million.

The reasons for confidence in assigning at this stage reaction systems to the hypothetical genes R_z and R_y may be made clear by rearranging the data of Table IV so as to find, not the reaction system of a given gene, but instead which gene has a given reaction system. In building up such a table we recognize that for the reagents Δ and H every gene must give either ++, +—, —+, or ——; we may therefore assign the rows of the table to these possibilities. Equally, with the reagents Γ and γ we must have the same four conceivable pairs of reactions, and these are the columns of the table. The table thus embraces every possible system of reactions to the four reagents. The systems actually found are set out below:

TABLE V—THE ALLELES HAVING GIVEN REACTIONS

	$\Gamma+$ $\gamma+$	$\Gamma+$ $\gamma-$	$\Gamma-$ $\gamma+$	$\Gamma-$ $\gamma-$
$\Delta+$ H+	—	R_z	R_2	—
$\Delta+$ H—	—	R_1	R_0	—
$\Delta-$ H+	—	(R_y)	R″	—
$\Delta-$ H—	—	R′	r	—

Obviously, every known gene reacts either with Γ or γ, never with both or neither. The corresponding antigens C and c are mutually exclusive alternatives, like a pair of allelomorphic genes. In combination with this pair we may have either D or not-D, and either E or not-E.

Setting, as in ordinary Mendelian analysis, *d* as the alternative to *D*, and *e* as the alternative to *E*, we recognize that each of what we have hitherto regarded as allelomorphs in reality contains three elementary antigens, capable of reacting, as R_2 and R_z certainly do, with three different antibodies.

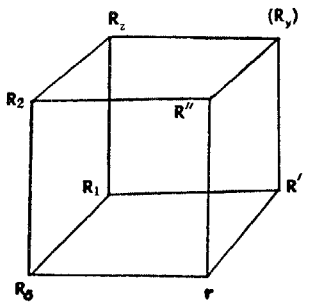
The theoretical reactions ascribed to the undiscovered gene or gene-complex R_y thus merely complete the system in which each reagent reacts positively with four and negatively with the other four complexes. Further, it appeared probable at this stage that the allelomorphs *d* and *e* were also capable of acting as antigens, and producing specific antibodies δ or anti-*d* and η or anti-*e*, having reactions antithetical to those of Δ and H, as shown in the extension on the right of Table IV. This gives three positive reactions to every allele. At least one writer felt this proposal to be so conjectural that he denied stoutly the possibility of these hypothetical antibodies, yet within the year (1944) in which Race [6] first published this proposal, Mourant [7] had discovered anti-*e* in a transfusion reaction, and very promptly recognised its conformity with the reactions predicted. Furthermore, before the end of the following year Diamond of the Harvard Medical School had confirmed the occurrence of anti-*d* in a case of haemolytic pregnancy reaction, following earlier experience of transfusion.

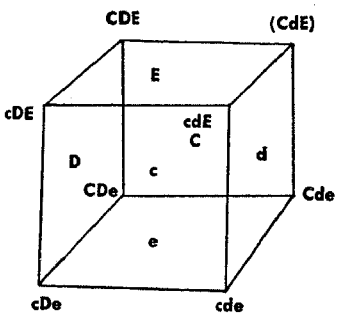
The discovery of these two reagents completing the set of six predicted was much more important than the finding of an exceptional variety such as R_y could possibly be, for these constitute new tools of research which at once enable distinctions to be made among genotypes previously indistinguishable. Indeed, it seems fair to say that Diamond's δ is the only reagent which can clear up the obscurity of the Rhesus situation in non-European peoples, for in all these *d* seems to be rare, and using Δ only it could only be recognised in homozygotes so rare that European ancestry cannot easily be excluded. With the aid of δ, however, every *d* present could be detected.

Although the advancement of knowledge of the Rhesus factor has not ceased at the point to which my narrative has reached, it is a convenient point at which to pause and take stock, for what at first seemed a bewildering complexity has now reduced itself to very simple terms. There are three distinct pairs of elementary antigens, each inherited like Landsteiner and Levine's *M* and *N*, but closely, or perhaps absolutely, linked in inheritance. Every gene complex of Table IV now reacts with three and only three different antibodies. For three such factors without linkage there would be 27 genotypes, all distinguishable by means of the six reagents available. For three such factors linked the number of genotypes is 36, and as we still have only 27 possible phenotypes, some pairs or sets of genotypes must be indistinguishable by individual tests unaided by family evidence.

We may represent the eight heritable antigen complexes geometrically as the corners of a cube, while the six elementary antigens are represented by the faces; each allelomorphic pair of antigens is then a pair

of opposite faces, and the three faces meeting in any point specify the antigens in each complex:





The genotype of each individual is determined by the pair of complexes supplied by the two parents. If these are wholly alike the person in question is triply homozygous; alternatively, he may be singly, doubly, or triply heterozygous. The possible genotypes may therefore be classified as follows:

	Number of antibodies with which +ve	*Number of antibodies which can be generated*	*Examples*
8 triply homozygous	3	3	R_1R_1 19%, rr 15%, R_2R_2 2%, all others rare.
12 singly heterozygous	4	2	R_1R_0 3%, R_0r 2.4%, all rather rare.
12 doubly heterozygous	5	1	Three common.
4 triply heterozygous	6	0	

The single heterozygotes correspond with the 12 edges of the cube, and like the first 8 are all recognisable by serological tests only.

Each face of the cube has two diagonals giving 6 pairs of genotypes all doubly heterozygous. Members of the same pair contain the same 5 antigens, and give the same reactions. There are therefore 12 genotypes and only 6 phenotypes. Three of these phenotypes are common, each having one common and one rare genotype:

R_1r	33%	R_1R_2	11%	R_2r	10%
R_0R'	.06%	R_0R_z	.006%	R_0R''	.08%

Individuals with these reactions will be classified as belonging to the more numerous genotypes with some confidence, but without family evidence about one in 300 will be misclassified. In the other three pairs both genotypes are exceedingly rare; estimated frequencies for these, taking 48 genes per million to be R_y, are shown below:

$R'R''$	277 per million	$R''R_z$	31 per million	$R'R_z$	21 per million
rR_y	37 per million	R_2R_y	12 per million	R_1R_y	42 per million

If the last class is identified, it is not improbable that family evidence will show it to contain R_y. Obviously, however, this complex R_y is not only rare, but also very well concealed.

Finally, a group of four genotypes should react with all six antibodies, these are:

R_1R'' 0.8% R_2R' 0.3% R_zr 0.1% R_0R_y 0.0003%

These constitute a single phenotype, with four genotypes all triply heterozygous.

The estimate adopted above is based on the possibility of occasional recombination of elementary antigens among the gametes from doubly and triply heterozygous genotypes. By crossing over, the common double heterozygotes could produce the four rarer complexes R_0, R', R'', R_z, but not R_y, as follows:

		3%	1.0%	1.2%	0.1%	
R_1r	33%	R_0	R'	—	—	by crossover c/d
R_1R_2	11%	R_0	—	—	R_z	by crossover c/e
R_2r	10%	R_0	—	R''	—	by crossover d/e

Since R_0 is in fact about as frequent in our population as R', R'' and R_z together, it seems possible that these frequencies are maintained by occasional crossing over against counterselection. The highest ratio appears to be that of R'' against R_2r, so that the order of arrangement in the chromosome is more likely to be

than either of the other two orders. In this case R_y could be produced only by the triple heterozygotes R_2R' and R_zr and their known frequencies supply a first rough estimate of the frequency of R_y to be expected.

Subsequent to the verification of the threefold nature of the Rhesus gene complexes the latest advances have shown that more than two alternative antigens may occur at some of the three loci. Race, Mourant and Callender [81], have recently established a third antigen C^w, and demonstrated the existence of all six possible combinations:

CC, Cc, cc, CC^w, cC^w, C^wC^w.

The frequencies in England are about:

c 55% C 44% C^w 1%.

The new elementary antigen has been found in the complexes

C^wDe and C^wde;

since many anti-C sera also contain anti-C^w, these would hitherto have been included with R_1 and R' respectively. The two other theoretical complexes

C^wDE and C^wdE

would have been included with R_z and R_y, and must therefore be exceedingly rare. I understand also that Dr. Stratton has found a third rare allelomorph at the *D* locus, but I have so far no particulars of this extension.

It appears, therefore, that the number of gene complexes which will need to be distinguished when types, rare in England, have to be diagnosed, is not 8 but perhaps 18 or more, and that the number of genotypes is not 36 but 171. Using reliable and unmixed testing fluids, each capable of identifying a single elementary antigen, this situation, though complex, is by no means unmanageable. It does show, however, that we must not hastily extend conclusions from the comparatively well studied populations of European origin to non-European races, but must instead study these with the full battery of tests already available.

REFERENCES

1. Landsteiner, K. *Zentralb. f. Bakteriol., 27,* 357 (1900).
2. Landsteiner, K., and Levine, P. *Jour. Exp. Med., 47,* 757 (1928).
3. Levine, P., and Stetson, R. E. *J. Amer. Med. Assn., 113,* 126–7 (1939).
4. Landsteiner, K., and Wiener, A. S. *Proc. Soc. Exp. Biol. and Med., 43,* 223 (1940).
5. Taylor, G. L., and Mollison, P. L. *Brit. Med. J., 1,* 561 (1942).
6. Race, R. R. *Nature, 153,* 771 (1944).
7. Mourant, A. E. *Nature, 155,* 542 (1945).
8. Race, R. R., Mourant, A. E., and Callender, S. *Nature, 157,* 410–11 (1946).