Contribution of Clinical Correlates and 13 C-Reactive Protein Gene Polymorphisms to Interindividual Variability in Serum C-Reactive Protein Level

Sekar Kathiresan, MD; Martin G. Larson, ScD; Ramachandran S. Vasan, MD; Chao-Yu Guo, PhD; Philimon Gona, PhD; John F. Keaney, Jr, MD; Peter W.F. Wilson, MD; Christopher Newton-Cheh, MD, MPH; Stacy L. Musone, BA; Amy L. Camargo, BA; Jared A. Drake, BA; Daniel Levy, MD; Christopher J. O’Donnell, MD, MPH; Joel N. Hirschhorn, MD, PhD; Emelia J. Benjamin, MD, ScM

Background—Serum C-reactive protein (CRP) level is a heritable complex trait that predicts incident cardiovascular disease. We investigated the clinical and genetic sources of interindividual variability in serum CRP.

Methods and Results—We studied serum CRP in 3301 Framingham Heart Study (FHS) participants (mean age 61 years, 53% women). Twelve clinical covariates explained 26% of the variability in CRP level, with body mass index alone explaining 15% (P<0.0001) of the variance. To investigate the influence of genetic variation at the CRP gene on CRP levels, we first constructed a dense linkage disequilibrium map for common single-nucleotide polymorphisms (SNPs) spanning the CRP locus (1 SNP every 850 bases, 26 kilobase [kb] genomic region). Thirteen CRP SNPs were genotyped in 1640 unrelated FHS participants with measured CRP levels. After adjustment for clinical covariates, 9 of 13 SNPs were associated with CRP level (P<0.05). To account for correlation among SNPs, we conducted forward stepwise selection among all 13 SNPs; a triallelic SNP (rs3091244) remained associated with CRP level (stepwise P<0.0001). The triallelic SNP (C→T→A; allele frequencies 62%, 31%, and 7%), located in the promoter sequence, explained 1.4% of total serum CRP variation; haplotypes harboring the minor T and A alleles of this SNP were associated with higher CRP level (haplotype P=0.0002 and 0.004).

Conclusions—In our community-based sample, clinical variables explained 26% of the interindvidual variation in CRP, whereas a common triallelic CRP SNP contributed modestly. Studies of larger samples are warranted to assess the association of genetic variation in CRP and risk of cardiovascular disease. (Circulation. 2006;113:1415-1423.)

Key Words: epidemiology ■ inflammation ■ genetics ■ C-reactive protein ■ risk factors

C-reactive protein (CRP) is a marker of systemic inflammation, and circulating CRP levels predict incident cardiovascular disease (CVD). Current evidence suggests that serum CRP level is a complex trait, influenced by both clinical and genetic factors. Clinical and environmental correlates of CRP reported in the literature include age, sex, smoking, body mass index, hormone replacement therapy, and statin use, among others. In addition, heritability estimates for CRP levels have ranged from 27% to 40%, which suggests a role for DNA sequence variation in determining serum protein level. However, the extent to which clinical and genetic variables together account for the total variation in serum CRP level is unclear.

Clinical Perspective p 1423

Recent advances in genomics allow for the comprehensive evaluation of candidate gene regions for roles in determining phenotypic variation. There are an estimated 11 million single-nucleotide polymorphisms (SNPs) with a minor allele frequency >1%, and it recently has been recognized that groups of neighboring sequence variants are correlated, a pattern termed “linkage disequilibrium” (LD). A gene-
mapping approach that takes advantage of LD offers promise in mapping genes for complex traits. Specifically, such an approach involves characterization of sequence variation in a reference sample, selection of tag SNPs (“tag” refers to a subset of SNPs that capture the underlying genetic variation in a reference panel and are taken forward to genotype in an association study), and testing of tag SNPs for association with phenotype.

CRP sequence variants that explain interindividual variability in CRP level may aid in the prediction of clinical CVD, may provide targets for drugs to alter gene expression, and represent strong candidate alleles to test for association with CVD. We hypothesized that clinical variables are related to serum CRP levels; that a limited number of tag polymorphisms capture the majority of common genetic variation present at the CRP locus; that after accounting for clinical variables, tag polymorphisms in the CRP gene are associated with serum CRP level; and that CRP polymorphisms related to serum CRP level will be associated with risk of CVD. Accordingly, we investigated LD structure and tag SNP selection at the CRP locus in a reference DNA panel and the clinical and genetic correlates of serum CRP level in Framingham Heart Study (FHS) participants.

