

HUMAN GENES FOR THREE COMPLEMENT COMPONENTS THAT REGULATE THE ACTIVATION OF C3 ARE TIGHTLY LINKED

BY SANTIAGO RODRIGUEZ DE CORDOBA, DOUGLAS M. LUBLIN,*
PABLO RUBINSTEIN, AND JOHN P. ATKINSON*

*From the Department of Immunogenetics, The Lindsley F. Kimball Research Institute of the New York Blood Center, New York 10021; and the *Howard Hughes Medical Institute and Department of Medicine, Division of Rheumatology, Washington University School of Medicine, St. Louis, Missouri 63110*

Two groups of complement components are involved in the activation of C3. The first includes C2, factor B (B), C4, and C3, i.e., the structural components of the C3 convertases of the classical and alternative pathways. The second contains the regulatory proteins C4-binding protein (C4-bp), factor H (H), the C3b/C4b receptor (CR1 or C3bR), and, probably, the decay-accelerating factor (DAF) and gp45-70.

C2, C4, and factor B are encoded by a cluster of genes inside the major histocompatibility complex (MHC),¹ a genetic association that is consistently found in mammalian species (reviewed in 1 and 2). The mechanisms leading to that linkage and its evolutionary stability are unclear, as are its functional and genetic consequences. Study of the linkage relationships of other genes that code for proteins involved in the activation of C3 is, therefore, of considerable interest.

We have recently demonstrated genetics variants of C4-bp (3), H (4), and C3bR (5, 6) in humans and have reported that the loci coding for C4-bp and C3bR are very closely linked (7). We proposed (7) the hypothesis that a cluster of genes encoded the regulatory complement components of the activation of C3. We now present evidence indicating that the locus coding for factor H in humans is also linked to the *C4BP*, *C3bR* cluster. In addition, we show that this cluster of regulatory proteins segregates independently of the *C2*, *Bf*, *C4A*, *C4B* gene cluster and of the human *C3* locus.

Materials and Methods

Materials. ϵ -Aminocaproic acid, Nonidet P-40 (NP-40), β -mercaptoethanol, phenylmethylsulfonyl fluoride, Coomassie Brilliant Blue R, and neuraminidase type V (*Clostridium perfringens*) were from Sigma Chemical Co., St. Louis, MO; EDTA from Fisher Scientific Co., Pittsburgh, PA; ampholine (pH 3.5-10, 4-6, and 6-8) from LKB Produkter, Bromma, Sweden; *N,N',N'*-methylenebis-acrylamide, acrylamide, *N,N,N',N'*-tetramethylethylenediamide, ammonium persulfate, sodium dodecyl sulfate (SDS) from Bio-

This work was supported in part by grant HL-09011 from the National Institutes of Health, and by a grant from the Monsanto Corporation and the Lottie Caroline Hardy Trust Fund.

¹ *Abbreviations used in this paper:* MHC, major histocompatibility complex; NP-40, Nonidet P-40; RCA, regulators of complement activation; SDS, sodium dodecyl sulfate; TCA, T cell A locus.

Rad Laboratories, Richmond, CA; ultrapure urea from Schwarz-Mann Biotech, Cambridge, MA.

Blood Samples. 230 families were included in these studies. The familial relationships were corroborated in all cases by several genetic polymorphisms. Some of the families were made available by Dr. F. Ginsberg-Fellner, Department of Pediatrics, Mount Sinai School of Medicine, New York and by Dr. G. Ercilla, Departamento de Inmunologia, Hospital Clinico y Provincial, Barcelona, Spain.

Blood samples were obtained by venipuncture, and lymphocytes, EDTA-plasma, and serum were separated by standard techniques. Red blood cells were also obtained. Tests were done on fresh specimens whenever possible. Otherwise, samples were frozen and kept either in the gas phase of liquid nitrogen (lymphocytes and erythrocytes) or at -80°C (serum and plasma) until used.

Allotyping of C4-bp and H. *C4BP* and *FH* phenotyping were performed as previously described (8) using EDTA-plasma or EDTA-serum. Briefly, samples were treated with neuraminidase, immunoprecipitated with a monospecific antibody, analyzed by isoelectric focusing under denaturing conditions using vertical 4.5% polyacrylamide gels, and stained with Coomassie Brilliant Blue.

