PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Phenotypic Genetics of Complement Components

M. J. Hobart

Phil. Trans. R. Soc. Lond. B 1984 **306**, 325-331
doi: 10.1098/rstb.1984.0093

Phil. Trans. R. Soc. Lond. B **306**, 325–331 (1984)
Printed in Great Britain

Phenotypic genetics of complement components

By M. J. Hobart

Mechanisms in Tumour Immunity Unit, M.R.C. Centre, Hills Road, Cambridge CB2 2QH, U.K.

Both charge and size-dependent electrophoretic techniques have been used to investigate genetic polymorphisms of complement proteins. Of the seventeen complement proteins, ten have been shown to have genetic variants and only one (C9) has been extensively investigated without revealing variants.

These investigations give information on the numbers of cistrons and their linkage relations. They demonstrate or confirm the linkage of C2, Factor B and C4 to the MHC. In the cases of both human and mouse C4, it has been shown that the loci are (usually) duplicated. C4 in humans is extremely polymorphic and exhibits a number of strong allelically associated haplotypes. Some of these have only one expressed C4 gene and are associated with disease susceptibility.

C8 has at least two cistrons coding for associating subunits. C6 and C7 are linked in several species and sometimes C7 is duplicated. This gene pair is discussed in relation to natural selection and gene conversion.

Information on the phenotypic genetics of complement components comes principally from two kinds of variation: deficiency, which is the subject of a later paper, and inherited structural variation. The latter is most often recognized by differences in electrophoretic mobility, usually in a charge-dependent separation system, or by antigenic differences.

Until the advent of gene cloning methods, phenotypic differences were the only tool by which the genes of the complement system could be investigated. This paper is concerned with reviewing the information that we have gleaned by these investigations and suggesting a role for future phenotypic approaches which will be complementary to advances by gene cloning.

Gene variation, revealed principally by structural variations, has been used to show the linkage relations of complement components and to assign them to chromosomes. It has been used to suggest how many loci are involved in coding for a protein. It might have been used, in the cases of C3 and C4, to suggest that these molecules are coded by single cistrons, although they are composed of more than one polypeptide chain. All these roles are now much better undertaken by investigation at the DNA level, with the exception of describing linkage relations at distances greater than a hundred or so kilobases. In addition, DNA investigations allow the acquisition of a wealth of data on the sequence, structure and control of the genes. There remain, however, areas where phenotypic methods are still quicker and more informative. Investigations of variation within the population are still much more readily done at the phenotypic level than at the DNA level. This is partly because there is very little labour in the preparation of the samples and comparatively little in their investigation, once a method has been established.

Investigation of inherited structural variation of the complement components began with the discovery of C3 polymorphism by three groups (Wieme & Demeulenaere 1967; Alper & Propp 1968; Azen & Smithies 1968). C3 is a sufficiently abundant component of the serum that it can be directly seen on stained electrophoretic separations. Since the other complement

components are substantially less abundant, their positions in electrophoretic separations must be discovered by the use of more sensitive and specific methods.

The first of these was the use of immunofixation, in which precipitating antisera are applied to the surface of the gel in the region where the complement component has migrated (Alper *et al.* 1972). When the antigen–antibody precipitate has formed, the gel is washed to remove the soluble proteins and the precipitate, which contains substantially more protein than the antigen alone, is stained.

The immunofixation method can be applied conveniently to separation methods in agarose gels, and these now include isoelectric focusing, which offers very high resolution of charge differences. However, this method of separation is still more conveniently done in polyacrylamide gels, which are inconvenient for immunoprecipitation detection methods. Two methods of component detection which work well with such gels are functional overlays (Hobart *et al.* 1975) and protein blotting, the latter being combined with detection with antibodies labelled either with radio-iodine or peroxidase (Whitehouse & Putt 1983).