Methods

Study Participants

The Framingham Offspring cohort (n=5124 participants) was recruited in 1971; the participants have been examined approximately every 4 to 8 years. Of the 3539 Offspring participants who attended the seventh examination (1998–2001), 3301 individuals were available for the analyses of clinical correlates of CRP, after the exclusion of participants with missing CRP level (n=238). DNA was available in 1809 unrelated individuals who provided blood samples for DNA extraction during the sixth examination cycle (1995–1998). The genotype-phenotype association sample consisted of the overlap of participants with available CRP level and available DNA (n=1640). The Institutional Review Board at Boston Medical Center approved the study, and all participants gave written informed consent.

Clinical Examination and CRP Testing

Risk factor information was collected from routine medical history, physical examination, and laboratory assessment as described previously (please see online-only Data Supplement for risk factor definitions). CVD was defined by the presence of coronary heart disease (recognized or unrecognized myocardial infarction, coronary insufficiency, angina pectoris, or coronary death), stroke, transient ischemic attack, intermittent claudication, or heart failure. CVD diagnoses were determined by a panel of 3 investigators who reviewed all available medical and hospital records, as described previously. CRP level was measured once for each participant with a Dade Behring BN100 nephelometer (Deerfield, Ill) on fasting morning serum samples that had been stored at −70°C and thawed at room temperature. The mean intra-assay coefficient of variation was 3.2%.

SNP Selection and Genotyping in a Reference Panel

To construct a dense SNP map at the CRP locus, we searched the dbSNP database (http://www.ncbi.nlm.nih.gov/SNP) as of June 2004 and selected every SNP within a 26-kilobase (kb) region spanning the CRP gene (2-kb coding region; 15 kb upstream and 9 kb downstream) on chromosome 1 (gene symbol CRP, accession number NM_000567). Ninety-nine SNPs were genotyped in a panel of 30 white trios (90 individuals total, with each trio consisting of 2 parents and an offspring) from Utah/Centre d’Etude du Polymorphisme Humain (CEPH) panel (Coriell Institute for Medical Research, Camden, NJ). Genotyping was performed with matrix-assisted laser desorption ionization–time of flight mass spectrometry in the Sequenom MassARRAY platform (San Diego, Calif) as described previously.

Assays were considered successful if they met the following criteria: (1) at least 75% success for genotyping calls; (2) Hardy-Weinberg equilibrium P>0.01; and (3) Mendelian transmission errors ≤1. In addition, we imposed a minor allele frequency threshold and defined “common” for the present study as a minor allele frequency ≥5%. Overall, 30 SNPs passed quality control and had a minor allele frequency ≥5% (Data Supplement Table I).

LD Structure in Reference Pedigrees

LD structure was assessed by calculating D’ and a logarithm of the odds score for each pair of SNPs as implemented in the publicly available Haploviz version 2.03 software. Haplotype blocks were defined by the criteria of Gabriel et al.

Tag SNP Selection

We used 3 approaches to select tag SNPs: (1) a pairwise correlation method; (2) R2, a method of selecting haplotype-tag SNPs to predict common haplotypes; and (3) Rs, a method of selecting haplotype-tag SNPs to predict unmeasured SNPs (see Data Supplement for details).

In each of the 3 methods, 7 tag SNPs were required to capture the underlying common genetic variation. These 3 sets of 7 SNPs partially overlapped and yielded a total of 10 tag SNPs (Data Supplement Table I). In addition, we genotyped 3 previously studied SNPs: rs1417938, rs1130864, and rs3091244. Overall, 13 CRP SNPs were genotyped in the FHS sample. All SNPs genotyped were in Hardy-Weinberg equilibrium (χ2 test P>0.05).

Statistical Analysis

Clinical Correlates

Serum CRP levels had a skewed distribution and were natural log transformed. We constructed stepwise multivariable linear regression models (SAS REG procedure) to identify clinical covariates associated with log CRP, with age and sex forced in as covariates in all models. Additional candidate covariates are detailed in Table 2. P<0.05 was the significance criterion for covariates to be retained in the forward stepwise-selection models.