Allotyping of C3bR on Erythrocytes. *C3bR* phenotyping was performed as previously described (5) using blood samples collected in standard ACD (acid-citrate-dextrose, formula A) anticoagulant. Briefly, erythrocytes were surface labeled with ^{125}I and solubilized with 1% NP-40. Samples were then immunoprecipitated with a monoclonal antibody against C3bR and analyzed by slab gel SDS-polyacrylamide gel electrophoresis according to the method of Laemmli using a 5% gel and autoradiography.

Allotyping of HLA Antigens. Typing of HLA-A, -B, and -C antigens was done by the standard National Institutes of Health cytotoxicity technique modified so that the reading was by the contrast fluorescence technique developed in this laboratory (9). Typing of HLA-DR antigens was done by the two-color fluorescence test described by van Rood et al. (10). All currently recognized HLA-A, -B, -C, and -DR antigens were identifiable with the sera ($n = 204$) in our routine battery.

Allotyping of C3. The procedure for testing this polymorphic protein was a modification of a previously described test (11).

Linkage Analysis. Segregation data for the different variants fit Mendelian expectations. Analysis of linkage was done with the Lod score method of Morton using the NYLIP computer program developed in this laboratory (12). Confirmation of these findings was done with the newly described method of multilocus analysis of linkage (13) using the LINKAGE computer program (14) kindly provided by Dr. J. M. Lalouel.

Results and Discussion

This report demonstrates the existence in humans of a cluster of genes for complement proteins that constitute the regulatory counterparts of the MHC-linked cluster of C3 convertase components genes. We have previously demonstrated that the *C4BP* and *C3bR* loci are linked in humans (7): no recombinants have been encountered between *C4BP* and *C3bR* in informative heterozygotes. We also reported that the less common allele *C4BP*2* is only found in haplotypes that also contain the less frequent allele *C3bR*B*, as shown by segregation studies in three pedigrees. Recently, a further individual with the genotype *C3bR*B,B; C4BP*1,2* was identified in a Dutch family. The complete inclusion of the *C4BP*2* allele in the *C3bR*B* haplotype demonstrates the presence of very strong linkage disequilibrium. This fact, in addition to the linkage data, indicates that the loci coding for C4-bp and C3bR must be closely linked. The more recently found polymorphism of factor H could, therefore, be analyzed for linkage to the haplotypes defined by *C4BP* and *C3bR* alleles, thus increasing the number of informative families.

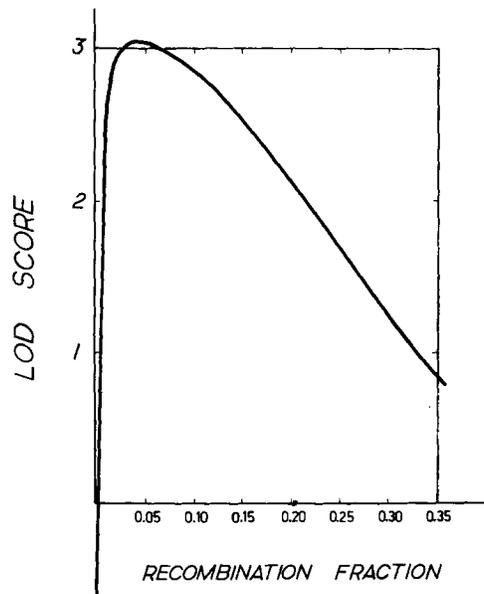


FIGURE 1. Linkage analysis between *FH* locus and the *C3bR*, *C4BP* linkage group. Maximum likelihood estimate of recombination fraction, θ , is 0.043, with corresponding maximum Lod score of 3.035.