To make full use of the potential of protein blotting methods, it would be convenient if the blot could be re-used. It is possible to elute the antibodies from a blot under mildly denaturing conditions, but it would be preferable to have a method for chemically tagging the antibody which can be reversed at a later date, with release of the label from the blot. The biotin–avidin labelling system is a good candidate and a cleavable biotinylating agent has been reported (Mouton *et al.* 1982). We have developed a different approach to cleavable biotinylation, involving treatment of the antibody with SPDP (Carlsson *et al.* 1978) and subsequent substitution with *N*-biotinyl penicillamine to yield a disulphide-linked biotin derivative. The blot is probed successively with the biotinylated antibody and ^{125}I-labelled streptavidin. After radioautography, the blot is reduced and the biotin–streptavidin–label complex is released. After alkylation of free –SH groups, the blot may now be re-probed with a similarly tagged antibody of different specificity. This method promises to provide a higher throughput of results than previous methods, with the additional advantages of economy of sample and simplified storage of family material as pre-run blots which can be probed with new antibodies as they become available. If the sensitivity of the system can be sufficiently raised, it offers the possibility of systematic investigation of many cell surface proteins such as complement receptors.

Molecular mass variation of the C3b (CR1) receptor has been described by Wong *et al.* (1983), who used immunoprecipitation of solubilized cell membranes. These methods do not readily permit subsequent investigations of the material collected, which is not abundant and it is here that the advantage of re-use techniques would be especially felt.

The use of phenotyping methods (and information from studies of deficient humans and laboratory animals) has permitted the localization of three of the complement components to the major histocompatibility complex, and the exclusion of many others from this region of the genome. The assignment of mouse C4 to the H-2 region and the subsequent discovery that the locus is duplicated was a notable landmark and is the subject of an accompanying paper by D. Schreffler.

The components of the complement system are surprisingly rich in variants. Only one of the complement components that has been extensively investigated has so far failed to reveal polymorphic differences: C9 (table 1). We have recently begun to extend this investigation using the cleavable biotin reagent, with similar results (M. J. Hobart, J. Tamerius & P. J. Lachmann, unpublished results). C5 shows polymorphism so far only among people of Melanesian

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

extraction (Hobart *et al.* 1981). Antibody-based detection techniques have been used to investigate Factor H (B. Larsen, personal communication and author's unpublished observations) and Factor I (unpublished observations). These very preliminary investigations have not revealed polymorphism.

In contrast, all the other components that have been investigated show polymorphism in at least one major racial group (table 2). In most cases, there is now linkage information for these genes, and only the linkage group C6–C7 is unassigned to a chromosome.

Table 1. Individuals tested for C9 variants

population	number
Caucasian (E Anglia)	714
Negroid (Gambia)	45
Melanesian (Papua New Guinea and Solomon)	121
Australian aborigine	50
Mongoloid (Burma)	23
total	953

No variants observed.

Table 2. Polymorphic complement components

component	gene frequency	linked to	chromosome	reference
C4	(2 loci)	HLA	6	O'Neill *et al.* (1978)
C2	0.05	HLA	6	Hobart & Lachmann (1976) Alper *et al.* (1976) Meo *et al.* (1976)
Factor B	0.28	HLA	6	Allen (1974)
C3	0.22	Le	19	
C8 α–γ β	0.4	PGM1	1	Mevag *et al.* (1984) Alper *et al.* (1983)
C6	0.33	C7	?	Hobart *et al.* (1975)
C7	0.01	C6	?	Hobart *et al.* (1978)
C5	0.05 (Mel)	—	—	Hobart *et al.* (1981)
Factor D	0.05 (Neg)		—	Hobart & Lachmann (1976)
C4-binding protein	0.02		—	Rodrigues de Cordoba *et al.* (1983)
CR1	0.1		—	Wong *et al.* (1983)

There is some conflict of evidence on the linkage relations of C8. This protein is composed of three polypeptide chains which form two freely associating subunits (Lachmann, this symposium). Polymorphism of the protein has been investigated both by the haemolytic overlay technique and by protein blotting, using isoelectric focusing separations in each case. The clarity of the bands seems better by blotting, though this is subject to uncertainties about the detailed specificity of the antibody which makes for some difficulty in interpreting the patterns. It is agreed that there is a polymorphism with gene frequencies of about 0.6–0.4 (Raum *et al.* 1979; Mevåg *et al.* 1984) and that this is coded at the α–γ locus. There is much less common polymorphism of the β locus (Alper *et al.* 1983) and it may be that this is not well revealed by protein blotting. Alper *et al.* (1983) report a convincing lack of linkage between the C8 loci and both Pericak-Vance *et al.* (1982) and Mevåg *et al.* (1984) find linkage of C8 with PGM1