Genotype–CRP Level Association

Log CRP values were adjusted for 15 covariates to create a multivariable-adjusted CRP residual level for each participant. A participant’s multivariable-adjusted CRP level served as the sole phenotype in genotype-phenotype and haplotype-phenotype association analyses.

For genotype-phenotype association analyses, we assumed a general model of inheritance (3 genotype classes by biallelic SNP and 6 genotype classes for triallelic SNP rs3091244). We conducted multivariable linear regression analyses (SAS Proc GLM) to test the null hypothesis that log CRP residuals did not differ by SNP genotype. To identify a subset of SNPs that significantly explained variance in CRP levels when adjusted for the effects of other SNPs, we conducted forward stepwise selection of the SNPs using a P<0.05 criterion. In exploratory analyses, we evaluated specific interaction terms in the multivariable-adjusted model, including an SNP×covariate interaction term for each clinical covariate significantly related to serum CRP level. Haplotype-based association analyses were conducted with weighted regression as implemented in the haplo.stats program.

Genotype–Cardiovascular Disease Association

Because we identified a single CRP polymorphism (triaallelic SNP rs30921244) that strongly related to CRP level, in a post hoc analysis, we hypothesized that this sequence variant was related to either prevalent or incident CVD among participants. Multivariable logistic
**TABLE 1. Framingham Study Sample Characteristics**

<table>
<thead>
<tr>
<th>Clinical Correlates Sample</th>
<th>Genetic Association Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Men (n=1536)</td>
</tr>
<tr>
<td>Age, y</td>
<td>61±10</td>
</tr>
<tr>
<td>Systolic blood pressure, mm Hg</td>
<td>128±17</td>
</tr>
<tr>
<td>Diastolic blood pressure, mm Hg</td>
<td>76±10</td>
</tr>
<tr>
<td>Body mass index, kg/m²</td>
<td>28.8±4.6</td>
</tr>
<tr>
<td>Total/HDL cholesterol ratio</td>
<td>4.5±1.4</td>
</tr>
<tr>
<td>Cigarette smoking, %</td>
<td>13</td>
</tr>
<tr>
<td>Fasting glucose, mg/dL</td>
<td>109±29</td>
</tr>
<tr>
<td>Triglycerides, mg/dL</td>
<td>143±100</td>
</tr>
<tr>
<td>Diabetes, %</td>
<td>17</td>
</tr>
<tr>
<td>Hypertension treatment, %</td>
<td>37</td>
</tr>
<tr>
<td>Lipid-lowering medication, %</td>
<td>25</td>
</tr>
<tr>
<td>Hormone replacement therapy, %</td>
<td>...</td>
</tr>
<tr>
<td>Aspirin treatment, %</td>
<td>43</td>
</tr>
<tr>
<td>Prevalent CVD, %</td>
<td>18</td>
</tr>
<tr>
<td>CRP, mg/L, median (Q1, Q3)</td>
<td>1.9 (0.9, 4.4)</td>
</tr>
</tbody>
</table>

Q1 indicates quartile 1; Q3, quartile 3.
Continuous measures, mean±SD, except CRP, for which median (1st and 3rd quartile) values are presented.

regression analysis was performed to test the null hypothesis that prevalent CVD status (n=210 at examination 7, 1998 to 2001) did not differ by marker genotype, and multivariable Cox proportional hazards modeling was performed to test the null hypothesis that incident CVD (n=52 at postexamination 7 to December 2004) did not differ by marker genotype (see Data Supplement for details of genetic statistical analysis). For all analyses, a nominal \( P<0.05 \) was considered significant.

The authors had full access to the data and take responsibility for its integrity. All authors have read and agree to the manuscript as written.

**Results**

**Framingham Heart Study Sample Characteristics**
Clinical characteristics of the FHS sample (mean age 61 years, 53% women) are shown in Table 1. The genotype-phenotype association study sample (n=1640) was similar to the sample with CRP levels.

**Clinical Correlates of Serum CRP Levels**
In stepwise multivariable linear regression models that included age and sex, 10 clinical factors were significantly related to serum CRP levels (Table 2). Clinical correlates explained 26% of the interindividual variability in serum CRP level, with body mass index alone explaining 15% of the total variance.

**LD Pattern at the CRP Locus in a Reference Panel**
In 90 CEPH individuals (30 trios) of European ancestry, we constructed a dense SNP map (1 SNP every 830 bases) at the CRP locus (genomic distance 26 kb, 15 kb upstream, 2-kb gene length, 9 kb downstream) using 30 common SNPs (minor allele frequency \( \geq 5% \); Data Supplement Table 1). We observed 1 extended haplotype block of 16 kb in this region (Figure). There was limited haplotypic diversity, with 7 haplotypes with \( \geq 5% \) frequency accounting for 94% of chromosomes.

**Performance of Tag SNPs**
We evaluated the performance of the tag SNPs genotyped in the present study by comparing the correlation between our tag SNPs and CRP sequence variation described in the SeattleSNPs Program for Genomics Applications (http://pga.gs.washington.edu/). In the SeattleSNPs project, 23 individuals of European ancestry were resequenced for 6.8 kb spanning the CRP gene. Ten biallelic SNPs were observed with a minor allele frequency \( \geq 5% \). This SeattleSNPs collection should represent complete ascertainment for common SNPs in 6.8 kb of genomic distance spanning the CRP gene. Six of the biallelic tag SNPs typed in the present study perfectly captured the 10 SeattleSNPs variants (mean \( r^2=1.0 \)).

**Associations of Single SNPs With Multivariable-Adjusted Serum CRP Level**
We examined the relations between each of 13 SNPs at the CRP locus and multivariable-adjusted serum CRP residual; 9 SNPs were nominally associated (each \( P<0.05 \), minimum \( P<0.0001 \); Table 3). Several SNPs were highly correlated with each other. For example, SNP rs3116653 and rs1417938 were near-perfect proxies for each other (mean \( r^2=0.98 \)).

To account for correlation among SNPs, we conducted stepwise selection among all 13 SNPs and found that triallelic SNP rs3091244 remained associated with CRP level (stepwise \( P<0.0001 \)). This triallelic SNP was highly correlated with haplotypes composed of 2 other SNPs (rs1205 and rs2808630). In stepwise models that excluded the triallelic SNP, these 2 SNPs, rs1205 and rs2808630, were also separately associated with CRP level (\( P<0.0001 \) and 0.003,
TABLE 5). Compared with all other haplotypes combined, 8 common haplotypes observed in FHS (global
Mean serum multivariable-adjusted CRP level differed across
Multivariable-Adjusted Serum CRP Level
Associations of Haplotypes With
variation in CRP level (ie, \(1.8\%\)). The triallelic SNP (as represented by the
TRIPLET SNP) was our best association result at the
CRP locus, we studied whether genetic variants highly
correlated to the triallelic SNP existed beyond the 26-kb
genomic segment surveyed.\textsuperscript{21} Using the HapMap catalog of
>3.5 million SNPs in the CEPH sample of European ancestry,\textsuperscript{22} we
assessed a 2-megabase genomic segment encompassing the
CRP locus and found that there were no perfect
proxies (\(r^2=1.0\)) for the triallelic SNP (as represented by the
A-G haplotype at rs2808630 and rs1205). Seven variants in
this 2-megabase region were correlated with the A-G haplotype
to a lesser degree (each \(r^2=0.74\)).

Triallelic SNP and Proxies in the International
Haplotype Map

Because triallelic SNP rs3091244 was strongly related to
CRP level, and CRP level has been consistently related to
CVD, we performed a post hoc analysis to evaluate the
relations between SNP rs3091244 genotype and prevalent (at
the time of FHS Offspring Study examination cycle 7) or
incident (after examination 7 to December 2004) CVD. SNP
rs3091044 was not associated with prevalent CVD (n = 210;
\(P=0.62\) in multivariable-adjusted model). Compared with the
major allele homozygotes (CC) as referent genotype, the OR
for the group that included major allele heterozygotes (CT
and CA) was 0.96 (95% CI 0.67 to 1.37), and for the group
with both minor alleles (TA, TT, and AA), it was 0.79 (95% CI
0.48 to 1.39). SNP rs3091044 was not associated with incident CVD (n = 52; \(P=0.18\) in
multivariable-adjusted model). Compared with the major allele homozygotes (CC)
as referent genotype, the hazard ratio for the group that
included major allele heterozygotes (CT and CA) was 0.82
respectively). The triallelic SNP (C→T→A, allele frequencies 62%, 31%, and 7%), located in the promoter sequence, explained 1.8% of the residual variation in a model that included all clinical covariates (Table 3). Because clinical covariates explained 26% of serum CRP variation, this 1.8% of variation in CRP residual represents 1.4% of the total variation in CRP level (ie, \([1−0.26]×1.8\%\)).

CRP level by genotype at the triallelic SNP rs3091244 is
presented in Table 4. Unadjusted values are presented for
ease of interpretation. Major allele homozygotes (CC geno-
type) had the lowest CRP level. Major allele heterozygotes
(CT or CA genotypes) had an intermediate level of CRP,
whereas individuals with only the minor T or A alleles had
the highest CRP level. In exploratory analyses, there were
nominal significant SNP×covariate interaction terms for the
following: hypertension treatment×rs3116654, triglycerides×rs1417938, triglycerides×rs1130864, and body mass
index×rs1205 (each interaction term \(P=0.04\); Data Supple-
mantent Table II).

Associations of Haplotypes With
Multivariable-Adjusted Serum CRP Level
Mean serum multivariable-adjusted CRP level differed across
8 common haplotypes observed in FHS (global \(P<0.0001\);
Table 5). Compared with all other haplotypes combined,
haplotypes A and B were associated with a higher CRP level
(haplotype \(P=0.0002\) and 0.004, respectively). Haplotype A
harbored the minor T allele, and haplotype B harbored the
minor A allele of the triallelic SNP. A 2-SNP combination
(rs2808630 and rs1205) served as a proxy for the triallelic
SNP and also distinguished haplotypes A and B from all
others (Table 5).
(95% CI 0.46 to 1.45), whereas it was 0.31 (95% CI, 0.09 to 1.07) for the group with both minor alleles (TA, TT, and AA).

**Statistical Power**

Because our sample included a modest number of CVD events (210 prevalent, 52 incident), we assessed power to detect an association between CRP SNP genotype and CVD using separate analyses of prevalence and incidence but combining results into a single, combined test (4 degrees of freedom, ie, 2 degrees of freedom each for the prevalence and incidence analyses). At the 5% significance level, we had power of 0.55 for an SNP genotype with an OR of 1.3 per risk allele (that is, OR 1.00, 1.30, and 1.69 for carriers of 0, 1 and 2 risk alleles), power of 0.80 for an OR of 1.40, and power of 0.94 for an OR of 1.5 per risk allele, assuming a risk allele frequency of 30% (the T allele of triallelic SNP rs3091244 has an allele frequency of 31%).

**Discussion**

**Principal Findings**

In a community-based sample, we sought to define the contribution of clinical correlates and 13 CRP gene polymorphisms to interindividual variation in serum CRP level. The present study yielded several findings: (1) in a reference sample of European ancestry, a single block of LD stretched across the CRP locus, and a limited number of tag polymorphisms captured the majority of common sequence variation; (2) the proportion of CRP variation explained by clinical variables was \( \approx 26\% \), and for CRP SNPs, it was \( \approx 1.4\% \); (3) of 13 CRP common sequence variants studied, a triallelic SNP in the promoter sequence was the strongest cis-acting variant associated with CRP level; (4) there were no perfect proxies for the triallelic SNP across a 2-megabase region that spanned the CRP locus\(^2\); and (5) the CRP triallelic SNP was not associated with CVD in analyses with
modest power to detect a genetic effect. The triallelic SNP association with CRP level strongly supports recent work described below and represents a reproducible association between genetic variation and this biomarker of high interest.

**Clinical Correlates of Serum CRP**

We confirmed prior reports on individual clinical correlates of CRP including age, sex, cardiovascular risk factors (body mass index, cigarette smoking, total/HDL cholesterol ratio, triglycerides, systolic blood pressure, and diastolic blood pressure), medications (lipid-lowering therapy, hormone replacement therapy, and prevalent CVD). We estimated the relative contribution of each of these factors to CRP level variation in the community and demonstrated that body mass index was by far the strongest correlate, explaining 15% of the overall variation in CRP. A possible mechanism behind the adiposity and CRP association has been suggested; adipocytes synthesize a substantial proportion of circulating interleukin-6, the principal regulator of CRP synthesis in the liver.

**Comparison of SNP Associations With Prior Studies**

Several prior studies have examined the relations between CRP polymorphisms and CRP levels and reported individual associations between rs1800947 (also referred to as 1059T>G) and prevalent CVD. We estimated the relative contribution of each of these factors to CRP level variation in the community and demonstrated that body mass index was by far the strongest correlate, explaining 15% of the overall variation in CRP. A possible mechanism behind the adiposity and CRP association has been suggested; adipocytes synthesize a substantial proportion of circulating interleukin-6, the principal regulator of CRP synthesis in the liver.

**TABLE 3.** Association of Individual CRP Gene Variants and Multivariable-Adjusted Serum CRP Level* in FHS

<table>
<thead>
<tr>
<th>Allelic Variant</th>
<th>Variant Type</th>
<th>Major Allele→Minor Allele Frequency</th>
<th>Partial R²†</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs3116653</td>
<td>5’ Upstream</td>
<td>C→G</td>
<td>0.31</td>
<td>0.009</td>
</tr>
<tr>
<td>rs3116654</td>
<td>5’ Upstream</td>
<td>T→C</td>
<td>0.12</td>
<td>0.004</td>
</tr>
<tr>
<td>rs2794517</td>
<td>5’ Upstream</td>
<td>C→T</td>
<td>0.30</td>
<td>†</td>
</tr>
<tr>
<td>rs3122012</td>
<td>5’ Upstream</td>
<td>A→G</td>
<td>0.31</td>
<td>0.008</td>
</tr>
<tr>
<td>rs3091244</td>
<td>5’ Upstream</td>
<td>C→T</td>
<td>0.31</td>
<td>0.018</td>
</tr>
<tr>
<td>rs1417938</td>
<td>Intron</td>
<td>A→T</td>
<td>0.31</td>
<td>0.008</td>
</tr>
<tr>
<td>rs1800947</td>
<td>Synonymous, coding</td>
<td>G→C</td>
<td>0.06</td>
<td>†</td>
</tr>
<tr>
<td>rs1130864</td>
<td>3’ Untranslated</td>
<td>C→T</td>
<td>0.31</td>
<td>0.009</td>
</tr>
<tr>
<td>rs1205</td>
<td>3’ Flanking</td>
<td>G→A</td>
<td>0.33</td>
<td>0.016</td>
</tr>
<tr>
<td>rs2808630</td>
<td>3’ Flanking</td>
<td>A→G</td>
<td>0.28</td>
<td>†</td>
</tr>
<tr>
<td>rs3093077</td>
<td>3’ Flanking</td>
<td>T→G</td>
<td>0.07</td>
<td>0.008</td>
</tr>
<tr>
<td>rs12029262</td>
<td>3’ Flanking</td>
<td>C→G</td>
<td>0.09</td>
<td>0.007</td>
</tr>
<tr>
<td>rs876538</td>
<td>3’ Flanking</td>
<td>G→A</td>
<td>0.20</td>
<td>†</td>
</tr>
</tbody>
</table>

*Phenotype was log-transformed serum CRP level adjusted for age, sex, and 13 candidate covariates listed in legend to Table 2. †Partial R² refers to the proportion of multivariable-adjusted CRP level explained by the specific SNP.

**TABLE 4.** CRP Level by Genotype for 6 Genotypes at Triallelic SNP rs3091244

<table>
<thead>
<tr>
<th>Allelic Variant</th>
<th>Frequency</th>
<th>Partial R²†</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs3116653</td>
<td>0.31</td>
<td>0.009</td>
<td>0.0008</td>
</tr>
<tr>
<td>rs3116654</td>
<td>0.12</td>
<td>0.004</td>
<td>0.04</td>
</tr>
<tr>
<td>rs2794517</td>
<td>0.30</td>
<td>†</td>
<td>0.99</td>
</tr>
<tr>
<td>rs3122012</td>
<td>0.31</td>
<td>0.008</td>
<td>0.001</td>
</tr>
<tr>
<td>rs3091244</td>
<td>0.31</td>
<td>0.018</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>rs1417938</td>
<td>0.31</td>
<td>0.008</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Q1 indicates quartile 1; Q3, quartile 3. Unadjusted values are presented for reference. Note that association analyses were conducted with multivariable-adjusted log-transformed CRP level as the phenotype.
TABLE 5. Association of CRP Locus Haplotypes With Multivariable-Adjusted Serum CRP Level* in FHS

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Frequency</th>
<th>Score</th>
<th>P</th>
<th>Global P</th>
</tr>
</thead>
<tbody>
<tr>
<td>A G C T A G</td>
<td>0.29</td>
<td>3.74</td>
<td>0.0002</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>B G C G A G C G A A A C T G T C</td>
<td>0.05</td>
<td>2.86</td>
<td>0.004</td>
<td></td>
</tr>
<tr>
<td>C A C T G G C G A A C A T T C</td>
<td>0.19</td>
<td>-0.28</td>
<td>0.78</td>
<td></td>
</tr>
<tr>
<td>D G C T G G C G A C A T T C</td>
<td>0.08</td>
<td>-0.91</td>
<td>0.36</td>
<td></td>
</tr>
<tr>
<td>E G C T A A C C A C A C C T C</td>
<td>0.05</td>
<td>-1.98</td>
<td>0.048</td>
<td></td>
</tr>
<tr>
<td>F G G T A A C G A C A C C T C</td>
<td>0.09</td>
<td>-2.21</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>G G C T A A C C G A C A C C C</td>
<td>0.12</td>
<td>-2.34</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>H G C T A A C C G A C A C C T C</td>
<td>0.05</td>
<td>-2.78</td>
<td>0.005</td>
<td></td>
</tr>
</tbody>
</table>

*Phenotype was log-transformed serum CRP level adjusted for age, sex, and 13 candidate covariates listed in legend Table 2.
†Haplotype score statistic tests whether trait differences exist between a single haplotype and all other haplotypes combined.
‡Note that in stepwise selection including all 13 CRP SNP variants, triallelic SNP rs3091244 or a 2-SNP proxy (rs2808630 and rs1205) was significantly associated with multivariable-adjusted CRP level. The 2 minor alleles of the triallelic SNP (T and A) or the A-G haplotype at rs2808630-rs1205 distinguishes haplotype A and haplotype B from all others.

Triallelic SNP Function

Several lines of evidence strongly suggest that the triallelic SNP rs3091244 is functional. First, prior in vitro studies have directly demonstrated a functional role for SNP rs3091244. Szalai et al reported that SNP rs3091244 resides in a consensus E-box element (ie, a DNA sequence that binds basic helix-loop-helix transcription factors such as upstream stimulatory factor-1) and that transcription factor binding occurred when the variant T allele was present. They went on to demonstrate that haplotypes harboring the variant T or A alleles were associated with higher promoter activity, as assessed by luciferase reporter assays, consistent with our findings in the present study. Second, SNP rs3091244 is located in promoter sequence, 301 bases upstream of the CRP gene transcription start site. Third, Cha-Molstad and colleagues have shown that an interplay of transcription factors binding to the proximal promoter segment regulates CRP gene expression. Lastly, we observed that the region surrounding this SNP represents a nonexonic evolutionarily conserved region, which suggests functional importance (http://ecrbrowser.dcode.org/). This evolutionarily conserved region represents a 75% sequence identity compared with murine sequence (www.ecrbrowser.com), extends for 287 bases, and encompasses SNP rs3091244.

Association and Linkage for Complex Quantitative Traits

In a genome-wide linkage scan that we recently conducted for the quantitative trait of CRP level, we did not observe evidence for linkage at the CRP locus on chromosome 1. The apparent discrepancy highlights recent concepts raised by researchers comparing association studies with linkage analysis. Association approaches offer greater statistical power to detect common alleles with modest genetic effects. Indeed, for a range of blood protein level phenotypes including plasminogen activator inhibitor-1 and monocyte chemoattractant protein-1, we and others have shown that SNPs in the gene encoding the respective blood protein explain, on average, 1% to 2.5% of phenotypic variation. Association studies have greater power than linkage studies in equivalently sized cohorts to detect such modest effect sizes.