Seven two-generation families, with a total of 39 members, each contained one parent who was heterozygous both at the *FH* locus and at the *C3bR-C4BP* region (i.e., *C3bR*- and/or *C4BP*-heterozygous). There were 23 informative meioses in these families with only one recombinant. The values for the Lod score (logarithm of odds in favor of linkage) were calculated for different values of the recombinant fraction, θ , using the NYLIP program (Fig. 1). The maximum Lod score value was estimated to be 3.035 (i.e., the odds on linkage are 1,084:1) at a θ value of 0.043 (i.e., 4.3 centimorgans [cM]). The only recombinant between *FH* and the *C3bR-C4BP* region belongs to a large pedigree (family PED) in which the father is a triple heterozygote *C3bR***A,B*; *C4BP***1,2*; *FH***1,2* (Fig. 2). This crossing-over event defines the position of the *FH* gene as separable from the other two genes in the cluster. The estimated distance (4.3 cM) is, of course, only approximate because of the scarcity of data.

These results have been confirmed by multilocus linkage analysis, developed by Meyers et al. (13), which also allowed us to derive information from an additional family with five offspring that was informative for the segregation of the *C4BP* and *C3bR* loci. According to this approach, likelihood ratios from pedigrees with a double-heterozygote parent (and hence informative for only one of the three possible pairs of loci) were combined with data from families with a triple-heterozygote parent. Maximum likelihood estimates were made for the recombination fractions under each of the three possible gene orders. (Parameters included the coefficient of coincidence c , a measure of interference in crossing-over; the results gave no evidence for interference, with maximum likelihood occurring at the boundary value, $c = 1$.) The most likely gene order was either *C4BP-C3bR-FH* or *C3bR-C4BP-FH*, with $\theta_{C4BP,C3bR} = 0.000$ and $\theta_{C4BP,FH}$

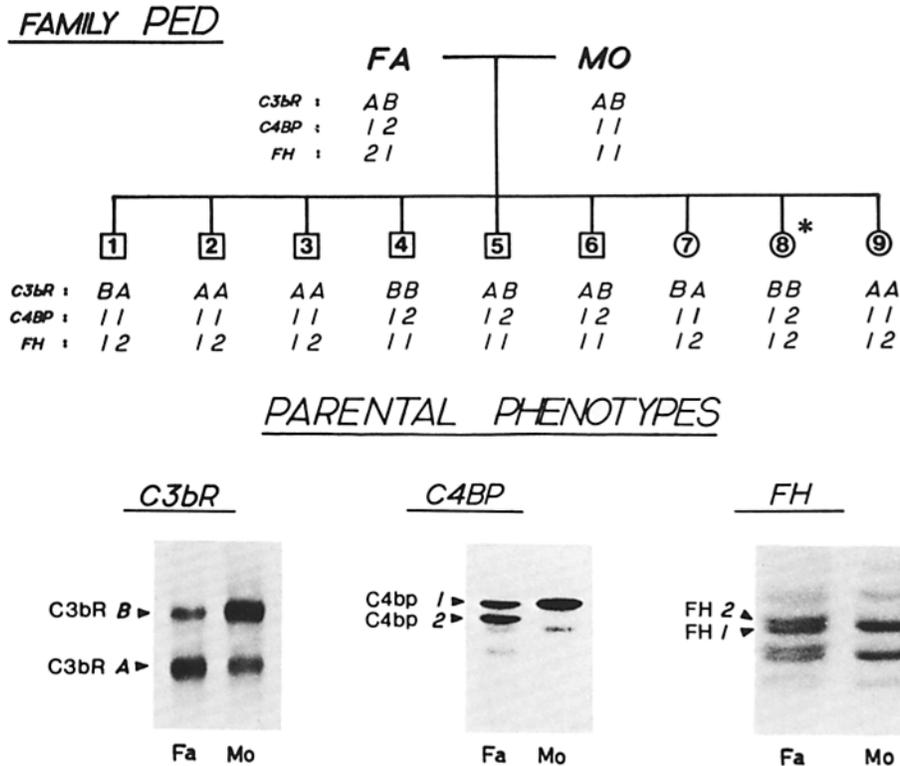


FIGURE 2. Pedigree of family PED showing *C3bR*, *C4BP*, and *FH* genotypes. Assigned haplotypes are A-1-1/B-1-1 for the mother and A-1-2/B-2-1 for the father. Note that sibling 8 (*) represents a cross-over between *C3bR*, *C4BP* on one side and *FH* on the other. Phenotypes for *C3bR*, *C4BP*, and *FH* were determined as described in Materials and Methods. Parental phenotypes are shown at the bottom of the figure for paternal (Fa) and maternal (Mo) samples.

= 0.044, as before, with a likelihood ratio in favor of linkage of 2.77×10^5 to 1. In accord with large-sample theory (15), $2 \log_e$ (likelihood ratio) is a χ^2 variate with two degrees of freedom, giving $P < 10^{-5}$ for this triple-linkage group. These findings, therefore, define a new linkage group in man, consisting of the regulatory proteins *C3bR*, *C4-bp*, and *H* of the complement system. They also provide information in relation to the ordering of the three loci.

We have previously reported that neither *C4BP* (3) nor *C3bR* (16) are closely linked to the *HLA*-linked gene cluster encoding the complement components *C2*, *B*, *C4A*, and *C4B*. These results are now updated and summarized in Table I, which indicates that the *C4BP*, *C3bR*, *FH* gene cluster segregates independently of the *C2*, *Bf*, *C4A*, *C4B* gene cluster, the *TCA* locus² and the *C3* locus.

The three proteins studied here, *C3bR*, *C4-bp*, and *H*, regulate the activities of the classical and/or alternative pathway *C3* convertases. They all cause decay acceleration of these enzymes and serve as cofactors for factor I-mediated cleavage of *C4b* and/or *C3b*. As members of a linkage group, they form

² *TCA*, a locus ~10 cM telomeric to *HLA-A*, codes for class I antigens in the surface of T lymphocytes and is probably homologous to the *Qa* locus in the mouse (17).

TABLE I
Lod Score Values for Linkage Between HLA, C3, or TCA Loci and RCA*

Loci	Number of informative families	Nonrecombinant/ recombinant [†]	Lod score at $\theta =$						
			0.0	0.05	0.10	0.15	0.20	0.25	0.30
HLA:RCA	42	103:42	$-\infty$	-23.07	-12.74	-6.87	-3.77	-1.91	-0.90
C3:RCA	11	23:10	$-\infty$	-5.63	-3.07	-1.79	-1.03	-0.57	-0.29
TCA:RCA	6	14:6	$-\infty$	-3.51	-1.94	-1.13	-0.64	-0.34	-0.15

* HLA, human leukocyte antigens; TCA, T cell A locus; RCA, regulators of complement activation; C3, locus encoding the third complement component.

[†] Under the assumption of linkage.

haplotypes analogous to the "complotypes" of Alper et al. (18), composed of the four closely linked loci for the complement proteins, C2, B, C4A, and C4B. The alleles of *C3bR*, *C4BP*, and *FH* will therefore be referred to as regulators of complement activation (*RCA*) haplotype. Both the complotype gene cluster and the *RCA* gene cluster reported here consist of functionally related proteins; the complotype cluster consists of elements of a C3 convertase, the *RCA* cluster, of regulatory proteins for C3 convertase. Linkage disequilibrium is present between several pairs of complotype alleles (18), and also between the *RCA* alleles *C4BP*2* and *C3bR*B*. Larger groups of unrelated people will need to be phenotyped to determine *RCA* haplotype frequencies and to ascertain whether there is additional linkage disequilibria in this group. The biological significance of this new complement gene cluster, perhaps in determining disease susceptibility, remains to be ascertained. In view of the proposed regulation of the trafficking of C3- and C4-bearing immune complexes (19) and the findings of abnormalities of these proteins in systemic lupus erythematosus (20–22), further understanding of the genetics and function of this system of linked loci should provide insights into the control of putative immune complex-mediated illnesses.

It is interesting that the genes coding for most of the complement components are arranged in clusters and that the two other linkage groups described thus far for human complement components (i.e., *C2-Bf-C4A-C4B* and *C6-C7*) have been extremely well preserved throughout evolution (1, 2, 23–25). The *RCA* system should, therefore, also be investigated from this viewpoint. An apparently inconsistent situation might exist in mice because *C4BP* was shown to be located within the *H-2D-Qa* interval (26, 27) while *FH* segregated independently of the *H-2* system (28). This would indicate that *C4BP* and *FH* are not linked in the mouse. However, we have recently presented data disputing the assignment of *C4BP* to the *H-2* region (29). Thus, the issue of the evolutionary conservation of this genetic unit remains unsettled.

This linkage group is also of interest because two of the proteins, C4-bp and H, are plasma glycoproteins synthesized and secreted by hepatocytes, while the other, C3bR, is an integral membrane glycoprotein characteristically found on erythrocytes and leukocytes. The functional similarity and genetic relationship among these molecules may well be due to their being encoded by genes with a common ancestral origin (7). This hypothesis, however, implies the additional existence of a significant structural homology, for which there are no direct

supporting data. On the other hand, whether or not this cluster was formed in that fashion, its existence may relate to possible evolutionary consequences of the close linkage. Close linkage, for example, may favor the evolution of the genetic mechanisms for the control of the integrated expression of the genes within the cluster, and thereby influence the homeostatic adaptability of their carriers.

Summary

A new cluster of complement component genes, including *C4BP*, *C3bR*, and *FH*, is described. Family segregation data indicate that *FH* is linked to the genes for C4-bp and C4bR, previously reported to be linked and to maintain linkage disequilibrium. This cluster is not linked to the major histocompatibility complex, which contains the genes for the complement components, C4, C2, and factor B, or to the *C3* locus. These data further suggest that the organization of genes for functionally related proteins in clusters may be a rule for the complement system.

We thank Dr. J. M. Lalouel for discussions of linkage analysis, Dr. A. van Leeuwen for providing us with TCA-typed families, Ms. Joann Sala for manuscript preparation, and Ms. Telervo Huima and Mr. Howard Wynder for photographic work.

Received for publication 27 December 1984.

References

1. Alper, C. A. Complement and the MHC. 1981. *In* The Role of the Major Histocompatibility Complex in Immunobiology. M. E. Dorf, editor. Garland Publishing, Inc., New York. 173-220.
2. Porter, R. R. 1984. The complement components of the major histocompatibility locus. *CRC Crit. Rev. Biochem.* 16:1.
3. Rodriguez de Cordoba, S., A. Ferreira, V. Nussenzweig, and P. Rubinstein. 1983. Genetic polymorphism of human C4-binding protein. *J. Immunol.* 131:1565.
4. Rodriguez de Cordoba, S., and P. Rubinstein. 1984. Genetic polymorphism of human factor H (B1H). *J. Immunol.* 132:1906.
5. Dykman, T. R., J. L. Cole, K. Iida, and J. P. Atkinson. 1983. Polymorphism of human erythrocyte C3b/C4b receptor. *Proc. Natl. Acad. Sci. USA.* 80:1698.
6. Wong, W., J. Wilson, and D. Fearon. 1983. Genetic regulation of a structural polymorphism of human C3b receptor. *J. Clin. Invest.* 72:685.
7. Rodriguez de Cordoba, S., T. R. Dykman, F. Ginsberg-Fellner, G. Ercilla, M. Aqua, J. P. Atkinson, and P. Rubinstein. 1984. Evidence for linkage between the loci coding for the binding protein for the fourth component of human complement (C4BP) and for the C3b/C4b receptor. *Proc. Natl. Acad. Sci. USA.* 81:789.
8. Rodriguez de Cordoba, S., P. Rubinstein, and A. Ferreira. 1984. High resolution isoelectric focusing of immunoprecipitated proteins under denaturing conditions. A simple analytical method applied to the study of complement component polymorphisms. *J. Immunol. Methods.* 69:165.
9. Martel, J. L., S. Jaramillo, F. H. Allen, Jr., and P. Rubinstein. 1974. Serology for automated cytotoxicity assays. Contrast fluorescence test. *Vox Sang.* 27:13.
10. Van Rood, J. J., A. Van Leeuwen, and S. Ploem. 1976. Simultaneous detection of two cell populations by two color fluorescence and application to the recognition of B-cell determinants. *Nature (Lond.).* 262:795.

11. Alper, C. A., and A. M. Johnson. 1969. Immunofixation electrophoresis: a technique for the study of protein polymorphism. *Vox Sang.* 17:445.
12. Falk, C. T., M. E. Walker, M. D. Martin, and F. H. Allen, Jr. 1975. Autosomal linkage in humans (methodology and results of computer analysis). *Ser. Haematol.* 8:153.
13. Meyers, D. A., P. M. Conneally, E. W. Lovrien, R. E. Magenis, A. D. Merritt, J. A. Norton, C. G. Palmer, M. L. Rivas, L. Wang, and P. L. Yu. 1976. Linkage group I: the simultaneous estimation of recombination and interference. *Natl. Found. March Dimes Birth Defects Orig. Art. Ser.* 12(7):335.
14. Lathrop, G. M., J. M. Lalouel, C. Julier, and J. Ott. 1984. Strategies for multilocus linkage analysis in humans. *Proc. Natl. Acad. Sci. USA.* 81:3443.
15. Rao, C. R. 1973. *Linear Statistical Inference and Its Applications.* John Wiley & Sons, New York. 351-374.
16. Hatch, J. A., J. P. Atkinson, B. K. Suarez, and T. R. Dykman. 1984. Evaluation of linkage of the human C3b/C4b receptor to HLA. *J. Immunol.* 132:2168.
17. Van Leeuwen, A. 1982. Di-allelic allo-antigenic systems on human T-lymphocyte subsets. Ph.D. thesis. London Hospital Medical College, London.
18. Alper, C. A., D. Raum, S. Karp, Z. L. Awdeh, and E. J. Yunis. 1983. Serum complement "supergenes" of the major histocompatibility complex in man (complotypes). *Vox Sang.* 45:62.
19. Medof, M. E., K. Iida, C. Mold, and V. Nussenzweig. 1982. Unique role of the complement receptor CR1 in the degradation of C3b associated with immune complexes. *J. Exp. Med.* 156:1739.
20. Miyakawa, Y., A. Yamada, K. Kosaka, F. Tsuda, E. Kosugi, and M. Mayumi. 1981. Defective immune-adherence (C3b) receptor on erythrocytes from patients with systemic lupus erythematosus. *Lancet.* 2:493.
21. Iida, K., R. Mornaghi, and V. Nussenzweig. 1982. Complement receptor (CR1) deficiency in erythrocytes from patients with systemic lupus erythematosus. *J. Exp. Med.* 155:1427.
22. Wilson, J. G., W. W. Wong, P. A. Schur, and D. T. Fearon. 1983. Mode of inheritance of decreased C3b receptors on erythrocytes of patients with systemic lupus erythematosus. *N. Engl. J. Med.* 307:981.
23. Hobart, M. J., V. Joysey, and P. J. Lachmann. Inherited structural variation and linkage relationships of C7. 1978. *J. Immunogenet.* 5:157.
24. Whitehouse, D. B. 1984. Genetic polymorphism and linkage of the sixth and seventh complement components (C6 and C7) in the common marmoset. *Biochem. Genet.* 22:51.
25. Eldridge, P. R., M. J. Hobart, and P. J. Lachmann. 1983. The genetics of the sixth and seventh components of complement in dog: polymorphism, linkage, locus duplication, and silent alleles. *Biochem. Genet.* 21:81.
26. Kaidoh, T., S. Natsuume-Sakai, and M. Takahashi. 1981. Murine binding protein of the fourth component of complement: structural polymorphism and its linkage to the major histocompatibility complex. *Proc. Natl. Acad. Sci. USA.* 78:3794.
27. Takahashi, S., M. Takahashi, T. Kaidoh, S. Natsuume-Sakai, and T. Takahashi. 1984. Genetic mapping of mouse C4-BP locus to the H-2D-Qa interval. *J. Immunol.* 132:6.
28. Sakai, S., T. Kaidoh, K. Sudo, J. Hayakawa, and M. Takahashi. 1983. Genetic polymorphism of murine complement component, B1H, controlled by a structural locus. *Proc. JPN Soc. Immunol.* 13:92.
29. Rodriguez de Cordoba, S., A. Ferreira, and P. Rubinstein. 1985. Does the murine C4-binding protein gene (C4BP) map in the H-2 region? *Immunogenetics.* In press.