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

on chromosome 1. It is here that there is a triangular conflict of evidence: the strong linkage of Mevåg *et al.* (1983) refers to the α–γ subunit, whereas the much weaker data of Pericak-Vance *et al.* (1982) refers to the beta subunit, which seems not to be linked to α–γ. The earlier report on C8 α–γ polymorphism gives negative linkage scores for the chromosome 1 markers Rh and Duffy (Raum *et al.* 1979).

The components that are linked to the major histocompatibility complex have attracted most attention, and the results of gene cloning experiments are presented in accompanying papers. The components that are linked to the MHC are C4 (Rittner *et al.* 1975), C2 (Fu *et al.* 1975) and Factor B (Allen 1974). Both C2 and Factor B appear to be coded by single loci, whereas C4 is coded by at least two loci in most haplotypes. To review in detail the current state of knowledge on C4 polymorphism is a major undertaking, but the following outline will suffice as an introduction to the gene cloning results.

(i) *C4* gene products occur in two families of variants, the products of the *C4A* locus usually having a faster mobility in alkaline electrophoresis and bearing antigens of the Rodgers blood group series. The haemolytic activity of these products is not seen under certain test conditions and their α chains are marginally larger than those of the *C4B* locus as assessed by SDS-polyacrylamide gel electrophoresis. The products of the *C4B* locus usually have a slower mobility and bear antigens of the Chido blood group series.

(ii) Haplotypes that fail to express one or other of the loci are quite common. It is not yet known whether the genes are absent or defective in these loci, or whether the loci duplicated identically and it may well prove that both cases occur. Rare haplotypes are found in which neither locus is expressed or in which a third product is coded.

(iii) Some alleles produce a product which is detected by immunoprecipitation but which has no haemolytic activity.

(iv) Within the loci there is extensive polymorphism with respect to electrophoretic mobility and some polymorphism of antigenic determinants, which present as partial inhibitors in Chido and Rodgers typing. The mobility of the products of the two loci overlap and the assignment of a particular product to a locus is based on evidence of its antigenic nature, haemolytic activity and molecular mass of the α chain. In many cases, the presence of other gene products in an individual prevent accurate assignment of a gene product to a locus, and family studies are essential.

(v) Although the overwhelming majority of genes can be assigned to one or the other locus, there are cases where molecules have features of both loci (Rittner *et al.* 1983). Gene conversion offers a plausible explanation of this.

C6 and C7 have been shown to be closely linked in a number of species: man (Hobart *et al.* 1978); dogs (Eldridge *et al.* 1983); marmosets (Whitehouse 1983); and mice (A. Orren, personal communication). There is reason to believe that they are the products of tandemly duplicated genes. Because they depend on each other for their principal biological action (the generation of the lytic complex), many of the evolutionary pressures that affect one may well be reflected in the other. However, such pressures cannot be extreme, since deficiency of either or both is not fatal, nor even inimical to good general health (see Lachmann, this symposium). Even pressures acting on nearby genes which might have the effect of accidentally selecting particular alleles would have very similar effects on both, since they are close together in the genome.

Human populations show a remarkable similarity in C6 allele frequencies, which are high

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

enough to make nearly half the human race heterozygous at this locus. One interpretation of this is that heterozygosity at the locus confers selective advantage, a conclusion which might be supported by observations in the Manx shearwater, where it is the only polymorphic locus among about 30 tested, and with very similar gene frequencies (Whitehouse 1982). However, if this is so, why should it not be equally so for C7, where only about 1% of Caucasoids are heterozygous? This gene pair represents a potentially good test for evidence of selection.

