Polymorphism at the C-reactive protein locus influences gene expression and predisposes to systemic lupus erythematosus

Andrew I. Russell[1], **Deborah S. Cunninghame Graham**[1], **Christopher Shepherd**[1], **Cheri A. Roberton**[1], **John Whittaker**[2], **John Meeks**[3], **Richard J. Powell**[4], **David A. Isenberg**[5], **Mark J. Walport**[1,†], and **Timothy J. Vyse**[1,*]

[1]Rheumatology Section, Imperial College Faculty of Medicine, Hammersmith Hospital, London W12 0NN, UK

[2]Department of Epidemiology and Public Health, Imperial College Faculty of Medicine, St Mary's Campus, Norfolk Place, London W2 1PG, UK

[3]Department of Pathology, Imperial College Faculty of Medicine, Hammersmith Hospital, London W12 0NN, UK

[4]Clinical Immunology Unit, Queen's Medical Centre, Nottingham NG7 2UH, UK

[5]The Middlesex Hospital, University College London, Arthur Stanley House, Tottenham Street, London W1T 4NJ, UK

Abstract

Relative deficiency of pentraxin proteins is implicated in the pathogenesis of systemic lupus erythematosus. The C-reactive protein (CRP) response is defective in patients with acute flares of disease, and mice with targeted deletions of the serum amyloid P component gene (*Sap*) develop a lupus-like illness. In humans, the genes for CRP (*CRP*) and SAP (*APCS*) map to 1q23.2 within an interval linked with SLE. We have investigated the candidate genes *CRP* and *APCS* in two cohorts totalling 586 UK simplex SLE families. The inheritance of an intronic dinucleotide repeat and seven single nucleotide polymorphisms in the *CRP* and APCS genes was examined by application of family-based tests of association and linkage. Basal levels of CRP were influenced independently by two polymorphisms at the *CRP* locus, *CRP*2 and *CRP*4. Furthermore, the latter polymorphism was linked/associated with SLE and antinuclear autoantibody production. Thus, the polymorphism associated with reduced basal CRP was also associated with the development of SLE. These data support the hypothesis that defective disposal of potentially immunogenic material is a contributory factor in lupus pathogenesis. The identification of polymorphisms that determine basal CRP levels has implications in ischaemic heart disease, where CRP level is an important predictor of risk.

INTRODUCTION

Systemic lupus erythematosus (SLE) is a systemic autoimmune disease with highly variable clinical features including glomerulonephritis, neuropsychiatric disease and cutaneous manifestations. SLE is characterized by autoantibodies directed against a number of constituents of the cell nucleus. These include double-stranded DNA (dsDNA), chromatin

*To whom correspondence should be addressed. Tel: +44 2083832339; Fax: +44 2087433109; t.vyse@imperial.ac.uk.
†Present address: The Wellcome Trust, Euston Road, London NW1 2BE, UK.

and small nuclear ribonucleoproteins (snRNPs), and antibodies to negatively charged phospholipids (reviewed in 1,2).

The aetiology of SLE remains incompletely understood; it includes hormonal, environmental and genetic factors (3, reviewed in 4). Genetic susceptibility to SLE is inherited as a complex trait, and recent independent genome-wide searches, in ethnically diverse multiplex families, have identified multiple intervals linked with SLE (5–9). When these data sets are compared, an interval on the long arm of chromosome 1, 1q23–24, is linked with SLE in multiple populations (5,8). This interval is orthologous to a region of distal chromosome 1 in the mouse (10), which is known to harbour susceptibility loci for murine lupus in several mouse strains (reviewed in 4). The genes for the pentraxin proteins C-reactive protein (CRP) and serum amyloid P component (SAP) (11) map to these intervals in both species. They are candidates as disease susceptibility genes for lupus by virtue of their position and the physiological activity of their products.

The genes for CRP (*CRP*) and SAP (*APCS*) have been physically mapped to 1q23.2 on an Ensembl contig AL445528 (www.ensembl.org). They are paralogues sharing at least 50% nucleotide identity and are separated by approximately 130 kb (chromosome 1 working draft contig NT_004406). Both genes consist of two exons, and have disproportionately long $3'$ untranslated regions ($3'$-UTR) (12,13), suggesting a possible regulatory role for this sequence.

Defective processing of cellular waste has been proposed as a contributory factor in SLE (reviewed in 14). There is evidence that the innate immune system, particularly the complement system, may be protective against autoimmunity. CRP and SAP are highly conserved mammalian acute phase reactants that could play a role in this protection (reviewed in 15). These pentraxins bind to specific membrane constituents and nuclear material, the major autoantigens in SLE, released during cell death or exposed on the cell surface during apoptosis (16). CRP binds to small nuclear ribonucleoproteins (17) and polar phospholipids (18) and SAP binds to and solubilizes native long chromatin (19). Characteristically, in human SLE there is relative failure of the acute phase CRP response during active disease despite evident tissue inflammation (20). Similarly, the inducible SAP response is defective in the New Zealand NZB/W F_1 murine model (21). Treatment of NZB/ W F_1 mice with CRP ameliorates the accelerated phenotype induced by immunization with chromatin (22). Furthermore, mice rendered null for *Sap* by targeted gene deletion (in the context of a permissive genetic background) developed spontaneous antinuclear autoimmunity and a lupus-like glomerulonephritis (23). It has recently been shown that NZB/W F_1 mice carrying a human transgene for CRP exhibit less severe disease (24).

Induction of the acute phase CRP response in hepatocytes is promoted by the synergistic action of IL-1β and IL-6. The proximal 250 nucleotides of $5'$ flanking sequence of *CRP* contain response elements for multiple transcription factors, including members of several HNF and C/EBP families (25–29). The intron of CRP contains a polymorphic GT dinucleotide repeat sequence (30), length polymorphism of which has been correlated with basal CRP levels (31).

We propose that genetic polymorphism at the pentraxin locus contributes to susceptibility to human SLE. We have tested this hypothesis by identifying single-nucleotide polymorphisms (SNPs) across the *CRP* and *APCS* genes and then conducting a family-based study of linkage and association in two large cohorts of SLE families. Finally, we provide data supporting a mechanism that underlies the genetic association with SLE.

RESULTS

SNP identification

Direct sequencing across the *CRP* gene identified four SNPs and confirmed the intronic GT repeat (Fig. 1 and Tables 1 and 2). The proximal promoter (extending to −250 nucleotides) was not polymorphic; *CRP* 1, a tri-allelic SNP at position −286, did not alter known transcription factor motifs. No coding polymorphisms were evident in *CRP*, a synonymous mutation in the third nucleotide of codon 188 (leucine), resulting from a guanine to cytosine transition at nucleotide 641, was named *CRP* 2. In the 3′ portion of *CRP* two SNPs were present with high frequency: *CRP* 3 and *CRP* 4. The latter SNP is present in only the longer of the two *CRP* transcripts described. Neither of these polymorphisms disrupted known consensus sequences associated with altered mRNA stability. A low frequency cytosine to adenine exchange eight nucleotides upstream of *CRP* 3 was in strong linkage disequilibrium with this SNP, which did not add to the association and haplotype analysis. The *APCS* coding region was not polymorphic apart from a synonymous mutation in the third nucleotide of the valine-specifying codon at residue 144 (*APCS* 2). *APCS* 1 and *APCS* 3 were identified within the 5′ and 3′ flanking sequence respectively. *APCS* 1 lies upstream of the transcription initiation site but not within known transcription factor binding motifs. *APCS* 3 is located 16 nucleotides downstream of the 3′ limit of known *APCS* mRNA.