Study Limitations and Strengths

The large community-based sample, comprehensive characterization of common sequence variation surrounding the CRP gene, unbiased ascertainment of measured covariates, multivariable analyses, and the combination of SNP-phenotype and haplotype-phenotype analyses are strengths of the present investigation. Despite these assets, our study has some limitations. First, we addressed the role of common CRP genetic variation in relation to CRP level. Multiple rare variants have been shown to influence quantitative cardiovascular traits such as HDL cholesterol. Resequencing of the CRP gene locus in individuals with extreme phenotypes will be required to address the possibility that aggregation of multiple rare variants contributes to interindividual variation in CRP level in the general population. Second, regulatory variants acting in trans (ie, sequence variants outside of the CRP locus) likely affect serum CRP level. The present study did not address this hypothesis. Third, because our genotype variant-phenotype interaction analyses were exploratory and only nominally significant, they need to be confirmed in another cohort.

Finally, our sample consisted of a modest number of CVD events, and thus, the lack of association between the CRP triallelic SNP genotype and CVD may be due to insufficient statistical power. Our estimates suggest that we had >90% power to detect an OR of 1.5 per risk allele if we assumed a risk allele frequency of 30%. Studies with larger number of individuals affected with CVD are warranted to assess for more modest effect sizes, and such efforts are currently in progress.

Implications

Our findings may have several implications. First, body mass index is by far the strongest correlate of CRP levels, and this finding suggests that maintaining a healthy weight may be the most effective measure to keep serum CRP levels low. Second, genetic association studies that use blood protein levels as phenotypes may effectively identify functional
noncoding alleles in the genome. Finally, genetic association studies using the finite set of CRP tag SNPs described in the present study could be used to clarify whether CRP is causally related to clinical disease. Whereas serum CRP level has been shown to predict a range of clinical phenotypes, including myocardial infarction, stroke, diabetes, and hypertension, it is debated whether serum CRP level marks an underlying inflammatory process or causally contributes to that clinical disease. A bona fide association between a CRP gene variant and clinical disease would suggest a causal role for CRP.

Acknowledgments
This study was supported by the National Heart, Lung, and Blood Institute (N01-HC-25195); CardioGenomics Programs for Genomic Applications (HL66582); R01HL076784 and R01HL064753 (Dr Benjamin); and 2K24HL004334 (Dr Vasan). D. K. Kathiresan is supported by the ACC Foundation/Merck Adult Cardiology Research Award. The authors acknowledge Itsik Pe’er, PhD, for his assistance and analyses of HapMap data.

Disclosures
None.

References
C-reactive protein (CRP) is a marker of systemic inflammation, and circulating CRP levels predict incident cardiovascular disease. Circulating CRP levels vary considerably between individuals, and at least part of this variability has been demonstrated to be due to genetic factors. In a community-based sample of 3301 individuals, we evaluated the extent to which clinical and genetic factors explained serum CRP level variability. Twelve clinical factors were significantly related to serum CRP level and combined explained ~26% of CRP level variation. Of these factors, body mass index alone was the strongest explanatory factor, contributing to 15% of the variance. To isolate genetic factors related to CRP level, we studied 13 common polymorphisms that represented the genetic variation at the CRP gene region and found that a triallelic polymorphism (most single-nucleotide polymorphisms have 2 alleles, whereas a few, such as that in the present study, have 3 alleles) was strongly related to CRP level. The triallelic CRP polymorphism explained ~1.4% of the overall variability in CRP level. Using a comprehensive approach that went beyond single polymorphisms, we were able to fully characterize the association of CRP genetic variation on circulating CRP level. The specific triallelic CRP SNP identified in the present study is of interest for future studies to assess whether this polymorphism potentially may aid in clinical disease risk prediction and in targeting therapeutic interventions. Our findings imply that maintaining a healthy weight may be the most effective measure to keep serum CRP level low.
 Contribution of Clinical Correlates and 13 C-Reactive Protein Gene Polymorphisms to Interindividual Variability in Serum C-Reactive Protein Level


_Circulation_. 2006;113:1415-1423; originally published online March 13, 2006; doi: 10.1161/CIRCULATIONAHA.105.591271

_Circulation_ is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2006 American Heart Association, Inc. All rights reserved.
Print ISSN: 0009-7322. Online ISSN: 1524-4539

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circ.ahajournals.org/content/113/11/1415

Data Supplement (unedited) at:
http://circ.ahajournals.org/content/suppl/2006/03/13/CIRCULATIONAHA.105.591271.DC1

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in _Circulation_ can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to _Circulation_ is online at:
http://circ.ahajournals.org/subscriptions/