In an investigation of 33 species, P. R. Eldridge *et al.* (unpublished results) sought to determine whether the loci are polymorphic. For some species, the numbers of samples were small, so the results are biased in favour of lack of polymorphism (table 3). It can be seen that there is no reliable relation between the existence of polymorphism at one locus or another, and the simplest explanation of this result is that, in this case, alleles arise and are fixed by chance rather than by selection based on their qualities.

TABLE 3. C6 AND C7 VARIATION IN 33 MAMMALS

		C6 polymorphism		
		+	−	assay failed
C7 polymorphism	+	4	5	0
	−	8	8	1
	assay failed	4	3	

A further interesting observation is that the *C7* locus is duplicated in dogs and Przwalski horses (Eldridge *et al.* 1983; D. B. Whitehouse and W. Putt, personal communication). In these cases there is extreme polymorphism of each of the loci, a feature which is shared by C4 in humans. Clearly, where loci are invariant, the possibility of duplication cannot be excluded as there are no differences in products to discern and the detection of these duplicates will demand DNA cloning. In this sense, the above observations are subject to ascertainment error. Nevertheless, it is notable that the duplication of loci is associated with florid polymorphism, exceeding that of the other components. One possible explanation for this fact is that the polymorphism is produced by mutations occurring in the course of gene conversion events (as has been shown for mouse immunoglobulin genes by Ollo & Rugeon 1983). This is an attractive hypothesis, since it is known that the *C7* locus is single and polymorphic in many species, especially primates (D. B. Whitehouse, personal communication). This presumably means that, when speciation occurs, *C7* is more commonly a single locus than a duplicated locus. Nevertheless, subsequent duplication is associated with greater polymorphism. This implies that the mutation rate in the duplicated loci is greater than single loci, and that an additional mechanism is operating. It will be of interest to investigate these polymorphic loci at the DNA level to test the hypothesis.

As it becomes more common to investigate genes at the DNA level, it will be appropriate for the phenotypic geneticist to concentrate on experiments where he is able to make a complementary rather than a competitive contribution. Clearly, investigation of the gene at the finest level and of its neighbours to a range of a few hundreds of thousands of bases can best be done with DNA techniques, as can chromosomal and regional assignments. The investigation of linkage relations at the ranges 0.5–10 centimorgans (cM) is still probably best achieved by phenotypic methods, especially where markers are as abundant as they are for most

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

complement components and where new techniques offer higher productivity than hitherto. In the near future, spot hybridization techniques using synthetic oligonucleotide probes will be in wide use for the investigation of specific sequence differences between alleles. However, in cases where gene conversion has been at work, these techniques may produce garbled results, owing to the presence of the same polymorphic sequence in more than one gene. In these cases, the phenotypic approach has an advantage, because, rather than investigating a specific nucleotide sequence or its context (restriction fragment length), phenotypic methods look at the integrated product of the whole gene. In investigations of rare *C4* alleles, this process has identified probable gene conversions leading to the transfer of antigenic determinants associated with the products of one *C4* locus into alleles at the other locus (Rittner *et al.* 1983). It is also tempting to believe that for some time to come, gene cloning will not be attempted simply 'because it is there', but because there are valid reasons for interest in a gene, and that phenotypic investigations will still provide some of the reasons.

References

Allen, F. H. 1974 Linkages of HL-A and GBG. *Vox Sanguinis* **27**, 382.

Alper, C. A., Boenish, T. & Watson, L. 1972 Genetic polymorphism in human glycine-rich beta-glycoprotein. *J. exp. Med.* **135**, 68.

Alper, C. A., Marcus, D., Raum, D., Petersen, B. H., Spira, T. J. 1983 Genetic polymorphism in C8 beta-chains: evidence for two unlinked genetic loci for the eighth component of human complement (C8). *J. clin. Invest.* **72**, 1526–1531.

Alper, C. A. & Propp, R. P. 1968 Genetic polymorphism of the third component of human complement (C'3). *J. clin. Invest.* **47**, 2181.

Azen, E. A. & Smithies, O. 1968 Genetic polymorphism of C'3 (beta 1c-globulin) in human serum. *Science, Wash.* **162**, 905.