Transmission disequilibrium test for single markers

Ethnicity influenced allele frequency at a number of SNPs (Tables 1 and 2). The transmission disequilibrium test (TDT) was therefore performed using genotype from the 299 European-Caucasian families from cohort 1 (Table 3), and also in supplementary analysis by inclusion all 344 families. There was a trend to over-transmission of *CRP* $(GT)_{16}$, most frequent allele of the *CRP* dinucleotide repeat ($\chi^2 = 3.82$, $P = 0.05$); there was compensatory under-transmission of $(GT)_{21}$ allele. We also demonstrated that patients with SLE were more likely to inherit the minor allele of SNP *CRP* 4 ($P = 0.05$). The inclusion of data from the 45 non-European families increased the transmission distortion for both markers ($P = 0.03$ and $P = 0.02$, respectively). Application of the pedigree disequilibrium test (PDT) confirmed the association of *CRP* 4 with SLE in European-Caucasian families ($P = 0.007$). Using PDT the trend to over-transmission of *CRP* $(GT)_{16}$ did not reach significance ($Z = 1.69$, $P = 0.09$). Segregation of *APCS* 3 alleles was distorted in European families, but this did not reach statistical significance. Other SNPs segregated at random with no association/linkage to SLE.

CRP 4 was also associated/linked with the presence of antinuclear autoantibodies (ANA) in probands and seropositive siblings ($\chi^2 = 4.13$, $P = 0.04$). However no associations were observed between other autoantibodies (anti-dsDNA, anti-Ro, anti-C1q, IgG ACA and IgM ACA) and pentraxin polymorphisms (data not shown). Stratification for renal involvement in the probands (128 families in cohort 1) was negative for association/linkage between renal lupus and any pentraxin polymorphism.

Linkage disequilibrium and haplotype analyses

Linkage disequilibrium (LD) analysis in Europeans identified strong LD across *CRP* (D' 0.91–0.98). A separate haplotype block was observed spanning the *APCS* locus. The degree of LD between these two genes was weak (D' 0.21–0.70; Fig. 2). The strength of LD within each pentraxin gene was reflected by restricted haplotype diversity. Five *CRP* haplotypes represented at least 85% all haplotypes present (Table 4). Using SNP genotypes only and the GENEHUNTER package to conduct TDT in trios there was strong support for association with *CRP* haplotype 2, carrying the rare allele at *CRP* 4 ($P = 0.0008$). This represents a statistically significant association even if the conservative Bonferroni correction of five is

applied reflecting the five common haplotypes across *CRP*. The microsatellite *CRP* (GT)$_{16}$ allele contributed to three common *CRP* haplotypes, of which only haplotype 2 (1.16.1.1.2) was associated with SLE when trios and single-parent families were studied (TRANSMIT *P* = 0.007). Since the association with SLE was observed with haplotype 2 and not with haplotype 1 (Table 4), the association between *CRP* (GT)$_{16}$ and SLE observed in cohort 1 was the result of its presence on a haplotype defined by the minor allele of *CRP* 4. Three *APCS* haplotypes accounted for nearly 100% of genotypes; there was a weak association of SLE with a haplotype carrying the rare allele at *APCS* 3 (*P* = 0.004). Using TDT analysis of all seven SNPs across the pentraxin locus, no haplotypes were associated with lupus or related phenotypes.

Cohort 2 and combined analysis

The rare allele at *CRP* 4 is a unique marker for the *CRP* haplotype that showed association with SLE. Thus, the inheritance of *CRP* 4 was studied in further families. As we describe later, *CRP* 4 and *CRP* 2 were shown to be associated with basal CRP serum levels (see below). Thus, *CRP* 2 and *CRP* 4 were tested for association in a second cohort of 242 families (cohort 2, Table 5). Genotyping of CRP 2 and CRP 4 defines three haplotypes corresponding to haplotypes 1, 2 and 5 (Table 4). The second cohort was slightly smaller than the first cohort with fewer transmissions subject to analysis. However, there was a suggestion of transmission distortion at *CRP* 4 towards lupus probands (analysis of 175 trios *P* = 0.03, analysis of all families TRANSMIT *P* = 0.08). The transmissions of *CRP* 2 and *CRP* 4 were then studied in all families from cohort 1 and cohort 2 together. This analysis increased the strength of association of the haplotype carrying the rare allele of *CRP* 4 with SLE was increased (TRANSMIT χ^2 = 7.4, *P* = 0.006), in 498 European-Caucasian families. This association was strengthened when all 586 families (independent of ethnicity) were analysed (TRANSMIT *P* = 0.003). When trios only were studied the association with *CRP* 4 was stronger (143 transmissions compared with 92 non-transmissions *P* = 0.0008) in all 394 trios. This result may reflect bias introduced by inferring missing parental haplotypes. There was no over-transmission of *CRP* 2 observed, but this SNP is at low frequency (0.07). As with cohort 1, no particular subset of SLE appeared to be particularly associated with *CRP* 4. However, we did run the TDT examining for association with ANA production in the combined cohorts, again there was a significant over-transmission of the rare allele of *CRP* 4 to ANA positive individuals (median of 25 TRANSMIT iterations χ^2 = 12, *P* = 0.002).

Comparison of basal CRP in lupus probands and healthy siblings

Having demonstrated evidence for association of the rare allele of *CRP* 4 with SLE we next sought to determine whether this polymorphism (and other allelic variants of *CRP*) influenced serum CRP levels. Thus, CRP levels were measured in the serum of 344 probands and 452 siblings. This analysis was constrained to those individuals without an apparent acute phase response in whom CRP levels were measured within the normal range, below 10 mg/l. We chose to investigate the potential effect of polymorphisms on basal CRP expression rather than acute phase expression, which most likely involves different regulatory elements at the *CRP* locus. Therefore, 27 probands and 31 siblings were excluded because of an evident acute phase response. Basal CRP levels were not significantly different between siblings and probands (*P*=0.24, two-tailed Mann–Whitney test).

The effect of CRP haplotype on serum CRP level

European-Caucasian individuals with unambiguous haplotype assignment by TRANSMIT were included in this analysis; 49 probands and 58 siblings were therefore excluded. Probands (*n* = 234) and siblings (*n* = 302) were analysed separately, but since the groups were not statistically different, the data sets were also combined (Fig. 3). Haplotype 2 (1.16.1.1.2) and haplotype 5 (1.16.2.1.2) were associated with low levels of CRP in serum

(mean 2.05 and 1.27 mg/l, respectively) compared with the common haplotype 1 (2.16.1.2.1), mean basal CRP 2.82 mg/l; these comparisons were significant, $P < 0.0001$ (haplotypes 2 and 5 versus haplotype 1), and $P = 0.02$ (haplotype 2 versus haplotype 5), in the combined analysis of 536 individuals. Haplotype 5 is likely to have arisen on the background of an ancestral haplotype 2, since these two haplotypes differ only at *CRP* 2, implicating both *CRP* 2 and *CRP* 4 in the association with low basal CRP levels.

The effect of *CRP* SNP alleles on serum CRP level

Analysis of basal CRP levels in individuals with differing *CRP* 4 genotype demonstrated a gene dose-specific effect (Fig. 4A); homozygosity for allele 2 was associated with the lowest mean level of CRP. Furthermore, an additional effect of *CRP* 2 genotype was also apparent (Fig. 4B). The effect of *CRP* 2 and *CRP* 4 allelic variants on basal CRP levels was also studied using modified linear regression. These analyses indicated correlation with *CRP* 4 genotype ($P = 0.003$) and an independent effect from *CRP* 2 ($P = 0.01$). The weaker association with *CRP* 2 reflects the lower allele frequency of this SNP. When affected status was added to the model, it was evident that the association of *CRP* 4 and *CRP* 2 with basal CRP levels was present in both SLE patients and in their siblings. We used the parameters derived from the regression model to predict basal CRP expression based on genotype at *CRP* 2 and *CRP* 4; inheritance of minor alleles at both SNPs predicted a CRP level of 0.9 mg/l, whereas inheritance of the common alleles at both SNP predicted a value of 3.4 mg/l.

DISCUSSION

We tested the hypotheses that polymorphism at the pentraxin locus might contribute to genetic susceptibility to human SLE, and that genetic variation in *CRP* could specify basal levels of C-reactive protein. The pentraxin proteins are important in innate immunity and in addition bind certain constituents of the cell nucleus, promoting their clearance and reducing their immunogenicity. Structural variants of the proteins have not been described; consistent with our observation that the coding sequences of both genes were monomorphic. Thus the aetiological effect of pentraxin gene polymorphism would be expected to be quantitative rather than qualitative.

We examined the inheritance of polymorphisms at the pentraxin locus in a family-based association study of SLE. Strong LD was present within each of the *CRP* and *APCS* genes, although minimal LD exists between the two genes. There are no functional genes that lie between *CRP* and *APCS.* We identified five common *CRP* haplotypes by characterizing the inheritance of four SNPs and one polymorphic microsatellite. By application of transmission testing we demonstrated that inheritance of *CRP* haplotype 2 by affected probands with SLE from their parents was not random. Considering the composition of the *CRP* haplotypes, it was apparent that the specific allelic variant associated with SLE was an allele of *CRP* 4. We demonstrated using both haplotype analysis and single marker inheritance that the less common A allele of the single nucleotide polymorphism *CRP* 4 in the 3′ region of the *CRP* gene is associated with the development of SLE. The association of this allele with SLE was detected in all 586 families ($P = 0.006$) using an analysis package TRANSMIT to infer missing parental genotypes. When only trios were examined, that is family units comprising both parents and an affected offspring, strong statistical support was obtained for the association in 394 families ($P = 0.0008$). The A allele at *CRP* 4 had a relatively high frequency in European and Asian-Indian populations (~0.3) and was present in Afro-Caribbean families too, albeit at a lower frequency (0.14). The association was present when all families were studied and when only European families were selected; there were insufficient numbers of non-European families to study separately.

SLE is a clinically heterogeneous disorder. Basic clinical and serological data were obtained on all affected family members and their siblings. We repeated the family-based association tests in SLE patients that had been categorized according to the presence or absence of renal disease, or by their autoantibody profiles. There was no evidence that any particular subgroup of SLE patients preferentially inherited the rare allele of *CRP*4, although there is an inherent loss of power in these subgroup analyses as the number of families analysed falls. Previous studies have indicated an elevated frequency of ANA in SLE families (32). Approximately 30% of the siblings in the SLE families had elevated antinuclear antibodies in the cohort investigated in this study and there was evidence of familial clustering of antinuclear antibody production (manuscript in preparation). Using the production of ANA to define affection status, there was evidence for distorted transmission of *CRP*4 alleles to offspring that produce antinuclear antibodies (*P*=0.002), suggesting that this allele may act to promote susceptibility to antinuclear autoimmunity.

Having shown that the rare allele of *CRP*4 was associated with the development of SLE, we undertook a study of basal CRP expression in the SLE patients and their siblings to determine whether this polymorphism had any effect on serum CRP levels. Basal CRP levels as measured by high-sensitivity CRP assay were compared between individuals to whom unambiguous CRP haplotypes could be assigned. Two haplotypes were defined that were significantly associated with reduced basal CRP expression. The SNPs that demarcated these haplotypes were *CRP*2 and *CRP*4. Linear regression analyses corroborated the conclusion that these two SNPs made an independent contribution to basal CRP levels and that their influence was independent of SLE status.

On the basis of these findings we suggest that reduced basal CRP expression predisposes to the development of SLE. However this would predict an association between the *CRP*2 locus and SLE, since the rare allele of *CRP*2 was associated with reduced basal CRP expression. The failure to detect an association between this SNP (frequency 0.07) and SLE may be explained by consideration of power. Simple power calculations indicate that for genotypic relative risk (GRR) of 1.75, 80% power (with $P=10^{-3}$) is achieved using 740 trios for a disease allele frequency of 0.07 (33). Thus our study is currently underpowered to reliably detect a weak effect from *CRP*2, if it exists. The numbers of families examined in this study are sufficient to produce 80% power to detect disease susceptibility alleles with frequency of 0.09 or greater with a modest effect (GRR=1.75). Therefore, single marker TDT and haplotype analysis should exclude the *APCS*1 and *APCS*2 polymorphisms examined here in genetic susceptibility to human SLE.

An allele of *CRP*4 was associated with ANA production. This observation has parallels in murine models since mouse strains deficient in production of the pentraxin SAP express antinuclear autoimmunity (23). It could also be relevant that autoantibody production in lupus-prone mice maps to an orthologous locus on distal mouse chromosome 1 (34,35). The association of the rare allele of *CRP*4 with antinuclear autoantibody production supports the 'waste disposal' hypothesis of SLE (16). Since there is evidence supporting the concept that CRP participates in the removal of apoptotic cells (36), our data support the view that low basal CRP expression impairs the 'physiological' removal of cellular debris. The data presented in this paper support the concept that this is a factor contributing to the development of SLE and antinuclear autoimmunity. Data from *Dnase* 1 and *Sap* knockout mice have been cited to support this hypothesis (reviewed in 14), but with the exception of rare mutations in *DNASE1* within the Japanese population (37) there has been little direct genetic evidence to support this concept in human SLE.

There are several potential mechanisms by which variants at *CRP*2 and *CRP*4 might influence basal CRP expression. *CRP* appears to generate two transcripts (12)

[www.ensembl.org, designated ENST00000325798 (longer transcript), and ENST00000255030 (shorter transcript)]; the longer of these carries *CRP* 4 in the 3′ untranslated region. The relative biological importance of these two transcripts is not known. The stability of mRNA species may depend on the sequences within the 3′-UTR, as exemplified by the AU-rich sequence motifs that are important in regulation of cytokine mRNA (38,39). It is of interest that both *CRP* and *APCS* have disproportionately long 3′-UTRs (Fig. 1), implying a regulatory role in pentraxin mRNA processing. In a second shorter *CRP* transcript, *CRP* 4 is located within 3′ flanking sequence just upstream of a putative polyA signal, consistent with the reported influence of the 3′ flanking sequence on CRP expression (40). The mechanism by which *CRP* 4 influences basal CRP expression remains to be established, but it may operate at the transcriptional or post-transcriptional stage.

Pro-inflammatory cytokines IL-1β and IL-6 drive the acute phase response of CRP via response elements located within the proximal 250 nucleotides of 5′ sequence (25–29). Since levels of these cytokines are known to be elevated in patients with active SLE (41), functional polymorphisms within the *CRP* promoter could explain the inappropriately normal pentraxin levels observed in lupus. However, we observed the sequence of this region to be non-polymorphic. Furthermore, polymorphism at *CRP* 1, located upstream of this sequence, did not alter consensus sequence of known motifs of transcription factors, was not associated with the presence of SLE, or influence basal CRP levels. It should be pointed out that we have not addressed the possible effect of the *CRP* 1 on the magnitude of the CRP rise during an acute phase reaction.

The GT dinucleotide repeat within the intron of *CRP* has been reported to influence *CRP* expression (31); the most frequent alleles, $(GT)_{16}$ and $(GT)_{21}$, were associated with low basal levels of the protein. The data presented in this paper refine the microsatellite association by considering the relation between CRP levels and CRP haplotypes (Table 4). The two *CRP* haplotypes associated with reduced basal CRP expression, haplotypes 2 and 5, also carry the $(GT)_{16}$ repeat. Thus our data are in part compatible with those previously reported (31). However, haplotype 3, which carries $(GT)_{21}$, was associated with intermediate levels of basal CRP (Fig. 3). Since the *CRP* SNPs described generate a more accurate definition of *CRP* haplotypes than the microsatellite alleles alone, it is unlikely that length polymorphism of this intronic repeat exerts a direct biological effect on CRP regulation.

The existence of genetic regulation of levels of CRP expression has implications beyond the predisposition to autoimmunity. Over the last five years there has been increasing evidence of the predictive value of basal CRP in atherosclerosis. An association between a SNP (rs1800947, or *CRP* 2 as defined in this paper) in *CRP* and basal CRP expression has been described (42). The data presented in this manuscript are therefore in agreement with this report, but provide a more comprehensive, haplotype-based approach to the issue of the genetics of basal CRP regulation. Our data implicate independent effects from *CRP* 2 and *CRP* 4. The rare allele frequency of *CRP* 2 in our study was 0.07, a similar frequency to that observed by Zee and Ridker (0.06). They did not find an association between the *CRP* 2 polymorphism and arterial thrombosis. However, polymorphism at *CRP* 2 is not the sole genetic cause at the *CRP* locus of variation in CRP basal levels. Recently it has been reported that the lower frequency allele a SNP in the 3′-UTR of *CRP* (rs1130864, our *CRP* 3) is associated with elevated basal CRP (43). The rare allele of this SNP is carried on haplotype 1, as defined in this study (Table 4). This SNP does not uniquely demarcate this haplotype, which is also characterized by a polymorphism in the *CRP* promoter, *CRP* 1. Our data are compatible, but amplify this finding and indicate that *CRP* 3 may not necessarily be the causal variant on this haplotype. Thus, it remains possible therefore that polymorphism

at the *CRP* locus influences the development and/or progression of atherosclerosis and we are investigating this hypothesis.

We conclude that polymorphism that a polymorphism in the 3′ flanking sequence of *CRP* gene is associated with SLE and antinuclear autoantibody production, perhaps by specifying low levels of the protein. Relative deficiency of the chaperoning properties of CRP therefore could predispose to antinuclear autoantibody production which in turn contributes to the development of human lupus. These findings have relevance in the study of other conditions characterised by variation of CRP responses, in particular ischaemic heart disease.

MATERIALS AND METHODS

SLE Families

We have established a large collection of UK SLE nuclear families with predominantly one affected offspring per family. The demographic details of these families are summarized in Table 5. Of the total of 586 families, samples from both parents were available in 69% of cases; samples were taken from siblings when available. Siblings were always sampled in single parent families. The mean number of siblings in two parent families was 1.81 and the mean number in single parent families was 1.97; 44% of the siblings were male. The collection comprises genomic DNA and serum from patients who satisfy the American College of Rheumatology revised criteria for SLE (44). Patients were categorized as having renal lupus using ACR criteria. Clinical information was obtained from all participants by interview and completion of a questionnaire. Genomic DNA was isolated by a standard phenol–chloroform extraction from peripheral blood mononuclear cells (45). The study protocol was approved by the London multi-centre research ethics committee (MREC).

The principal study cohort represents a random sample of 354 SLE families. Ten families were excluded from analysis due to ex-paternity of the pedigree, incorrect diagnosis or withdrawal from the study. The remaining 344 families are hereafter referred to as cohort 1. When a preliminary positive finding was obtained for the inheritance of a marker in cohort 1 replication was sought in a second cohort of 242 families (cohort 2).

The stored serum of all probands and their siblings was assayed for antinuclear autoantibodies and CRP. CRP was measured using a commercially available, automated high sensitivity immunoturbidimetric method (Olympus, UK). Briefly, 2 μl of serum sample was mixed with a 160 μl of a 0.05% suspension of latex particles coated with goat anti-human CRP antibodies, in the presence of MOPSO buffer (pH 7.5). The presence of the resultant insoluble aggregates was detected at 800 nm and is proportional to the CRP concentration of the sample. Samples were run on an Olympus AU640e Chemistry Immunoanalyser (Olympus, UK). Antinuclear antibodies (ANA) were assayed by indirect immnuofluorescence; anti-double stranded DNA antibodies (anti-dsDNA), antibodies to extractable nuclear antigens (Ro, La, Sm, RNP) and IgG and IgM isotypes of anticardiolipin antibodies were assayed by ELISA. The methods employed are described elsewhere (46–48).

SNP identification

SNPs were identified by direct sequencing of the complete genomic region of *CRP* and *APCS* in 45 SLE patients representative of the study population (Table 5). Direct sequencing was performed in both orientations in order exclude sequencing artefacts. SNPs were then checked for uniqueness by the computer program BLAST (www.ncbi.nlm.nih.gov/BLAST) and cross-referenced against publicly available SNP databases, the SNP Consortium (http://snp.cshl.org) and dbSNP (www.ncbi.nlm.nih.gov/SNP).

CRP was amplified in two overlapping 1154 and 1343 bp amplicons, using reference sequence reported by Woo *et al.* (accession M11880.1). Oligonucleotide primers designed from chromosome 1 working draft sequence NT_004406 were generated to amplify a further 1076 bp of 5′ flanking sequence. *APCS* was amplified with flanking sequence as a single 1354 bp PCR product using primers based on NCBI reference sequence (accession D00097.1) (13). The oligonucleotide sequences used are available on request. Polymerase chain reaction (PCR) amplifications were performed on 50 ng of genomic DNA using standard conditions (available on request) and were subjected to thermal cycling using a Peltier Thermal Cycler (MJ Research). Resultant PCR product was purified using a 96-well format microplate vacuum system according to manufacturer's instructions (MultiScreen®-PCR Plates, Millipore). A 30–90 ng aliquot of purified PCR product was subjected to a cycle sequencing reaction using a fluorescent-labelled dideoxynucleotide kit (Big Dye Terminator Cycle Sequencing Ready Reaction Kit, Applied Biosystems), and extension products purified by 75% isopropanol precipitation. Sequenced samples were electrophoresed through denaturing 5% polyacrylamide gels (Long Ranger™) on an ABI PRISM 377 DNA Sequencer. Sequence data was analysed using Sequence Analysis version 3.4.1 and the alignment program Autoassembler™ (Applied Biosystems). SNPs with a minor allele frequency of more than 5% were then investigated in the study cohorts.

SNP genotyping

The preferred method of genotyping was polymerase chain reaction–restriction fragment length polymorphism (PCR-RFLP). Where this method was unsuitable an amplification refractory mutation system–polymerase chain reaction (ARMS-PCR) assay was used.

PCR-RFLP

The genotyping assays employed are described in Table 6. In general, restriction maps were generated for each allele of a given SNP using REMAP [EMBOSS sequencing software package, via UK HGMP Resource Centre (www.hgmp.mrc.ac.uk)]. Oligonucleotide primers (Sigma Genosys) were designed to amplify 100–300 bp flanking the SNP and genomic DNA from each individual was amplified as described above. PCR product was digested to completion in appropriate buffer according to the manufacturer's specifications. The resulting fragments were size separated on 3% agarose or 10% polyacrylamide gels and visualised by ethidium bromide staining. *APCS* 2 was genotyped by a reverse primer-determined PCR-RFLP using the enzyme *Fok* I. To enhance allelic discrimination a 20meric tail of random nucleotides (ACA ATC TTG GGC TGA GCA TC) was added to the reverse primer increasing the size of the smallest fragment.

ARMS-PCR

APCS 3 was genotyped using ARMS-PCR. Paired amplifications were performed using a common forward primer and one of each allele specific reverse primer (designed with an allele-specific 3′ nucleotide and the insertion of a penultimate 3′ base mismatch which increased allele specificity). PCR conditions included 1.5 mM $MgCl_2$, and an annealing temperature 61°C. An internal control primer pair was included in each reaction in order to ensure the efficiency of amplification. Primer sequences were as follows: common forward primer, 5′-GGA GTC CTC ATC AGG TAT TG-3′; G allele-specific reverse primer, 5′-GAT ATG GAC TAT GAC AAT C-3′; T allele-specific reverse primer, 5′-GAT ATG GAC TAT GAC AAT A-3′; control forward primer, 5′-GTC TCA GCT CAC GAT CAC CA-3′; control reverse primer, 5′-CCG TAG CTG CCT GTA CCA AT-3′. The resultant 542 bp (and 249 bp control) amplicons were electrophoresed through 1% agarose gels and visualized by ethidium bromide staining. The genotype of those individuals with a positive T allele PCR was confirmed by direct sequencing.

CRP dinucleotide repeat genotyping

Fluorescent primer PCR was performed using the primer pair described in [30], with 5′ fluorescent labelling of the forward primer: forward primer, 5′-FAM-GAT CTA TCC CCT CAC TTA CG-3′; reverse primer, 5′-TAT GAA CAG AAC AGT GGA GC-3′.

PCR was performed as described above and amplification products were size-separated on a 3700 ABI PRISM DNA Sequencer (Applied Biosystems) against the Genescan-ROX 400HD size standard. Data was collected using GENESCAN 3.1 software (Applied Biosystems) and allelic assignment was semi-automated, using Genotyper 2.0 software (Applied Biosystems).

Statistical analysis

Allele frequencies were calculated from parental chromosomes for each ethnic group and compared using the chi square statistic. The pattern of LD and the presence of haplotype blocks across both pentraxin genes and across the locus as a whole were defined in the predominant European-Caucasian group using haplotype frequencies obtained from Transmit analysis (49) and from an estimation of D' (50) obtained from an extension of Genehunter 2 (Courtesy of Dr Mark Daly, Whitehead Institute).

Genotype data from SLE families were analysed by application of two family-based tests of association. In the TDT, transmission of marker alleles (or haplotypes) from heterozygous parents to affected individuals is examined and significant distortion away from random segregation is used to infer linkage and association (51). TDT was performed using the computer program Transmit version 2.5 (49) which is allows the reconstruction of missing parental genotype (52). Transmission of single markers was examined in addition to haplotype analyses. In order to optimise the information provided by the SLE family genotype data, analysis including genotype data from siblings was also performed using the PDT version 3.12 (53). This test integrates information from pedigrees irrespective of structure and size and is based on a random variable measuring transmission distortion across the entire pedigree.

Cohort 1 was analysed using SLE as the definition of 'affection' status, and also with families stratified for ethnicity and the presence of renal involvement. Markers with putative association/linkage in cohort 1 were also tested in cohort 2 and by combination of both data sets. The presence of autoantibodies was observed to cluster in families (manuscript in preparation) with autoantibody-positive siblings present in 30% of families. The presence of different autoantibodies was also used to define affection status for TDT analysis. To achieve this, TRANSMIT was constrained to select a single affected individual from each family at random. Multiple iterations of TRANSMIT were performed and the chi square values obtained for each SNP and their distribution was tested using the PRISM software package. The arithmetic mean of the chi square values was then used to derive the point P value for this analysis. Power calculations for TDT were performed based on the methods described by Risch and Merikangas (33).

Regression analysis comparing basal CRP with *CRP* genotype was conducted on a linear model. Since the serum CRP concentrations were not normally distributed, Box–Cox analysis (54) indicated that the data be transformed using the function $([CRP]^{0.17} - 1)/0.17$. We assumed a random family effect to represent correlation between sibs that might result from shared environmental factors or other genetic influences.

Acknowledgments

REFERENCES

1. Morrow, WJ.; Nelson, JL.; Watts, R.; Isenberg, DA. Autoimmune Rheumatic Disease. 2nd edn. Oxford University Press; Oxford: 1999. p. 56-103.

2. Pisetsky DS. Anti-DNA and autoantibodies. Curr. Opin. Rheumatol. 2000; 12:364–368. [PubMed: 10990170]

3. Deapen D, Escalante A, Weinrib L, Horwitz D, Bachman B, Roy-Burman P, Walker A, Mack TM. A revised estimate of twin concordance in systemic lupus erythematosus. Arthritis Rheum. 1992; 35:311–318. [PubMed: 1536669]

4. Vyse TJ, Kotzin BL. Genetic susceptibility to systemic lupus erythematosus. A. Rev. Immunol. 1998; 16:261–292.

5. Moser KL, Neas BR, Salmon JE, Yu H, Gray-McGuire C, Asundi N, Bruner GR, Fox J, Kelly J, Henshall S, et al. Genome scan of human systemic lupus erythematosus: evidence for linkage on chromosome 1q in African-American pedigrees. Proc. Natl Acad. Sci. USA. 1998; 95:14869–14874. [PubMed: 9843982]

6. Gaffney PM, Kearns GM, Shark KB, Ortmann WA, Selby SA, Malmgren ML, Rohlf KE, Ockenden TC, Messner RP, King RA, et al. A genome-wide search for susceptibility genes in human systemic lupus erythematosus sib-pair families. Proc. Natl Acad. Sci. USA. 1998; 95:14875–14879. [PubMed: 9843983]

7. Gaffney PM, Ortman WA, Selby SA, Shark KB, Ockenden TC, Rohlf KE, Walgrave NL, Boyum WP, Malmgren ML, Miller ME, et al. Genome screening in human systemic lupus erythematosus: results from a second Minnesota cohort and combined analysis of 187 sib-pair families. Am. J. Hum. Genet. 2000; 66:547–555. [PubMed: 10677315]

8. Shai R, Quismorio FP Jr, Li L, Kwon OJ, Morrison J, Wallace DJ, Neuwelt CM, Brautbar C, Gauderman WJ, Jacob CO. Genome-wide screen for systemic lupus erythematosus susceptibility genes in multiplex families. Hum. Mol. Genet. 1999; 8:639–644. [PubMed: 10072432]

9. Johanneson B, Steinsson K, Lindqvist AK, Kritjansddottir J, Grondal G, Sandino S, Tjernstrom F, Sturfelt G, Granados-Arriola J, Alcocer-Varela J, et al. Comparison of Genome-scans performed in multi-case families with systemic lupus erythematosus from different populations. J. Autoimmun. 1999; 13:137–141. [PubMed: 10441178]

10. Kingsmore SF, Watson ML, Howard TA, Seldin MF. A 6000 kb segment of chromosome 1 is conserved in human and mouse. EMBO J. 1989; 8:4073–4080. [PubMed: 2591369]

11. Walsh MT, Divane A, Whitehead AS. Fine mapping of the human pentraxin gene region on chromosome 1q23. Immunogenetics. 1996; 44:62–69. [PubMed: 8613143]

12. Woo P, Korenberg JR, Whitehead AS. Characterization of genomic and complementary DNA sequence of human C-reactive protein, and comparison with the complementary DNA sequence of serum amyloid P component. J. Biol. Chem. 1985; 260:13384–13388. [PubMed: 3840479]

13. Ohnishi S, Maeda S, Shimada K, Arao T. Isolation and characterization of the complementary and genomic DNA sequences of human serum amyloid P component. J. Biochem. (Tokyo). 1986; 100:849–858. [PubMed: 3029048]

14. Pickering MC, Botto M, Taylor PR, Lachmann PJ, Walport MJ. Systemic lupus erythematosus, complement deficiency, and apoptosis. Adv. Immunol. 2000; 76:227–324. [PubMed: 11079100]

15. Carroll M. Innate immunity in the etiopathology of autoimmunity. Nat. Immunol. 2001; 2:1089–1090. [PubMed: 11725293]

16. Casciola-Rosen L, Andrade F, Ulanet D, Wong WB, Rosen A. Cleavage by granzyme B is strongly predictive of autoantigen status: implications for initiation of autoimmunity. J. Exp. Med. 1999; 190:815–826. [PubMed: 10499920]

17. Du Clos TW. C-reactive protein reacts with the U1 small ribonucleoprotein. J. Immunol. 1989; 143:2553–2559. [PubMed: 2477447]

18. Volanakis JE, Wirtz KW. Interaction of C-reactive protein with artificial phosphatidylcholine bilayers. Nature. 1979; 281:155–157. [PubMed: 471064]

19. Pepys MB, Butler PJG. Serum amyloid P component is the major calcium-dependent specific DNA binding protein of the serum. Biochem. Biophys. Res. Commun. 1987; 148:308–313. [PubMed: 3675579]

20. Pepys MB, Lanham JG, de Beer FC. C-reactive protein in SLE. Clin. Rheum. Dis. 1982; 8:91–103. [PubMed: 6749401]

21. Rordorf C, Schnebli HP, Baltz ML, Tennent GA, Pepys MB. The acute phase response in (NZB X NZW)F1 and MRL/lpr mice. Contrasting patterns resembling systemic lupus erythematosus and rheumatoid arthritis respectively. J. Exp. Med. 1982; 156:1268–1273. [PubMed: 7153712]

22. Du Clos TW, Zlock LT, Hicks PS, Mold C. Decreased autoantibody levels and enhanced survival of (NZB X NZW)F1 mice treated with C-reactive protein. Clin. Immunol. Immunopathol. 1994; 70:22–27. [PubMed: 8261665]

23. Bickerstaff MC, Botto M, Hutchinson WL, Herbert J, Tennent GA, Bybee A, Mitchell DA, Cook HT, Butler PJ, Walport MJ, et al. Serum amyloid P component controls chromatin degradation and prevents antinuclear autoimmunity. Nat. Med. 1999; 6:694–697. [PubMed: 10371509]

24. Szalai AJ, Weaver CT, McCrory MA, van Ginkel FW, Reiman RM, Kearney JF, Marion TN, Volanakis JE. Delayed lupus onset in (NZB × NZW)F_1 mice expressing a human C-reactive protein transgene. Arthritis Rheum. 2003; 48:1602–1611. [PubMed: 12794828]

25. Ganter U, Arcone R, Toniatti C, Morrone G, Ciliberto G. Dual control of C-reactive protein gene expression by interleukin-1 and interleukin-6. EMBO J. 1989; 8:3773–3779. [PubMed: 2555173]

26. Majello B, Arcone R, Toniatti C, Ciliberto G. Constitutive and IL-6-induced nuclear factors that interact with the human C-reactive protein promoter. EMBO J. 1990; 9:457–460. [PubMed: 2154372]

27. Li SP, Goldman ND. Regulation of human C-reactive protein gene expression by two synergistic IL-6 responsive elements. Biochemistry. 1996; 35:9060–9068. [PubMed: 8703909]

28. Zhang D, Sun M, Samols D, Kushner I. STAT3 participates in transcriptional activation of the C-reactive protein gene by interleukin-6. J. Biol. Chem. 1996; 271:9503–9509. [PubMed: 8621622]

29. Agrawal A, Cha-Molstad H, Samols D, Kushner I. Transactivation of C-reactive protein by IL-6 requires synergistic interaction of CCAAT/enhancer binding protein (C/EBPbeta) and Rel p50. J. Immunol. 2001; 166:2378–2384. [PubMed: 11160296]

30. Weber JL, Kwitek AE, May PE. Dinucleotide repeat polymorphism at the *CRP* locus. Nucl. Acids Res. 1990; 18:4635.

31. Szalai AJ, McCrory MA, Cooper GS, Wu J, Kimberly RP. Association between baseline levels of C-reactive protein (CRP) and a dinucleotide repeat polymorphism in the intron of the *CRP* gene. Genes. Immun. 2002; 3:14–19. [PubMed: 11857055]

32. Isenberg DA, Shoenfeld Y, Walport M, Mackworth-Young C, Dudeney C, Todd-Pokropek A, Brill S, Weinberger A, Pinkas J. Detection of cross-reactive anti-DNA antibody idiotypes in the serum of systemic lupus erythematosus patients and of their relatives. Arthritis Rheum. 1985; 9:999–1007. [PubMed: 3876101]

33. Risch N, Merikangas K. The future of genetic studies of complex human diseases. Science. 1996; 273:1516–1517. [PubMed: 8801636]

34. Rozzo SJ, Allard JD, Choubey D, Vyse TJ, Izui S, Peltz G, Kotzin BL. Evidence for an interferon-inducible gene, *Ifi202*, in the susceptibility to systemic lupus. Immunity. 2001; 15:435–443. [PubMed: 11567633]

35. Morel L, Blenman KR, Croker BP, Wakeland EK. The major murine systemic lupus erythematosus susceptibility locus, *Sle1*, is a cluster of functionally related genes. Proc. Natl Acad. Sci. USA. 2001; 98:1787–1792. [PubMed: 11172029]

36. Mold C, Baca R, Du Clos TW. Serum amyloid P component and C-reactive protein opsonize apoptotic cells for phagocytosis through Fc receptors. J. Autoimmun. 2002; 19:147–154. [PubMed: 12419285]

37. Yasutomo K, Horiuchi T, Kagami S, Tsukamoto H, Hashimura C, Urushihara M, Kuroda Y. Mutation of *DNASE1* in people with systemic lupus erythematosus. Nat. Genet. 2001; 28:313–314. [PubMed: 11479590]

38. Bakheet T, Frevel M, Williams B, Greer W, Khabar K. ARED: human AU-rich element-containing mRNA database reveals an unexpectedly diverse functional repertoire of encoded proteins. Nucl. Acids Res. 2001; 29:246–254. [PubMed: 11125104]

39. Conne B, Stutz A, Vassalli JD. The 3′ untranslated region of messenger RNA: a molecular 'hotspot' for pathology? Nat. Med. 2000; 6:637–641. [PubMed: 10835679]

40. Murphy C, Beckers J, Ruther U. Regulation of the human C-reactive protein gene in transgenic mice. J. Biol. Chem. 1995; 270:704–708. [PubMed: 7822299]

41. Dean GS, Tyrrell-Price J, Crawley E, Isenberg DA. Cytokines and systemic lupus erythematosus. Ann. Rheum. Dis. 2000; 59:243–251. [PubMed: 10733469]

42. Zee RY, Ridker PM. Polymorphism in the human C-reactive protein (CRP) gene, plasma concentrations of CRP, and the risk of future arterial thrombosis. Atherosclerosis. 2002; 162:217–219. [PubMed: 11947917]

43. Brull DJ, Serrano N, Zito F, Jones L, Montgomery HE, Rumley A, Sharma P, Lowe GD, World MJ, Humphries SE, Hingorani AD. Human CRP gene polymorphism influences CRP levels. Implications for the prediction and pathogenesis of coronary heart disease. Arterioscler. Thromb. Vasc. Biol. 2003 (July Epub).

44. Tan EM, Cohen AS, Fries JF, Masi AT, McShane DJ, Rothfield NF, Schaller JG, Talal N, Winchester RJ. The 1982 revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum. 1982; 25:1271–1277. [PubMed: 7138600]

45. Sambrook, J.; Fritsch, EF.; Maniatis, T. Analysis and Cloning of Eukaryotic Genomic DNA, Molecular Cloning: a Laboratory Manual. Nolan, C.; Ford, N.; Ferguson, M., editors. Vol. vol. 9. Cold Spring Harbor Laboratory Press; Cold Spring Harbor, NY: 1989. p. 16

46. Chapman JR, Charles PJ, Venables PJW, Thompson PJ, Haslam PL, Maini RN, Turner-Warwick MEH. Definition and clinical relevance of antibodies to nuclear ribonucleoprotein and other nuclear antigens in patients with cryptogenic fibrosing alveolitis. Am. Rev. Respir. Dis. 1984; 130:439–443. [PubMed: 6433759]

47. Charles PJ, Smeenk RJT, De Jong J, Feldmann M, Maini RN. Assessment of antibodies to double-stranded DNA induced in rheumatoid arthritis patients following treatment with infliximab, a monoclonal antibody to tumour necrosis factor alpha: findings in open-label and randomized placebo-controlled trials. Arthritis Rheum. 2001; 43:2383–2390. [PubMed: 11083258]

48. Loizou S, McCrea JD, Rudge AC, Reynolds R, Boyle CC, Harris EN. Measurement of anti-cardiolipin antibodies by an enzyme-linked immunosorbent assay (ELISA): standardization and quantitation of results. Clin. Exp. Immunol. 1985; 62:738–745. [PubMed: 4085154]

49. Clayton D. A generalization of the transmission/disequilibrium test for uncertain haplotype transmission. Am. J. Hum. Genet. 1999; 65:1170–1177. [PubMed: 10486336]

50. Lewontin RC. The interaction of selection and linkage. I. General considerations; heterotic models. Genetics. 1964; 49:49–67. [PubMed: 17248194]

51. Spielman RS, McGinnis RE, Ewens WJ. Transmission test for linkage disequilibrium: the insulin region and insulin-dependent diabetes mellitus (IDDM). Am. J. Hum. Genet. 1993; 52:506–516. [PubMed: 8447318]

52. Cervino ACL, Hill AVS. Comparison of tests for association and linkage in incomplete families. Am. J. Hum. Genet. 2000; 67:120–132. [PubMed: 10841813]

53. Martin ER, Monks SA, Warren LL, Kaplan NL. A test for linkage and association in general pedigrees: The pedigree disequilibrium test. Am. J. Hum. Genet. 2000; 67:146–154. [PubMed: 10825280]

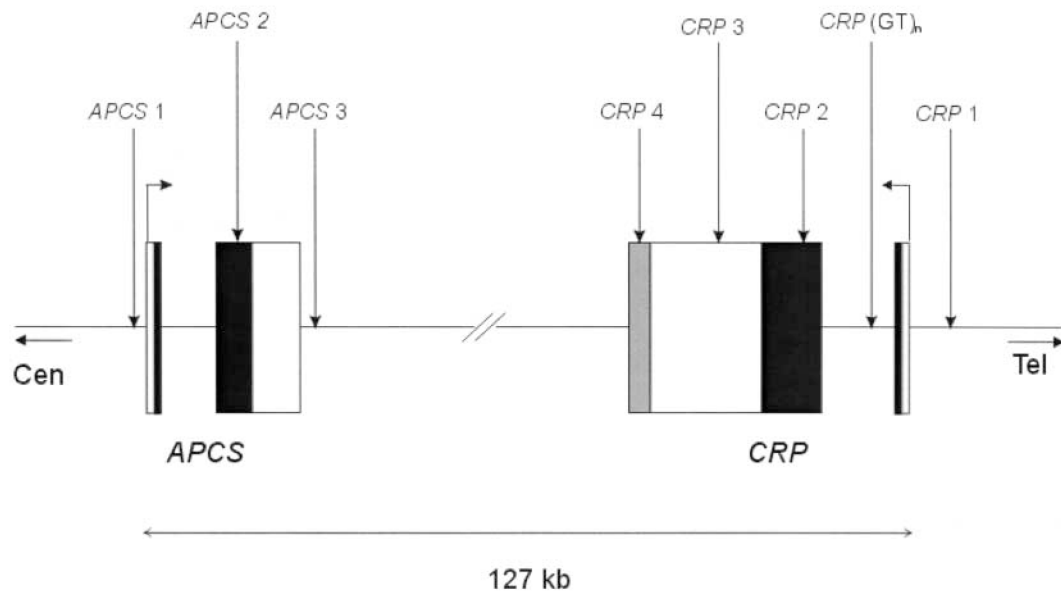54. Box GEP, Cox DR. An analysis of transformations. J. R. Stat. Soc. 1964; 26:211–252.

Figure 1.
The structure and orientation of the pentraxin genes. The pentraxin genes, *CRP* and *APCS* are paralogues sharing approximately 60% nucleotide identity. They both consist of two exons separated by a single intron. The first exon encodes the leader peptide and first two amino acids of the mature peptide and the second exon encodes the remaining 202 (SAP) and 204 (CRP) amino acids. The orientation of these genes in centromeric (Cen) and telomeric (Tel) directions is marked by arrows. Exons are represented by boxes, and introns and intergenic regions by lines. Filled boxes denote translated regions; open boxes untranslated regions. The region boxed in grey represents a portion of the gene downstream from the limit of the 3′-UTR, but which contains a possible polyadenylation signal. Arrows mark the locations of SNPs within the CRP and APCS genes. The GT dinucleotide repeat located within the intron of *CRP* is represented by *CRP*(GT)$_n$.
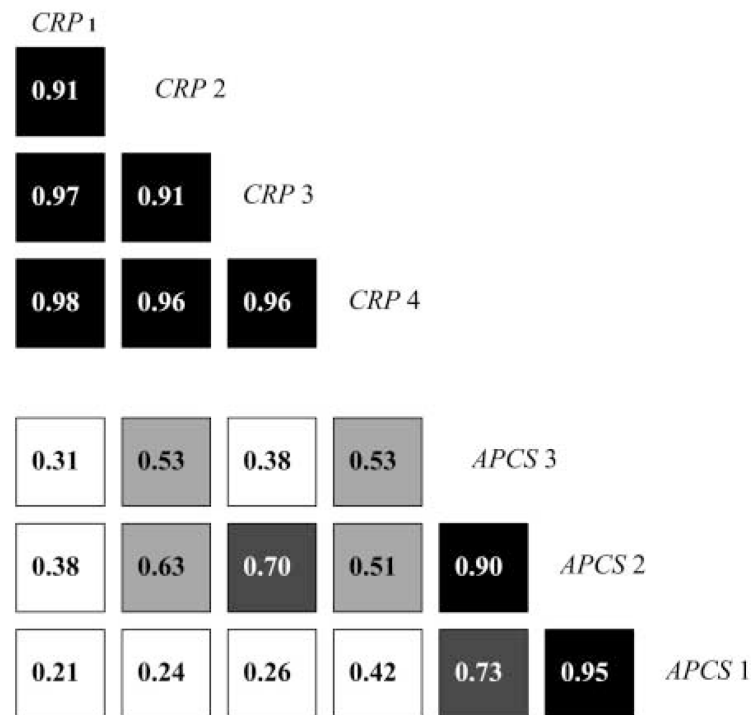
Figure 2.
LD across the pentraxin locus. The results of linkage disequilibrium mapping using GENEHUNTER 2 are presented in this matrix. The values for D' between each SNP are presented in each box. Black boxes denote strong LD ($D' > 0.8$), dark grey boxes intermediate LD (D' 0.7–0.8), pale grey boxes denote weak LD (D' 0.5–0.7) and white boxes very weak or negligible LD ($0.2 > D' < 0.5$).

Figure 3.
The association of *CRP* haplotype with basal serum CRP levels. The left-hand part of the figure shows the five common *CRP* haplotypes with corresponding allelic composition at the five markers across the *CRP* locus. To the right of this the mean serum basal CRP levels are shown for each haplotype. Significant differences ($P < 0.001$) are indicated by double asterisks.
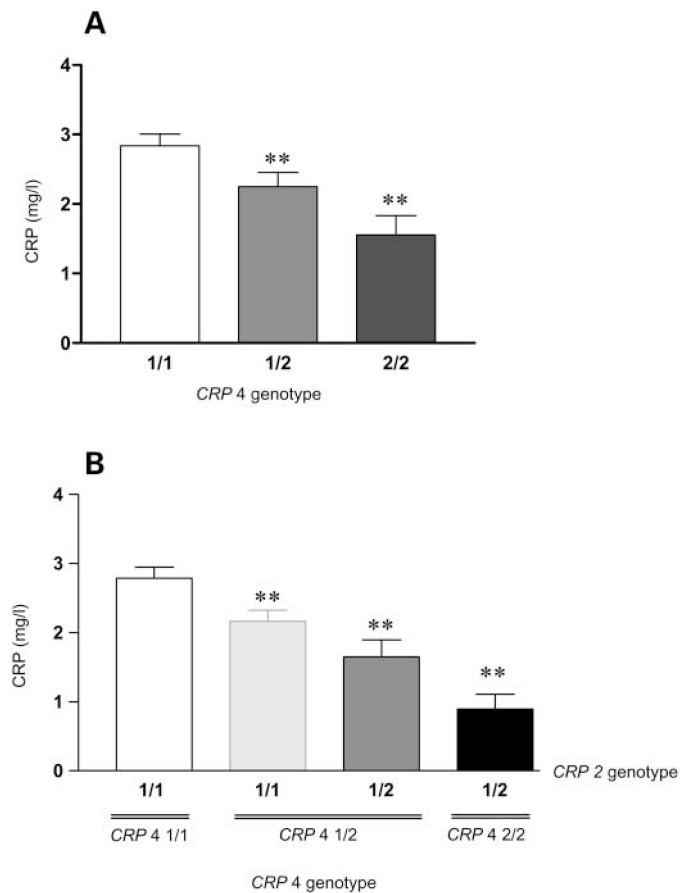
Figure 4.
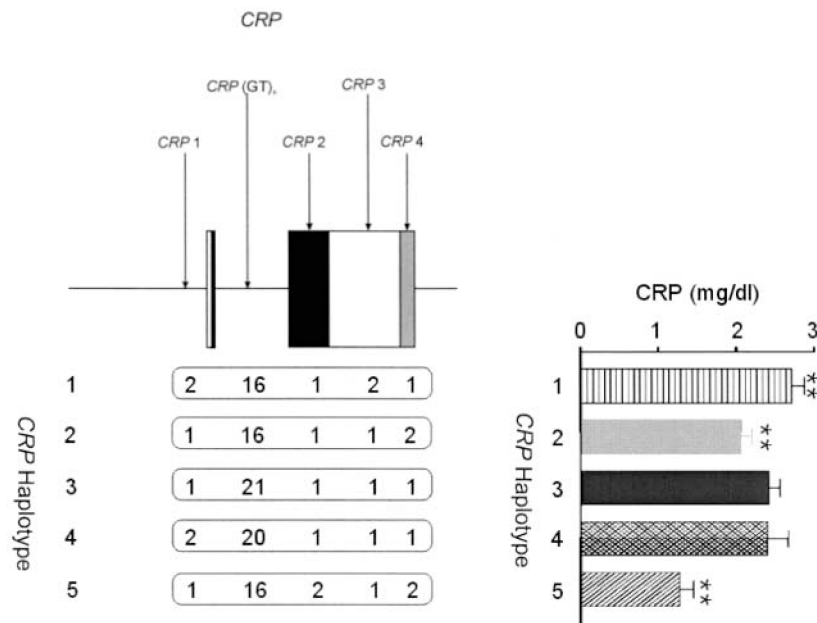(**A**) *CRP* 4 genotype influences CRP levels. This bar chart illustrates the effect on *CRP* 4 genotype on serum CRP levels in the combined analysis of Caucasian probands and siblings. Median CRP levels in the heterozygous and 2/2 homozygous groups were both significantly lower than in 1/1 homozygotes ($P < 0.0001$, double asterisk). (**B**) The interaction of *CRP* 2 and *CRP* 4 genotype on basal CRP. In this analysis Caucasian probands and siblings were divided by their *CRP 2* and *CRP 4* genotypes. The results are graphed on this bar chart. Individuals homozygous for the major (1) allele at box loci are represented by the white bar, individuals homozygous for 1 at *CRP* 2 but heterozygous at *CRP* 4 by a pale grey bar. The dark grey bar represents double heterozygotes and the black bar, those individuals homozygous for 2 at *CRP* 4, and heterozygous at *CRP* 2. There were too few double homozygotes for meaningful analysis.

Table 1

Pentraxin SNPs. SNPs were denoted numerically from 5′ to 3′ within each gene. The nucleotide position within mRNA and location within human chromosome 1 working draft sequence NT_04406 are specified. Allele frequencies for the minor alleles at each SNP are listed in the three major ethnic groups comprising our study cohorts: European-Caucasian (E), Indo-Asian (I) and Afro-Caribbean (AF)

SNP	dbSNP no.	Nucleotide Position		Position/function	Alleles	Minor allele frequency		
		mRNA	NT_04406			E	AF	I
					Total	956	54	68
CRP 1	3091244	−286	722901	5′ flanking region	C>T	0.32	0.37	0.24[a]
					C>A	0.06	0.24	0.10
CRP 2	1800947	641	724128	Synonymous codon 188	G>C	0.07	0.02	0[a]
CRP 3	1130864	988	724475	3′ untranslated region	C>T	0.31	0.16[a]	0.24
CRP 4	1205	1846	725333	3′ flanking/untranslated region	G>A	0.33	0.14[a]	0.32
APCS 1	3753868	−189	850076	5′ flanking region	G>A	0.20	0.48[a]	0.29
APCS 2	2808661	528	850793	Synonymous codon 144	G>A	0.18	0.03[a]	0.25
APCS 3		—	851049	3′ flanking region	G>T	0.08	0.02	0.03

[a]SNPs in which the allele frequency in other ethnic groups differs significantly (*P*<0.01) from that observed in Europeans.

Table 2

Length polymorphism in the *CRP* dinucleotide repeat. Allele frequency was calculated in the parental chromosomes in the major ethnic groups [European-Caucasian, Afro-Caribbean (AF), Indian-Caucasian] in our study. The alleles are named according to the number of GT repeats as described in Szalai *et al.* (31)

$(GT)_n$	Allele frequency		
	European	AF	Indian
14	0.01	0.00	0.00
15	0.01	0.01	0.01
16	0.58	0.35	0.55
17	0.03	0.09	0.04
18	0.03	0.15	0.01
19	0.00	0.07	0.00
20	0.07	0.19	0.13
21	0.24	0.10	0.25
22	0.01	0.02	0.02
23	0.01	0.02	0.00
24	0.01	0.00	0.00

Table 3

TDT of *CRP* and *APCS* polymorphisms in Cohort 1 European families. Results of single marker transmission disequilibrium analysis, using TRANSMIT, in cohort 1 (299 families). SNP alleles at a given locus are referred to as 1 (major allele) and 2 (minor allele). Although transmission of all alleles of the dinucleotide repeat was analysed, only the three most frequent *CRP* $(GT)_n$ alleles (16, 20 and 21) are presented. Values for observed and expected transmissions and the resulting chi square value (χ^2) with 1 degree of freedom are given

Polymorphism	Allele	Observed	Expected	χ^2	*P*-value
CRP 1	1	367	361	0.59	>0.05
	2	170	175		
	3	37	38		
CRP $(GT)_n$	16	345	329	3.82	0.05
	20	42	46	0.91	>0.05
	21	131	141	2.18	>0.05
CRP 2	1	532	526	1.53	>0.05
	2	36	42		
CRP 3	1	402	401	0.01	>0.05
	2	166	167		
CRP 4	1	363	380	3.81	0.05
	2	199	182		
APCS 1	1	450	449	0.01	>0.05
	2	114	115		
APCS 2	1	475	472	0.27	>0.05
	2	107	110		
APCS 3	1	518	525	2.56	>0.05
	2	52	45		

Table 4

TRANSMIT TDT analysis for haplotypes across the *CRP* genes in the European families of cohort 1. Haplotypes are numbered according to their frequency, and the allelic variants that make up each haplotype are listed from 5′ to 3′ in columns 2–6. Observed (O) and expected (E) transmissions are listed in columns 8 and 9. This analysis has been constrained to those haplotypes with frequencies of at least 5%. The *CRP* 1.16.1.1.2 haplotype (haplotype 2) was over-transmitted to affected offspring (European families $P = 0.017$, all families $P = 0.07$).

CRP Polymorphisms							Transmissions			
Haplotype	1	$(GT)_n$	2	3	4	Frequency	O	E	χ^2	*P* value
1	2	16	1	2	1	0.28	161	160	0.03	>0.05
2	1	16	1	1	2	0.24	145	129	5.61	0.017
3	1	21	1	1	1	0.23	133	136	0.15	>0.05
4	3	20	1	1	1	0.06	34	36	0.25	>0.05
5	1	16	2	1	2	0.05	28	33	1.51	>0.05

Table 5

Features of the study cohorts. Details of the demographic features of the 45 SLE patients whose DNA was used for SNP screening (SNP screening cohort) and the two cohorts of SLE families examined in this study. Glomerulonephritis refers to the documented involvement of the kidneys in the proband

		SNP screen	Cohort 1	Cohort 2
Total		45	344	242
Family structure	Nuclear simplex	—	229	175
	Single parent simplex	—	115	67
Ethnicity	European Caucasian	14	299	197
	Afro-Caribbean	14	18	14
	Indo-Asian	14	17	13
	Oriental	3	3	3
	Mixed ethnicity	—	7	5
Glomerulonephritis	Lupus nephritis	23	128	46
	No lupus nephritis	22	216	196
Gender of proband	Male	3	24	22
	Female	42	320	220

Table 6

PCR-RFLP assays employed to genotype SNPs in the pentraxin genes. The oligonucleotide primer sequences used for each assay and the discriminating restriction enzyme are given in this table

SNP	Forward primer	Reverse primer	Enzyme	Inhibitory allele		Permissive allele	
				Allele	Fragment	Allele	Fragment
CRP 1	CATCTCCAGAAGCTGTCAG	TCAGAGCTTCGGGAAGAG	*Hin*4I	T, A	287	C	86 + 201
			*Spe*I	T, C	287	A	95 + 182
CRP 2	TCAGGGATCGTGGAGTTC	GCCTTGCACTTCATACTTC	*Tsp*45I	C	279	G	90 + 189
CRP 3	GTCTCTGGGCCTTTGTTC	TAGAAGATCAGCGCTTCCT	*Hpy*CH4III	T	290	C	111 + 179
CRP 4	GGAGTGAGACATCTTCTTG	CTTATAGACCTGGGCAGT	*Hpy*CH4III	A	227	G	97 + 130
APCS 1	CCATTGATTTTCCAGCTCAG	AGGAAGAGCACCAATAATATGTG	*Mnl*I	G	91	A	41 + 50
APCS 2	GGAGTCCTCATCAGGTATTGCTG	CTGGCAATACGGCAGCTAATCAGG	*Fok*I	A	106	G	24 + 82