Eldridge, P. R., Hobart, M. J. & Lachmann, P. J. 1983 The genetics of the sixth and seventh components of complement in the dog: polymorphism, linkage, locus duplication and silent alleles. *Biochem. Genet.* **21**, 81–91.

Fu, S. M., Kunkel, H. G., Brusman, H. P., Allen, F. H. & Fotino, M. 1974 Evidence for linkage between HL-A histocompatibility genes and those involved in the synthesis of the second component of complement. *J. exp. Med.* **140**, 1108.

Hobart, M. J., Joysey, V. & Lachmann, P. J. 1978 Inherited structural variation and linkage relationships of C7. *J. Immunogenet.* **5**, 157–163.

Hobart, M. J., Lachmann, P. J. & Alper, C. A. 1975 Polymorphism of human C6. In *22nd protides of the biological fluids* (ed. H. Peeters), p. 575. Oxford: Pergamon Press.

Hobart, M. J., Vaz Guedes, M. A. & Lachmann, P. J. 1981 Polymorphism of human C5. *Ann. Hum. Genet.* **45**, 1–4.

Mevåg, B., Rogde, S. Olaisen, B., Gedde-Dahl, T., Teisberg, P. & Tedesco, F. 1984 Genetic polymorphism of complement component C8. Proceedings, Xth Congress of Forensic Haemogenetics, Munich 1983.

Mouton, C. A., Pang, D., Natraj, C. V. & Shafer, J. A. 1982 A reagent for covalently attaching biotin to proteins via a cleavable connection arm. *Archs Biochem. Biophys.* **218**, 101–108.

Ollo, R. & Rougeon, F. 1983 Gene conversion and polymorphism: generation of mouse immunoglobulin 2a chain alleles by differential gene conversion by 2b chain genes. *Cell* **32**, 515–523.

Pericak-Vance, M. A., Elston, R. C., Spira, T. J. & Band, J. 1982 Segregation and linkage analysis of immunochemical C8 levels in a family with C8 beta-chain deficiency. *Am. J. Hum. Gen.* **34**, 190.

Raum, D., Spence, A., Balavitch, D., Tideman, S., Merritt, D., Taggart, R. T., Petersen, B. H., Day, N. K. & Alper, C. A. 1979 Genetic control of the eighth component of complement. *J. clin. Invest.* **64**, 858–865.

Rittner, C., Hauptmann, G., Grosse-Wilde, H., Grosshans, E., Tongio, M. M. & Mayer, S. 1975 Linkage between HL-A (major histocompatibility complex) and genes controlling the synthesis of the fourth component of complement. In *Histocompatibility testing* (1975 edn). Copenhagen: Kissmeyer–Nielsen F. Munksgaard.

Rittner, C. H., Roos, M. H., Mollenhauer, E. & Kolvenback, H. 1983 Detection and segregation of rare duplicated human C4 haplotypes. *Immunobiology* **164**, 288 (abstract).

Rodriguez de Cordoba, S., Ferreira, A., Nussenzweig, V. & Rubenstein, P. 1983 Genetic polymorphism of human C4 binding protein. *J. Immunol.* **131**, 1565–1569.

Whitehouse, D. B. 1982 Genetic variation of the sixth component of complement (C6) in the Manx shearwater (*Puffinus p. puffinus*). *Anim. Bld Groups Biochem. Genet.* **13**, 179–187.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Whitehouse, D. B. 1983 Genetic linkage of C6 and C7 in the common marmoset. *Cytogenet. cell. Genet.* **37**, 606 (abstract).

Whitehouse, D. B. & Putt, W. 1983 Immunological detection of the sixth component of complement (C6) following flat bed polyacrylamide gel isoelectric focusing and electrophoretic transfer to nitrocellulose filters. *Ann. Hum. Genet.* **47**, 1–8.

Wieme, R. J. & Demeulenaere, L. 1967 Genetically determined electrophoretic variant of the human complement component C'3. *Nature, Lond.* **214**, 1042.

Wong, W. W., Wilson, J. G. & Fearon, D. T. 1983 Genetic regulation of a structural polymorphism of human C3b receptor. *J. clin. Invest.* **72**, 685–693.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES