EVIDENCE FOR HAPLOTYPE-BASED ASSOCIATION IN SLE AT THE C-REACTIVE PROTEIN LOCUS:
POPULATION-BASED AND FAMILY-BASED ASSOCIATION STUDIES

by

Pei-an Betty Shih

B.A., North Carolina School of the Arts, 1995
M.P.M., Carnegie Mellon University, 2000

Submitted to the Graduate Faculty of
Graduate School of Public Health in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

University of Pittsburgh

2007
In memory of my father

and

In honor of my mother
Systemic lupus erythematosus (SLE) is a major public health problem in the U.S. Cardiovascular disease (CVD) risk increases significantly in SLE patients, resulting in serious morbidity and mortality. Accelerated atherosclerosis and markedly higher prevalence of CVD risk factors (intermediate phenotypes) are thought to directly contribute to these consequences. Given the significant mortality and morbidity associated with SLE and high prevalence of CVD in SLE, identifying genetic variation associated with both SLE risk and intermediate phenotypes of CVD is of significant importance.

C-reactive protein (CRP) is a sensitive marker of inflammation. Increased CRP levels have been found to be associated with cardiovascular events in a large number of healthy populations and may contribute to atherosclerosis. The gene coding for CRP is located on chromosome 1q23.2, which falls within a linkage region thought to harbor a systemic lupus erythematosus (SLE) susceptibility gene. Moreover, two single nucleotide polymorphisms (SNPs) in the CRP gene have recently been shown to be associated with CRP levels and/or SLE risk in a British family-based cohort. This study was aimed to assess the genetic association between five CRP tagSNPs with SLE risk and intermediate phenotypes of CVD.

The association between CRP and SLE risk, assessed in two independently-ascertained SLE cohorts, was tested in a case-control Caucasian sample of 337 SLE and 448 healthy controls
from Pittsburgh and a family-based sample of 203 Caucasian SLE trios from Los Angeles. While none of the SNPs were found to be associated with SLE risk individually, global haplotype statistics revealed significant association \( (p < 0.000001) \) in the Pittsburgh cohort whereas all those haplotypes containing two potentially functional SNPs (-390 and +90) showed association with SLE risk in the Los Angeles cohort \( (p = 0.01 – 0.06) \). The association study between \( CRP \) and intermediate phenotypes of CVD and stroke risk was tested in 237 of the SLE women from the Pittsburgh cohort. Four of the five tagSNPs (-861, -390, +90, and +838) examined revealed significant association with risk of intermediate phenotypes of CVD \( (p < 0.001 \) to 0.04).

In summary, our data did not confirm previously observed individual SNP association with SLE, but suggested that unique haplotype combinations in the \( CRP \) gene may modify the risk of developing SLE, and that variation in \( CRP \) may contribute to the accelerated atherosclerosis in SLE.
TABLE OF CONTENTS

ACKNOWLEDGEMENT .......................................................................................................... xv

1.0 INTRODUCTION CHAPTER ................................................................................... 1

1.1 SYSTEMIC LUPUS ERYTHEMATOSUS (SLE) ...................................................... 1

1.1.1 Epidemiology and clinical features ................................................................. 2

1.1.2 Etiopathogenesis ............................................................................................ 7

1.1.2.1 A Pathway of Gene → Environment → Abnormal Immune Response .......................................................................................................................... 7

1.1.2.2 Autoantibodies – Hallmark of SLE .............................................................. 7

1.1.2.3 B-cells and T-cells in SLE ........................................................................ 9

1.1.2.4 Cytokines .................................................................................................. 9

1.1.2.5 Environmental Risk Factors and SLE ...................................................... 10

1.1.2.6 Genetics of SLE ...................................................................................... 11

1.2 C-REACTIVE PROTEIN (CRP) .......................................................................... 14

1.2.1 CRP as an acute phase protein and implication for CVD ......................... 15

1.2.2 Gene encoding CRP .................................................................................... 17

1.3 PROJECT METHODS SUMMARY ...................................................................... 20

1.4 CHAPTER 1 REFERENCES ............................................................................... 22
3.1 ABSTRACT................................................................................................. 59
3.2 INTRODUCTION ....................................................................................... 60
3.3 SUBJECTS AND METHODS ..................................................................... 61
  3.3.1 Human Subjects and Data Collection. ............................................... 61
  3.3.2 SNP selection and Genotyping............................................................ 62
  3.3.3 Statistical and In Silico Analysis.......................................................... 63
3.4 RESULTS .................................................................................................... 64
  3.4.1 Gene structure, Hardy-Weinberg equilibrium, and linkage disequilibrium in CRP ................................................................. 64
  3.4.2 Marginal over-transmission of -390T and +90T alleles....................... 64
  3.4.3 Significant “Mini-Haplotype” Associations with SLE risk................... 65
  3.4.4 Marginal Global Haplotype association with SLE risk....................... 65
  3.4.5 Association of CRP SNPs with SLE glomerulonephritis..................... 66
3.5 DISCUSSION............................................................................................... 66
3.6 ACKNOWLEDGEMENTS ......................................................................... 73
3.7 MANUSCRIPT 2 TABLES AND FIGURES................................................. 74
3.8 MANUSCRIPT 2 REFERENCES ............................................................... 79
4.0 MANUSCRIPT 3. PLEIOTROPIC EFFECTS OF C-REACTIVE PROTEIN POLYMORPHISMS INFLUENCING INTERMEDIATE TRAITS OF CARDIOVASCULAR DISEASE RISK IN SYSTEMIC LUPUS ERYTHEMATOSUS.... 83
  4.1 ABSTRACT.............................................................................................. 84
  4.2 INTRODUCTION ..................................................................................... 85
  4.3 SUBJECTS AND METHODS .................................................................... 86
4.3.1 Subjects .................................................................................................................. 86
4.3.2 Cardiovascular disease intermediate phenotypes measurement ............ 87
4.3.3 Inflammatory markers measurement ............................................................... 87
4.3.4 Vascular disease measurement .................................................................... 88
4.3.5 Statistical analysis ......................................................................................... 89
4.3.6 Genotyping ..................................................................................................... 89

4.4 RESULTS ................................................................................................................. 90
4.4.1 Clinical characteristics of SLE women ......................................................... 90
4.4.2 Allele and genotype frequencies of CRP SNPs and corresponding CRP levels ................................................................................................................. 90
4.4.3 Correlation between circulating CRP levels and cardiovascular disease intermediate phenotypes ................................................................. 91
4.4.4 Pleiotropic effects of CRP with CVD intermediate phenotype .............. 92
4.4.5 Associations of SNP -861 with BMI and SNP +838 with Ischemic Stroke ......................................................................................................................... 93

4.5 DISCUSSION .......................................................................................................... 94

4.6 MANUSCRIPT 3 TABLES AND FIGURES ......................................................... 98

4.7 MANUSCRIPT 3 REFERENCES ....................................................................... 105

5.0 CONCLUSION CHAPTER .................................................................................. 108

5.1 SUPPLEMENTARY DATA OF THIS DISSERTATION ......................... 108

5.1.1 Genetic and environmental contributions of serum CRP ...................... 108
5.1.2 Correlation between SLE disease activity and CRP levels ................. 111

5.2 DISSERTATION SUMMARY ............................................................................. 114
5.2.1 Does variation in CRP gene confer susceptibility to SLE? .................. 114
5.2.2 Does CRP variation affect intermediate phenotypes of CVD in SLE? 115
5.2.3 Does CRP variation affect CVD risk in SLE? .............................. 115
5.2.4 How does environmental factors and CRP variation contribute to circulating CRP in SLE, and is CRP level associated with disease activity? ..... 116

5.3 FUTURE WORK .................................................................................. 116
5.4 PUBLIC HEALTH SIGNIFICANCE AND CONCLUDING REMARKS 119
5.5 REFERENCES FOR CONCLUSION CHAPTER .............................. 121

BIBLIOGRAPHY .................................................................................. 123
LIST OF TABLES

Table 1.1. The 1982 American College of Rheumatology Criteria for Classification of SLE ...... 5
Table 1.2. The Biomarkers used to assess SLE disease activity, adapted from (17) ................. 6
Table 1.3. Chromosomal regions exhibiting confirmed significant linkage to SLE, adapted from (59) ........................................................................................................................................ 12
Table 1.4. CRP heritability reported in various population ..................................................... 17
Table 2.1. Genotype and allele frequencies of CRP SNPs ........................................................ 47
Table 2.2. Pairwise Linkage Disequilibrium between CRP SNPs ........................................... 48
Table 2.3. CRP Haplotype Case-Control Comparison .............................................................. 49
Table 2.4. Association of CRP Polymorphisms and Mean logCRP level (± SD) ..................... 50
Table 2.5. Association of CRP Haplotype with Serum logCRP Levels in SLE ....................... 51
Table 3.1. Pairwise Linkage Disequilibrium between CRP SNPs in founders ....................... 74
Table 3.2. Genotype frequencies of CRP SNPs – comparison between UCLA cohort and Pitt Cohort ........................................................................................................................................ 75
Table 3.3. Single-marker association analysis in the CRP gene ............................................. 76
Table 3.4. Haplotype association analysis of CRP gene ......................................................... 77
Table 4.1. Descriptive statistics of White SLE Women (mean ± sd) ...................................... 98
Table 4.2. Genotype and allele frequencies of CRP SNPs and associated mean logCRP levels (± SD) in 237 SLE women ................................................................. 99
Table 4.3. Pearson's product-moment correlation between CRP and various intermediate phenotypes of CVD .......................................................... 100

Table 4.4. Inferential statistics for association between CRP polymorphisms and various intermediate phenotypes ............................................................................................................. 101

Table 4.5. Pleiotropic effects of CRP SNPs with cluster of CVD risk factors. ...................... 102

Table 4.6. Multivariate regression model for Body Mass Index (BMI) in 237 SLE patients. .. 103

Table 4.7. Multivariate regression model for Stroke in 11 SLE patients. ................................. 104

Table 5.1. Multivariate regression model for Body Mass Index (BMI) in 237 SLE patients ... 110

Table 5.2. Distribution of SLE patients by disease activity score (SLAM) and CRP levels..... 112

Table 5.3. No significant association between CRP SNPs and SLAM–CRP correlation ....... 113
LIST OF FIGURES

Figure 1.1. Sample autoantibodies in SLE................................................................. 8
Figure 1.2. Contributions of interferon-α (IFNα) to autoimmunity in systemic lupus
erythematosus, adapted from (40)........................................................................... 10
Figure 1.3. Human CRP 3D structure. Adapted from (85)........................................... 16
Figure 1.4. C-reactive protein gene and seven TagSNPs examined in this project........... 18
Figure 3.1. CRP gene and position of five TagSNPs..................................................... 78
Figure 5.1. Scatter plot of CRP levels versus SLE disease activity................................. 111
ACKNOWLEDGEMENT

The past four years have been a memorable and enjoyable journey of learning and growth. My success as an epidemiologist specializing in human genetic research was directly attributed to my two advisors: Dr. Susan Manzi and Dr. Ilyas Kamboh. I am thankful to Dr. Manzi for providing me the opportunity to take up research topics that are of significant personal interest to me. I am especially grateful for Dr. Kamboh “adopting me” into the department of Human Genetics, teaching me from the ground up, and inspiring me to become an independent researcher. I am truly grateful for their continual trust and support in my work.

I would like to also thank my entire doctoral committee for providing tremendous input to me every step of the way. Dr. Candace Kammerer has served as my statistical genetics mentor, patiently guiding me at every road block, allowing me the opportunity to gain expertise in various novel statistical techniques, and most importantly, the analytical ability to develop an optimum study design. Drs. Chau-Ching Liu and Joseph Ahearn from the School of Medicine have served as my mentors in immunology and basic science topics related to my research. Dr. Liu was not only a supportive mentor but also a true friend providing me with scientific and personal advice. Dr. Joe Zmuda has served not only as my committee member but was also my go-to person for Hapmap related updates during the course of my study. I am thankful for having this diversified panel of experts for my doctoral committee. Additionally, I thank Drs.
Michael Barmada and Eleanor Feingold from the department of Human Genetics for having
suggested additional statistical methods and interpretations that enhanced my studies.

I would also like to express my sincere thanks to Dr. Betty Tsao at UCLA for allowing
me the opportunity to work under her mentorship in the summer of 2006, and utilized her
research cohort to achieve scientific goals of my work. She and Dr. Chris Carlson from the
University of Seattle were a joy to collaborate with and I thank them for their teachings. I would
like to thank my friends for influencing me positively in so many ways both in scientific and
personal growth. These included Margaret Kenney, Jay White, Brian Henson, and Ryan Minster
whom have taught me various skills and provided laughter during my studies. I would like to
especially thank Thomas Kiesling for believing in me, and for his unconditional support which
contributed significantly to the completion of this work.

Finally, I would not be here today without the love and support of my family. I thank my
parents especially for instilling a love of learning in me, enduring and encouraging me to be the
best in everything I dream to accomplish. They are the source of my inspiration and motivation
to succeed in life.
1.0 INTRODUCTION CHAPTER

1.1 SYSTEMIC LUPUS ERYTHEMATOSUS (SLE)

Lupus (Latin for wolf) was first described by the thirteenth century physician Rogerius who noted the erosive facial lesions were reminiscent of a wolf's bite (1). The shape of the typical mid-face lupus rash was then allegorized again by von Hebra in 1845 in a published illustration to a butterfly spreading its red wings (Gr. erythema, flash), while the lesions now referred to as discoid lupus were described in 1833 by Cazenave under the term erythema centrifugum (2). The systemic nature of lupus was first observed in 1872 by Kapos, whom noted “…even dangerous constitutional symptoms may be intimately associated with the process in question (facial rash), and that death may result from conditions which must be considered to arise from the local malady” (3).

The existence of a disseminated or systemic form of lupus was firmly established by Osler in 1904 (4). This important establishment of a disseminated form of lupus allowed extensive pathologic studies of today’s SLE, resulting in the milestone discoveries, including the lupus erythematosus (LE) cell (a neutrophil or macrophage that has phagocytized the denatured nuclear material of an injured cell) in 1948 (5), the presence of antinuclear antibodies (6), and antibodies to deoxyribonucleic acid (DNA) (7) in SLE patients.
1.1.1 Epidemiology and clinical features

SLE is the prototypic systemic inflammatory autoimmune disease affecting predominantly younger premenopausal women at onset. Despite significant medical advances made in the past decade, SLE continues to cause devastating disability, morbidity and mortality in the more than 1.5 million Americans affected by this chronic disease.

The complex nature of SLE is well known for its heterogeneous and unpredictable phenotypes, therefore this disease model lacks “clean” epidemiological features across the general population to easily define and capture SLE cases. Using the 1982 revised American College of Rheumatology (ACR) criteria to define SLE cases (8), epidemiological studies found the prevalence of SLE in the population to be about 54 cases per 100,000 in white females, and 40 to 50 per 100,000 in the general population (9). Improved detection of milder disease in the past decades resulted in tripling of the incidence in the last 40 years to 5.56 per 100,000 (10). A recent study in Allegheny County, Pennsylvania found the 5-year incidence rate to be 3.4 per 100,000 (11).

SLE is a complex, chronic, multi-system autoimmune disease involving both humoral and cellular aspects of the innate and acquired immune systems. While disease states can be as simple as troublesome alopecia, they can also manifest into life-threatening kidney and brain damage. Due to the improved diagnosis and treatment of the disease, survival in SLE has improved dramatically from 50% after 4 years in 1954 to 97% after 5 years and 90% after 10 years (9). The extended lifespan of SLE patients allows the observation of a bimodal mortality pattern. Patients who die early in the course of their disease likely expire due to active lupus and/or infection, while those who die late in the course of the disease often expire from strikingly high incidence of myocardial infarction due to atherosclerotic heart disease (12).
SLE is known to impact social productivity because sixty-five percent of patients have onset between the ages of 16 and 55 (13). SLE disproportionately affect minorities, the prevalence of SLE is higher among Asians, Afro-Americans, Afro-Caribbean, Hispanic Americans, than Caucasians, and the disease activity appears to be most severe and difficult to treat in African American women (14). Young African American females are at the highest risk to develop SLE, which was reported to have a prevalence of 1 in 245 black women age 15 to 64 in a California metropolitan area (15). SLE has the most striking impact on females with an incident rate of 9 to 1 over males. In child bearing age SLE disproportionately affects women up to 15 times more than men. Even in pediatric SLE, the female-to-male ratio has been reported to be 3:1. Taking these data together, it is clear that hormones play a significant role in the SLE etiology (14).

Predicting the characteristics of the clinical phase of SLE is a difficult task for this autoimmune disease. SLE not only can affect any system and organ of the patient, the heterogeneous clinical presentations almost always differ from patient to patient, as well as differ within the same patient from one phase of disease progression to another in her/his lifetime. Therefore, the diagnosis and management of SLE has long been a challenge for health care providers. To have a confirmed diagnosis of SLE, patients must have at least 4 of the 11 criteria given in Table 1.1 at some point in time. Since patients go in and out of remission for different “organ involvement” and/or serological readings, it is often difficult for patients with milder cases of the disease to recall correctly all the symptoms that have occurred. Therefore it is likely for physicians to miss the cases with milder symptoms at onset, which might develop into full-blown disease eventually. In 1997, the Diagnostic and Therapeutic Criteria Committee of the
ACR reviewed the 1982 criteria and recommended revisions to the 10th criterion of immunologic disorders based on committee consensus (16):

1. Delete item 10(a) (Positive LE cell preparation), and
2. Change item 10(d) to “Positive finding of antiphospholipid antibodies based on (1) an abnormal serum level of IgG or IgM anticardiolipin antibodies, (2) a positive test result for lupus anticoagulant using a standard method, or (3) a false-positive serologic test for syphilis known to be positive for at least 6 months and confirmed by *Treponema pallidum* immobilization or fluorescent treponemal antibody absorption tests.” (Table 1.1).
Table 1.1. The 1982 American College of Rheumatology Criteria for Classification of SLE

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Malar rash</td>
<td>Fixed erythema, flat or raised, over the malar eminences, tending to spare the nasolabial folds</td>
</tr>
<tr>
<td>2. Discoid rash</td>
<td>Erythematous raised patches with adherent keratotic scaling and follicular plugging: atrophic scarring may occur in older lesions</td>
</tr>
<tr>
<td>3. Photosensitivity</td>
<td>Skin rash as a result of unusual reaction to sunlight, by patient history or physician observation</td>
</tr>
<tr>
<td>4. Oral ulcers</td>
<td>Oral or nasopharyngeal ulceration, usually painless, observed by a physician</td>
</tr>
<tr>
<td>5. Arthritis</td>
<td>Nonerosive arthritis involving ≥2 peripheral joints, characterized by tenderness, swelling, or effusion</td>
</tr>
<tr>
<td>6. Serositis</td>
<td>(A) Pleuritis: convincing history of pleuritic pain or rub heard by a physician or evidence of pleural effusion, or (B) Pericarditis: documented by ECG or rub or evidence of pericardial effusion</td>
</tr>
<tr>
<td>7. Renal disorder</td>
<td>(A) Persistent proteinuria &gt;0.5 g/d or &gt;3+ if quantitation not performed, or (B) Cellular casts: may be red blood cell, hemoglobin, granular, tubular, or mixed</td>
</tr>
<tr>
<td>8. Neurologic disorder</td>
<td>(A) Seizures: in the absence of offending drugs or known metabolic derangements (eg, uremia, ketoacidosis, or electrolyte imbalance), or (B) Psychosis: in the absence of offending drugs or known metabolic derangements (eg, uremia, ketoacidosis, or electrolyte imbalance)</td>
</tr>
<tr>
<td>9. Hematologic disorder</td>
<td>(A) Hemolytic anemia: with reticulocytosis, or (B) Leukopenia: &lt;4000/mm(^3) total on ≥2 occasions, or (C) Lymphopenia: &lt;1500/mm(^3) on ≥2 occasions, or (D) Thrombocytopenia: &lt;100,000/mm(^3) in the absence of offending drugs</td>
</tr>
<tr>
<td>10. Immunologic disorder</td>
<td>(A) Positive LE cell preparation, or (B) Anti-DNA: antibody to native DNA in abnormal titer, or (C) Anti-Sm: presence of antibody to Sm nuclear antigen, or (D) False positive serologic test for syphilis known to be positive for at least 6 months and confirmed by (T) pallidum immobilization or fluorescent treponemal antibody absorption test</td>
</tr>
<tr>
<td>11. Antinuclear antibody</td>
<td>An abnormal titer of antinuclear antibody by immunofluorescence or an equivalent assay at any point in time and in the absence of drugs known to be associated with “drug-induced lupus” syndrome</td>
</tr>
</tbody>
</table>

The proposed classification is based on 11 criteria. For the purpose of identifying patients in clinical studies, a person shall be said to have systemic lupus erythematosus if any 4 or more of the 11 criteria are present, serially or simultaneously, during any interval of observation.

SLE disease manifestation can be categorized in two phases, active phase and chronic damage phase. Active phase of clinical symptoms often requires intense immune-suppressive agents to quickly “slow-down” the disease activity to prevent chronic organ damage that is often
irreversible and sometimes fatal. Preventing damage from occurring is the most important goal in clinical practice. The significant risk factors physicians aim to reduce in order to prevent morbidity and mortality include: infection, hypertension, diabetes, osteoporosis, and atherosclerosis. It is therefore important to manage disease activity in SLE patients using several serological biomarkers listed in Table 1.2, in addition to routine physical examinations. These serological biomarkers helped physicians to identify asymptomatic patients who are at increased risk of flaring; hence treatment can be administered in a timely fashion to prevent the severe irreversible damage from occurring.

Table 1.2. The Biomarkers used to assess SLE disease activity, adapted from (17)

<table>
<thead>
<tr>
<th>Biomarker</th>
<th>Activity measure</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Autoantibodies</strong></td>
<td></td>
</tr>
<tr>
<td>Anti-C1q</td>
<td>SLEDAI</td>
</tr>
<tr>
<td>Anti-dsDNA</td>
<td>SLAM</td>
</tr>
<tr>
<td>Anti-ssDNA</td>
<td>SLAM</td>
</tr>
<tr>
<td><strong>Complement</strong></td>
<td></td>
</tr>
<tr>
<td>C3</td>
<td>PGA, M-LAI, M-SLEDAI, M-BILAG, SLAM</td>
</tr>
<tr>
<td>C4d</td>
<td>Not available</td>
</tr>
<tr>
<td>CR1/E</td>
<td>SLEDAI</td>
</tr>
<tr>
<td>CRP</td>
<td>SLEDAI</td>
</tr>
<tr>
<td><strong>Cytokines and soluble cell surface molecules</strong></td>
<td></td>
</tr>
<tr>
<td>CD27^{high}</td>
<td>SLEDAI, ECLAM</td>
</tr>
<tr>
<td>IL-1 receptor antagonist</td>
<td>BILAG</td>
</tr>
<tr>
<td>IL-6</td>
<td>SLEDAI</td>
</tr>
<tr>
<td>IL-10</td>
<td>BILAG</td>
</tr>
<tr>
<td>IL-12</td>
<td>BILAG</td>
</tr>
<tr>
<td>Perinuclear-ANCA</td>
<td>SLEDAI</td>
</tr>
<tr>
<td>Urine MCP-1</td>
<td>BILAG</td>
</tr>
</tbody>
</table>

*a Activity measure refers to the comprehensive scoring systems used to assess lupus disease activity.
*b Abbreviations: BILAG, British Isles lupus assessment group; ECLAM, European consensus lupus activity measurement; LAI, lupus activity index; M-, modified; PGA, physicians global assessment; SLAM, systemic lupus activity measure; SLEDAI, systemic lupus erythematosus disease activity index.
1.1.2 Etiopathogenesis

1.1.2.1 A Pathway of Gene → Environment → Abnormal Immune Response

Despite the significant advancement made in immunobiology, the etiopathogenesis of SLE is still not yet completely understood. Familial aggregation of lupus and disease concordance rates in monozygotic twins suggest the involvement of both genetic and environmental factors in the predisposition to lupus (18). It is believed that gene products are influenced by environmental stimuli and that most genes only have deleterious effects in the presence of particular environmental stimulus (19, 20). The resulting abnormal immune response which leads to disease phenotype are well known for certain distinguished immunological and cellular characteristics: over production of autoantibodies, antigens that autoantibodies direct against, exaggerated B cell responses, altered activation of T-cells resulting in increased apoptosis, and cytokine-related innate immunity dysfunctions (17, 21).

1.1.2.2 Autoantibodies – Hallmark of SLE

SLE is characterized by an overproduction of autoantibodies. Recent data have showed the presence of over 100 autoantibodies in patients with SLE (22), many of which have been associated with disease activity and are being used as biomarkers (Table 1.2). More surprisingly, it is estimated that 88% of SLE patients have serological evidence of some types of autoantibodies before the diagnosis of their disease (23). Significant autoantibodies commonly seen in SLE patients include dsDNA, ANA, Ro (SSA), La (SSB), and antiphospholipidantibodies (aPL). SLE sub-phenotypes that were found to be associated with specific antibodies are illustrated in Figure 1.1. Antigens in SLE patients are material of DNA/RNA plus protein, which include chromatin components such as dsDNDA, nucleosomes
and histones. It is believed that these antigens are “mistaken” by immune system of SLS patients as a “foreign target”, therefore a wave of autoantibodies are produced to direct against these antigens (17). Recently, Katzav et al. (24) injected human anti-ribosomal P antibodies extracted from an SLE patient into healthy female mice. These mice were twice as likely as the control mice to show typical signs of depressive cognitive and motor functions related to lupus, suggesting the mechanistic role of anti-ribosomal P antibodies on central nervous system (CNS) involvement of lupus, one of the most serious organ involvement in patients.

Figure 1.1. Sample autoantibodies in SLE
1.1.2.3 B-cells and T-cells in SLE

The contributions of B- and T-cell abnormalities to SLE are well established (25). It is believed that lymphoid architecture has a key role in the development of the humoral immune response. Since variations in program death 1 (PD1) and protein tyrosine phosphatase non-receptor type 22 (PTPN22) genes were found to be associated with SLE risk and that they are both involved in lymphocyte activation, the hyperactivity of B-cell may be a key mechanism for the disease development. Altered T-cell activation thresholds and accompanied increased co-stimulations are found in SLE patients. These abnormal activations of T-cells may contribute to defects both in helper T-cell functions and the T regulatory cell (Treg) mechanism in response to self-antigen presentation (26). Better understandings of the B- and T-cell and the immune system in SLE have helped scientist to develop more specific biologics therapies to treat SLE. Biologic therapies are a new class of drugs produced through genetic manipulation. They include standard single molecule drugs, as well as antibodies and vaccines. Examples of these biologics for treating SLE include B-cell depletion agents (Lymphostat-B, Rituximab), agent inhibiting T–B interaction (IDEC-131), blockade of cytokines (anti-IL-10 antibodies), tolerance induction to DNA and to Ig-peptides and peptide therapy (Riquent) (27), and gene therapy (28).

1.1.2.4 Cytokines

Cytokines profile in SLE patients is believed to consist of both increased pro-inflammatory cytokines (IFN-α, IFN-γ, IL-4, IL-6, and IL-10) and decreased anti-inflammatory cytokines (TGFb and IL-2). The pathogenic role of type I interferons (IFN) in SLE has been well established in both human clinical studies and murine models (29-35). Evidence of association between IFN and SLE include: observed elevated levels of IFN in SLE sera (36), IFN-α levels
correlating with SLE disease activity and organ involvement (37, 38), increased expression of IFN-α-induced genes in SLE (IFN signature) (39-41), induction of lupus-associated autoantibodies and clinical lupus upon IFN-α therapy for some malignancies (42, 43), and differential expression of a IFN-induced gene in lupus and control mice (44). IFN-α is perhaps the “boss cytokine” and many groups are working on understanding the etiology of SLE based interferon pathways in order to develop more therapeutic agents (45, 46).

Figure 1.2. Contributions of interferon-α (IFNα) to autoimmunity in systemic lupus erythematosus, adapted from (40)

1.1.2.5 Environmental Risk Factors and SLE

Environmental risk factors for SLE may include infectious agents (20, 47, 48), environmental pollutants and occupational exposures such as silica (49-51) that can modulate immune responses, and behavioral factors such as smoking (52), diet (53) and sun exposure (54, 55).
Exposure to these factors may induce the production of autoreactive T cells, autoantibodies and the stimulation of pro- and anti-inflammatory cytokines that induce the onset of SLE. Additionally, sex hormones in hormone replacement therapy have been associated with mild flares in SLE, which confirmed the role of female hormones in SLE development (56).

1.1.2.6 Genetics of SLE

Family and twin studies strongly suggest the involvement of genetic factors in the predisposition to lupus. First degree relatives of SLE patient has 20 to 40 times more at risk than those without an immediate family member diagnosed with SLE (18, 57). Using a polygenic inheritance model, SLE has a heritability of 66 ± 11% in the Caucasians (9).

The study of human lymphocyte antigen (HLA) class genes has been conducted to determine specific amino acid sequences in the cell surface molecules that are involved in antigen presentation to T-helper cells in patients with lupus. Using a dense map of polymorphic microsatellites across the HLA region, three distinct haplotypes that encompassed the class II region exhibited transmission distortion. SLE risk HLA haplotypes include: DRB1*1501/DQB1*0602, DRB1*0801/DQB1*0402, and DRB1*0301/DQB1*0201. It is found that having one haplotype increases the relative risk of SLE by 2 to 3 times, by having any two haplotypes the relative risk increases to 4 to 6 times (58).

Single gene dysfunctions that may cause lupus development include genes involved in antigen presentation in the innate immunity (C1q, SAP, PARP, MBL), loci found to associate with tolerance breaking (Sle1 [Mus musculus]), and increased immunological response (Sle2/3 [Mus musculus], IL-10, PD-1, IRF5, PTPN22). Other genes found to be associated with SLE risk include Ifi202, Fas, FasL, p21, bcl2, which are also associated with autoreactive cells, and FcγR, CR1, MBL which are associated with antibody/immune complexes persistence (59, 60).
Genes with the strongest influence on SLE risk are known to be complement-related defects. Though extremely rare, homozygous deficiency of \textit{C1q}, \textit{C1r}, and \textit{C1s} are at up to 95% increased risk of developing SLE, while homozygous deficiency for \textit{C4} have up to 60% risk of SLE. Even partial \textit{C4A} deficiency/defect has been associated with 15% risk of the disease (61). Indeed, association studies showed that complement components \textit{C2} (62) and \textit{C4} (63) genes also have been implicated in SLE. Chromosomal regions, and the candidate genes found associated with risk of SLE within the linked intervals, are shown in Table 1.3.

\textbf{Table 1.3.} Chromosomal regions exhibiting confirmed significant linkage to SLE, adapted from (59)

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>6</th>
<th>12</th>
<th>16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cytogenetic location</td>
<td>1q23</td>
<td>1q25-31</td>
<td>1q41-42</td>
<td>2q35-37</td>
<td>4p16-15.2</td>
<td>6p11-21</td>
<td>12q24</td>
<td>16q12</td>
</tr>
<tr>
<td>Locus name</td>
<td>SLEB1</td>
<td>SLEB2</td>
<td>SLEB3</td>
<td>SLEB4</td>
<td>SLEB4</td>
<td>SLEB4</td>
<td>OAZ</td>
<td></td>
</tr>
<tr>
<td>Candidate genes*</td>
<td>CRP, FCGR2A, FCGR2B, FCGR3A, PBX1</td>
<td>PARP</td>
<td>PDCD1</td>
<td>MHC haplotypes, C4Q0, TNFA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Candidate genes within the linked interval that has been shown to be associated with SLE.

The best-supported SLE susceptibility genome regions exhibiting significance levels in single studies and/or confirmed evidence in an independent sample are: 1q23, 1q41, 2q37, 4p16, 6p21, 11p13, 12q24 and 16q13 (64-70). Recent meta-analysis of genome-wide linkage studies identified the strongest evidence for linkage in regions to be 6p21 and 16q13 (66), confirmed once again the importance of complement pathway genes and importance of loci interaction effects in the genome (69).

A number of genes have also been found to be associated with specific manifestations of SLE phenotypes. These include \textit{FCR2A/cA, MCP1, MBL, PDCD1} associated with lupus
nephritis, \textit{FCR3A} with end stage renal disease \textit{MCRP1} associated with vasculitis and arthritis, and \textit{CRP} associated with atherosclerosis (25, 59, 60).
1.2 C-REACTIVE PROTEIN (CRP)

Pentraxins are highly conserved cyclic pentamers such as CRP, Serum Amyloid P component (SAP) and PTX3 (a cytokine modulated molecule). CRP and SAP bind to chromatin, small nuclear ribonucleoproteins (snRNP), as well as to apoptotic cells that are exposed on the cell membrane. CRP and SAP act as scavengers for dying cells and cell debris. Therefore, defects in these factors may impact the development of SLE in humans and mice in a variety of ways. These arguments also provide a possible rationale for using them in the treatment of autoimmune diseases such as lupus. In fact, treatment of lupus-prone mice with CRP or transgenic over-expression of CRP protects against development of lupus (71). Recent studies also suggest an immune modulatory role of CRP through induction of IL-10 and binding with Fcγ receptors. Consistent with their role in development of autoimmune and inflammatory diseases, CRP has been used as a biomarker of inflammation (72). Given the overwhelming association between CRP and inflammation, CRP genetic variations may also modify risk of SLE itself and excessive cardiovascular disease (CVD) risk in SLE.
1.2.1 CRP as an acute phase protein and implication for CVD

CRP, discovered in 1930 by Tillett and Francis (73), is a major acute-phase protein whose levels increase significantly after tissue injury or inflammation. CRP is believed to be critically important for organism survival, as it is found in both vertebrates and invertebrates phylogenetically spanning 400 million years of evolution. Important functions CRP serves in the immune system include opsonization, activation of the classical pathway of complement, immune-modulation, induction of phagocytosis, and possibly atherogenesis.

Given CRP’s capability of recognizing foreign pathogens and/or damaged cells of the host and to initiate their elimination by interacting with humoral and cellular effector systems in the blood (74), CRP is known to be a sensitive marker of systemic inflammation. In fact CRP has been one of the most widely studied inflammatory markers in the past decade, resulting in at least 14,114 published articles since year 1997 (as documented in NCBI PubMed database).

Mounting evidence indicates that systemic inflammation is linked to the emergence of CVD disease. Several studies have successfully demonstrated associations between increased CRP levels and marked CVD risk in healthy population (75-77). Ridker et al.(78) have found that elevation of high sensitivity CRP (hsCRP) was not only a independent predictor of cardiovascular event in 28,263 healthy postmenopausal women, cardiovascular event prediction model which includes hsCRP was much better than a model with lipids profile alone. Indeed, recent studies using human CRP-transgenic mice have provided convincing evidence that CRP may be an active participant in thrombosis and atherogenesis (79, 80), instead of simply a down-stream inflammatory marker alone.
CRP serves not only as a powerful inflammatory marker, significant risk factor in predicting cardiovascular mortality and potential contributor of atherogenesis, it is also associated with major CVD risk factors including inflammation (81), metabolic syndrome (82), hypertension (83), and obesity (84). These significant associations with well-established Framingham CVD risk factors suggest that circulating CRP may play a significant role in the early development of CVD by affecting the CVD intermediate phenotype progression, which then leads to cardiovascular events.

Native CRP (nCRP) is composed of 5 identical subunits arranged as a cyclic pentamer (Figure 1.3). As subunits separate and dissociate into individual nomonomic units, they undergo a conformational change that modifies CRP’s solubility and antigenicity as monomeric CRP (mCRP). As oppose to nCRP which is a serum-based protein, mCRP is primarily detected in fibrous tissues of normal human blood vessel intima (86). Several groups have attempted to determine the differencial effects nCRP and mCRP have on development of artherosclerosis. While some groups found mCRP to be more proinflammatory than nCRP (87), other found

Figure 1.3. Human CRP 3D structure. Adapted from (85)
nCRP to be responsible for increasing of artherosecrosis (88). Therefore, precise differential biological functions and pathogenetic mechanisms nCRP and mCRP have on artherosecrosis and possibly SLE risk are not certain.

CRP is a highly heritable multi-factorial trait. The heritability for CRP has been reported ranging from 35% to 56% in various populations (Table 1.4). The high of 56% strongly confirm the significance of the role CRP gene plays in various phenotypic associations discussed above.

<table>
<thead>
<tr>
<th>Author</th>
<th>Sample Description</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parkow et al. 2001</td>
<td>N = 2163 (101 pedigrees, 517 sibships, 454 unrelated probands)</td>
<td>35-40% G, 13-30% E</td>
</tr>
<tr>
<td>Vickers et al. 2002</td>
<td>588 members of 98 nuclear families</td>
<td>39%</td>
</tr>
<tr>
<td>Retterstol et al. 2003</td>
<td>155 MZ twins</td>
<td>40%</td>
</tr>
<tr>
<td>Friedlander et al. 2005</td>
<td>142 kindreds</td>
<td>Segregation analysis revealed major environmental factor</td>
</tr>
<tr>
<td>O’Connor et al. 2006</td>
<td>75 MZ twins &amp; 32 DZ twins</td>
<td>56%</td>
</tr>
</tbody>
</table>

1.2.2 Gene encoding CRP

CRP is a relatively small gene with approximately 6800 base pairs. The gene contains two exons (black boxes in Figure 1.4) with a 104 amino acid peptide coding regions. The gene contains a short 5’-untranslated region (89 bp, grey box), with two exons separated by a short intron region (286 bp), and an unexpectedly longer (1.2 kb) 3’-untranslated region (grey box).
Figure 1.4. C-reactive protein gene and seven TagSNPs examined in this project

CRP has been shown to bind chromatin (89, 90), histones (90), and apoptotic cells (91). Given these specific immune-modulating abilities, CRP is also thought to modify the autoimmune disease phenotype by promoting the removal of necrotic and apoptotic cells, and its ability to recruit complement and FcyR mediated effector pathways (92). In SLE patients, defective apoptotic cell clearance and immune complex deposition are thought to cause organ damage and cascading inflammation. Therefore the increased clearance of apoptotic cells and their derived nuclear contents by phagocytic cells via CRP opsonization may possibly prevent the development of potential nuclear antigen-specific autoimmune responses (91, 93).

Recent in vivo studies have shown that lupus-prone BW mice carrying the CRP transgene had reduced proteinuria, lived longer than non-transgenic BW, and accumulation of IgM and IgG in their renal glomeruli was delayed (71). Additionally, injecting CRP to another lupus strain mouse NZB/NZW delayed the onset of high-grade proteinuria and prolonged survival (94). With its unquestionable ties to inflammation, increasing evidence of its participation in atherogenesis, and unique ability to modify the disease phenotypes of SLE, CRP has clearly proven to be a functional candidate gene for SLE.

Abundant epidemiological and functional studies have investigated association between plasma or serum CRP and SLE. Genetic studies on CRP and SLE are, however, very limited to
date. In 2002, Szalai et al. (95) found a polymorphic GT-repeat in the intron of the CRP gene to be associated with baseline CRP levels in SLE patients. In 2005, the same group found that SLE patients with vascular events had a greater number of the CRP GT20 variant compared with SLE patients without vascular events in African-Americans and Hispanics, but not in Caucasians (96). In 2004, Russell et al. investigated CRP as a candidate gene for SLE in two cohorts totalling 586 UK simplex SLE families. They found the basal levels of CRP to be influenced independently by two polymorphisms at the CRP locus, one of which was also associated with SLE and antinuclear autoantibody production (97).

Given the limited literature on association finding between CRP polymorphisms and SLE, it is important to carry out additional genetic studies to determine the role of CRP genetic variation in relation to SLE risk. The objective of this genetic association research project was to investigate the role of CRP as a potential disease-susceptibility locus or biomarker, for SLE using two independent SLE cohorts of North American caucasians. The two cohorts consisted of differentially ascertained case-control sample from Pittsburgh and a family-based sample from Los Angeles. In addition, the role of CRP polymorphisms in circulating CRP levels and additional well-established intermediate phenotypes for cardiovascular disease (CVD) in SLE patients were examined to deepen our understanding of the role of CRP gene in SLE.
Detailed methodology for this project is clearly explained in Method sections in each of the three manuscripts (Chapters 2, 3, and 4). In summary, this project used the candidate gene approach, combining multiple epidemiological study designs to test our hypothesis that variations in the *CRP* gene are associated with SLE risk, SLE clinical manifestations, and CVD risk.

*CRP* tagSNPs for this work were chosen from the SeattleSNPs Program for Genomic Applications (http://pga.gs.washington.edu/) European population, using the LDSelect algorithm (98) using a linkage disequilibrium threshold of $r^2 = 0.64$ and a minor allele frequency threshold of five percent. Five tagSNPs were selected relative to the ATG codon of the CRP translation site in the FASTA database (Genbank NC_000001.9). The positions of these SNPs are -861, -390, +90, +838 and +2043. SNPs -861 and -390 positioned in the promoter region, +90 located in intron/exon boundary, +838 positioned within exon 2, and + 2043 in the 3’ untranslated region. For clarification, reference numbers from NCBI Single Nucleotide Polymorphism are provided for each of the five SNPs: -861 is rs3093059, -390 is rs3091244, +90 is rs1417938, +838 is rs1800947 and +2043 is rs1205.

Allelic and haplotype association tests were used in the case-control cohort based at the University of Pittsburgh (Pitt). Case-control study analysis were conducted using R.2.1.1 and EH. Combined linkage and association tests were used in the familial cohort based at the University of California of Los Angeles (UCLA). Family-based study analyses were conducted using FBAT and GENEHUNTER, ver.2.0.
This research project included patients defined using both the 1982 and 1997 American College of Rheumatology criteria for definite or probable SLE (8, 16) to ensure capturing all SLE patients.


90. Du Clos TW, Zlock LT, Rubin RL. Analysis of the binding of C-reactive protein to histones and chromatin. Journal of Immunology 1988;141(12):4266-70.


92. Szalai AJ. C-reactive protein (CRP) and autoimmune disease: facts and conjectures. Clinical & Developmental Immunology 2004;11(3-4):221-6.


2.0 MANUSCRIPT 1. GENETIC VARIATION IN THE C-REACTIVE PROTEIN (CRP) GENE MAY BE ASSOCIATED WITH THE RISK OF SYSTEMIC LUPUS ERYTHEMATOSUS AND CRP LEVELS

Manuscript in Preparation

P. Betty Shih\textsuperscript{1,2}, Susan Manzi\textsuperscript{1,2}, Penny Shaw\textsuperscript{2}, Margaret Kenney\textsuperscript{1}, Amy Kao\textsuperscript{2}, Frank Bontempo\textsuperscript{3}, M. Michael Barmada\textsuperscript{1}, Candace Kammerer\textsuperscript{1}, M. Ilyas Kamboh\textsuperscript{1}

\textsuperscript{1} Graduate School of Public Health, University of Pittsburgh, PA; \textsuperscript{2} Lupus Center of Excellence, University of Pittsburgh School of Medicine, PA; \textsuperscript{3} Department of Medicine, University of Pittsburgh School of Medicine, PA
2.1 ABSTRACT

The gene coding for C-reactive protein (CRP) is located on chromosome 1q23.2, which falls within a linkage region thought to harbor a systemic lupus erythematosus (SLE) susceptibility gene. Recently, two single nucleotide polymorphisms (SNPs) in the CRP gene (+838, +2043) have been shown to be associated with CRP levels and/or SLE risk in a British family-based cohort. This study was aimed to confirm the reported association in an independent population-based case-control cohort, and also to investigate the impact of four additional CRP tagSNPs (-861, -860, -390, +90) on SLE risk and serum CRP levels. Altogether we examined the association of six tagSNPs in the CRP gene in 337 Caucasian SLE women and 448 healthy controls. While none of the SNPs were found to be associated with SLE risk individually, global haplotype statistics revealed significant association (p<0.000001). Three SNPs (-861, -390, +90) were found to significantly influence serum CRP level in SLE cases. Haplotypes consisting of these three SNPs also confirmed significant impact these SNPs have on CRP levels. However, global haplotype test revealed no significant association between CRP and CRP levels. Our data did not confirm previously observed individual SNP associations with either SLE risk or CRP levels, but suggested that unique haplotype combinations in the CRP gene may modify the risk of developing SLE.

2.2 INTRODUCTION

The pathogenesis of systemic lupus erythematosus (SLE) is complex and multi-factorial, involving interactions between multiple genes, hormones and several environmental factors.
Even though the etiology of SLE remains elusive, it is believed that impaired handling of antigen–antibody complexes and subsequent tissue deposition leading to release of inflammatory mediators and an array of inflammatory cells can induce a broad spectrum of clinical manifestation (1). Epidemiological studies have shown that the prevalence of SLE is up to 15 times higher in women than in men (2), and that premature atherosclerosis and cardiovascular disease (CVD) are significantly increased in SLE women as compared to the general population (3-6). Traditional risk factors alone are insufficient to fully explain this significant increased burden of CVD in young SLE women. Of a range of additional factors that are thought to be contributing to the premature atherosclerosis observed in SLE, chronic inflammation plays a pivotal role in the pathogenesis of both SLE and CVD.

The genetics predisposition of SLE is complex, possibly involving multiple genes, hormonal and environmental factors, and interactions among them. Family and twin studies suggest that genetic factors play a significant role in the predisposition to SLE (7, 8). The estimated heritability of SLE in Caucasian is 66% (9). Recent genome wide linkage analyses in multiplex SLE families have provided many chromosomal regions for exploration of disease-predisposing genes, including a region on the q-arm of chromosome 1 (10). The gene coding for C-reactive protein (CRP) is located at 1q23, which falls within the 1q23-43 region thought to harbor a susceptibility gene for SLE in multiple independent genome scans of both mice and humans (11-15). The unique position of the CRP gene makes it a logical positional candidate gene to investigate as a susceptibility locus for SLE.

CRP is also a functional candidate gene based on the physiological activity of its products. CRP is an important liver-derived acute-phase protein that can increase up to 1000-fold in serum as a response to diverse stimuli such as infection or injury (16). Important
functions of CRP in the immune system include opsonization (17), activation of the classical pathway of complement (18), immune-modulation, induction of phagocytosis (19), and possibly atherogenesis (20). In the past decade, CRP has been the most widely studied inflammatory marker in predicting CVD risks. Recent *in vitro* studies have provided convincing evidence that CRP is likely an active participant in thrombosis and atherogenesis (20, 21).

In addition to its strong association with inflammation, CRP has been shown to bind chromatin (22), histones (23) and apoptotic cells (24). These unique characteristics of CRP are thought to contribute to its ability to modify the autoimmune disease phenotype by promoting the removal of necrotic and apoptotic cells and recruiting complement and FcyR-mediated effector pathways (25). In the host, the increased clearance of apoptotic cells and their derived nuclear contents by phagocytic cells via CRP opsonization may prevent the development of potential nuclear antigen-specific autoimmune responses (24, 26). Recent *in vivo* studies have shown that lupus-prone BW mice carrying the *CRP* transgene had reduced proteinuria, lived longer than non-transgenic BW, and had delayed accumulation of IgM and IgG in their renal glomeruli (27). Injecting CRP to another lupus strain mouse, NZB/NZW, also delayed the onset of high-grade proteinuria and prolonged survival (28). CRP’s autoimmunity prevention ability may come from its ability to prevent activation of autoreactive B cells by promoting clearance of antoantigens to non-antigen presenting sites (25).

Several studies have shown that CRP levels in SLE patients are abnormally elevated both in the absence and presence of infection (29-33). The value of using CRP to monitor SLE disease activity has remained controversial given the inconsistent correlation between circulating CRP and disease activity from numerous studies (34-38). The abnormal elevation pattern of CRP in SLE patients provided the first clinical clue that variation in the *CRP* may contribute to
the pathogenesis of SLE. With CRP’s unquestionable tie to inflammation, association with atherogenesis, its unique ability to modify the disease phenotypes of SLE and its genetic candidacy, CRP serves as a promising susceptibility gene for SLE.

Russell et al.(39) recently found basal levels of CRP to be influenced independently by 2 CRP polymorphisms (+838 & +2043), and the latter was also associated with SLE and antinuclear autoantibody production. They hypothesized that defective disposal of potentially immunogenic material, indicated by low basal CRP levels, may be a contributory factor in lupus pathogenesis. In the present study, we examined six tagSNPs both individually and as haplotypes to investigate the associations of CRP with SLE risk and serum CRP levels in SLE patients. We hypothesized that the variation in the CRP gene may contribute to the genetic susceptibility of SLE and impact CRP levels in patients with SLE.

2.3 SUBJECTS AND METHODS

2.3.1 Subjects

A total of 337 white female SLE cases and 448 healthy female controls were included in this study. All cases were 18 years of age or older (mean age 43 ± 11 years), and met the 1982 and 1997 American College of Rheumatology criteria for definite or probable SLE(40, 41). All subjects were recruited from the Pittsburgh Lupus Registry, and have been seen either at the University of Pittsburgh Medical Center or by independent rheumatologists in the Pittsburgh metropolitan area. This mixed cohort provides a population-based representation of SLE compared to a cohort that is recruited strictly from a tertiary referral center, providing us with
better presentation of heterogeneity of this disease. Controls were geographically matched and obtained from the Central Bank of Pittsburgh, and had no apparent history of SLE (mean age 45 ± 13 years). This study was approved by the University of Pittsburgh Institutional Review Board, and all subjects provided written informed consent.

2.3.2 SLE clinical and laboratory characteristics

C-reactive protein was measured using high sensitivity enzyme-linked immunoabsorbent assay. Details of the assay have been previously described (42, 43). SLE disease activity and cumulative damage were measured by the same physician (SM) in all patients, using the Systemic Lupus Activity Measure (SLAM) (44) and the Systemic Lupus International Collaborating Clinics (SLICC) damage index (45), respectively. Renal disease among SLE patients was defined using the ACR criteria, which requires (a) renal biopsy showing lupus nephritis, or (b) persistent proteinuria greater than 0.5 grams per day or greater than 3+ if quantification is not performed, or (c) evidence of cellular casts in the urine. Central nervous involvement (CNS) among SLE patients was defined by history of seizure or psychosis due to SLE. Joint involvement was defined as inflammatory arthritis. Laboratory studies included ds-DNA, antiphospholipid antibodies, serum C3, and C4.

2.3.3 TagSNP selection and genotyping

Informative tagSNPs in the CRP gene were chosen by utilizing the SeattleSNPs Program for Genomic Applications web site (http://pga.gs.washington.edu/education.html). SNPs -861, -860, and -390 are located in the promoter region of CRP, SNP +90 is located in intron/exon boundary,
+838 is present within exon 2, and +2043 maps in the 3’ untranslated region. We have designated our SNPs based on their position relative to the ATG codon of the CRP translation site in the FASTA database. For clarification, reference numbers from the NCBI Entrez SNP database are provided for each of our six SNPs: -861 is rs3093059, -860 is rs3093060, -390 is rs3091244, +90 is rs1417938, +838 is rs1800947 and +2043 is rs1205.

Genotyping for +838 and +2043 was obtained using polymerase chain reaction restriction fragment length polymorphism (PCR-RFLP). SNPs -861 and -860 were genotyped using pyrosequencing assays. SNPs -390 and +90 were genotyped using TaqMan assays.

2.3.4 Statistical analyses

Allele frequencies were calculated by the allele counting method. Goodness of fit to Hardy–Weinberg expected proportions was examined by $\chi^2$ test. The pair-wise linkage disequilibrium (LD) between markers was estimated using the D’ method(46). The differences in genotype frequencies between cases and controls were tested by Fisher's Exact test. Multivariate logistic regression models were used to assess minor allele carrier effects of each SNP with SLE risk, adjusting for age as a covariate. Common haplotype frequency was estimated using the expectation-maximization algorithm in the EH software program(47) in both cases and controls. To assess the association between CRP genetic variation and SLE clinical characteristics, we performed either analysis of variance (ANOVA) for quantitative clinical characteristics of SLE (C3, C4, SLAM and SLICC) or logistic regression analysis for categorical variables (renal disease, joint inflammation, CNS involvement and antiphospholipid antibodies). Covariates adjusted for in the models included age, BMI, and smoking.
To assess the association between CRP genetic variation and serum CRP in SLE subjects, CRP values were log-transformed to achieve the normal distribution of the variable. The mean log-transformed CRP (logCRP) levels between different genotype groups were compared using ANOVA and adjusted for the effects of age, BMI, and smoking. Tests of haplotype association with logCRP were conducted using the haplo.stats package for R(48), with age, BMI, and smoking included as covariates. Haplo.stats tests association by means of a generalized linear-regression framework that incorporates haplotype phase uncertainty by inferring a probability matrix of haplotype likelihoods for each individual (derived by use of the EH haplotype-inference algorithm) rather than by assignment of a most likely haplotype. All computations were performed using R version 2.1.2. A nominal $p$-value of 0.05 was considered significant in all analyses. Analysis results were presented on all SNPs except for one uninformative SNP, -860, due to its extreme low minor allele frequency (0.1% in cases and 0% in controls).

2.4 RESULTS

2.4.1 Association of CRP SNPs with SLE risk

Of the total 785 subjects (337 cases and 448 controls) genotyped for six CRP SNPs, we repeated genotyping on 10% of the subjects for each SNP a second time and had higher than 99% concordance rate in all SNPs. No statistically significant deviations from Hardy–Weinberg equilibrium were found in any of the SNPs. Table 2.1 presents the genotype and allele frequencies in our cases and controls for the six CRP SNPs examined. Allele and genotype
frequencies were not significantly different between cases and controls (using the \( p \)-value of 0.05) in any of the individual six SNPs examined.

### 2.4.2 Association of CRP haplotype with SLE risk

We conducted pair-wise linkage disequilibrium (LD) analysis using four SNPs (excluding the uninformative SNP -860 and the tri-allelic SNP -390), and found different patterns of LD association in cases versus controls. In cases, with the exception of the -861/+90 and +90/+838 pairs, all SNP pairs are in significant LD. Among controls, all pairs were in significant LD except for the –861/+838 pair (Table 2.2). Given the potential effects LD may have on SLE risk, we further assessed the distribution of CRP haplotypes between cases and controls. Five SNPs (-861, -390, +90, +838, +2043) were included in our global haplotype analysis using 222 cases and 313 controls (Table 2.3). A total of 8 haplotypes were observed at a frequency of 2% or greater from either case or control groups. The overall haplotype distribution was significantly different between cases and controls (\( \chi^2 = 138.86, p < 0.000001 \)) (Table 2.3). Haplotype 5 appears to be the most pronounced risk haplotype for SLE while haplotypes 2, 4 and 8 seem to convey protection against SLE. However, since no single allele at any locus defined and was restricted to a given risk or protective overall haplotype, no specific haplotype-tagging SNP could be identified to account for the significant overall haplotype associations.

### 2.4.3 Association of CRP SNPs with SLE Clinical Characteristics

We performed either ANOVA for quantitative clinical characteristics of SLE or logistic regression analysis for categorical variables adjusting for the effects of age, BMI, CRP levels,
smoking and medications specifically correlated with the dependent phenotypes. The risk of
CNS involvement was significantly increased in cases with GC genotype at +838 (OR = 4.7 (1.6
– 13.7); \( p = 0.005 \)). Individuals with +838 GC genotype also exhibited significantly higher C4
levels compared to GG individuals (23.46 ± 8.35 vs. 20.67 ± 7.68, \( p = 0.033 \)). No significant
associations were observed between any of the individual SNPs and SLAM, SLICC, C3,
creatinine, renal disease, joint arthritis, and antiphospholipid antibodies (data not shown).

2.4.4 CRP SNPs associations with serum C-reactive protein levels

We performed both single-site and haplotype analyses to assess the association between CRP
SNPs and log-transformed serum CRP levels (logCRP) in a subgroup of SLE patients (\( n = 237 \)).
In the single-site analyses, minor alleles of two SNPs revealed significant associations with
increased logCRP in SLE patients (+90, \( p = 0.0032 \); -390, \( p = 0.012 \)), and one SNP was
marginally significant (-861, \( P = 0.159 \)) when age, BMI and smoking are controlled for as
covariates. Since BMI was a potential effect modifier between CRP levels and SNP-861,
association test was also performed with BMI removed from the model. Without BMI as a
covariate in the model, association between -861 and logCRP became statistically significant (\( p
= 0.02 \)) (Table 2.4). Heterozygotes of -861 were associated with increased logCRP levels
compared to homozygotes of the wild type allele (T). Homozygotes of the less common allele
(T) at +90 had the highest logCRP level (1.544 ± 1.048) compared to homozygotes of the wild
type allele (0.639 ± 0.976) and heterozygotes (0.623 ± 1.054). Mean logCRP levels were
significantly higher in homozygotes of T allele at the triallelic promoter SNP-390 (1.305 ±
1.128) and heterozygotes with an A allele (CA) (1.356 ± 0.947) when compared to homozygotes
of the wild type (CC) (0.519 ± 0.947) (Table 2.4).
2.4.5 *CRP* haplotype association with serum CRP levels

Given the significant individual effects SNP -861, -390, and +90 have on CRP level, we performed 3-SNP haplotype analysis consisting of these three potentially functional SNP to evaluate the significance of the CRP promoter region has on CRP levels. Three-SNP haplotypes were inferred using the haplo.glm function in the haplo.stats package in R. Haplotype T-861C X C-390TA X A+90T exhibited the most significant impact and associated with an increase of 1.171 logCRP units compared to the reference haplotype \( (p = 0.0161) \). Haplotype T-861C X C-390TA X A+90T also associated with an increase of logCRP by 0.2928 \( (p = 0.0423) \) (Table 2.5).

Haplotypes consisting of all five *CRP* SNPs were then inferred to determine “gene-wide” haplotype effects on CRP levels. Eight haplotypes with frequency greater than or equal to 0.02 were identified (Table 2.6). No individual 5-site haplotype showed associations with CRP levels.

### 2.5 DISCUSSION

We examined the association of *CRP* tagSNPs in relation to SLE risk and CRP levels of SLE patients. Circulating CRP has been the most widely studied inflammatory marker in predicting CVD risks in the past decade. Several studies have demonstrated associations between CRP levels and CVD in healthy population(49-53). Furthermore, recent studies using human CRP-transgenic mice have provided the first direct proof that CRP may be an active participant in thrombosis and atherogenesis (20, 21). Given the importance of inflammation as underlying SLE etiology and overwhelming associations observed between CRP and inflammation, we
hypothesized that variation in the gene encoding for CRP may contribute to the risk of SLE and modify the CRP levels in SLE patients.

An abundance of epidemiological studies and functional studies have investigated the association between CRP levels and SLE. Genetic association studies on CRP variation and SLE and related clinical manifestations are, however, very limited to date (39, 54, 55). Here, we screened six CRP tagSNPs using 337 white SLE women from the Pittsburgh Lupus Cohort and 448 sex and demographically matched controls. Given the reference of the recent family-based study (39), we set out to replicate these associations, but instead took a gene-wide, comprehensive approach by screening four additional CRP tagSNPs using a population-based case-control study design.

Individually none of the examined SNP showed significant association with SLE risk. Therefore we are unable to confirm findings by Russell et al. that the minor (A) allele of +2043 was associated with SLE risk. We also did not observe significant association between decreased CRP with +2043 and +838 as shown in their cohort. Interestingly, the +838 SNP demonstrated significant association with CNS involvement (OR = 4.7 (1.6 – 13.7); p = 0.005) and increased serum C4 (p = 0.033) in our SLE patients. While decreased C4 activity is believed to be a marker of active SLE, the observed elevated C4 levels in SLE patients with CNS involvement most likely result from CNS-specific systemic inflammation. The association observed with CNS involvement, though statistically significant, should be interpreted with caution given the small number of CNS positive patients (n = 23).

In contrast to the single-site analysis, the CRP haplotype analyses yielded significant associations with SLE risk. The global 5-site CRP haplotype distribution was remarkably different between cases and controls (p < 0.000001), presenting strong evidence of association
with SLE risk. However no haplotype-tagging SNP was found to explain the significant haplotype association with SLE risk. The absence of haplotype-tagging SNP confirmed our individual SNP analysis that no polymorphisms we examined in the CRP gene individually alter SLE susceptibility. The observed significant haplotype association in the absence of individual SNP association may be explained by the unique characteristic haplotype-based analyses offer on detecting unique chromosomal segments that harbor multiple disease-predisposing alleles. Further, the use of multilocus analyses in the SNP setting can improve the information content of genomic regions (56) and capture effects from higher number of polymorphisms (versus single SNP analysis) and their subtle interaction effects (epistasis) (57) within the given haplotype block. Even though the individual SNP approach has been the gold standard for association studies for many years, it was designed to easily detect polymorphisms with significant individual genotypic effect (OR > 3) on single-gene disorders that follow a Mendelian inheritance. Given the polygenic and multifactorial nature of SLE pathogenesis, the haplotype approach may be more sensitive and accurate at detecting genotype-phenotype associations in comparison to the individual SNP approach.

It remains a possibility that CRP itself does not directly contribute to SLE susceptibility, rather one or more as yet unidentified functional alleles may be in strong LD with one or more of the SNPs we examined. These functional alleles are likely to be located in one or more nearby genes, thereby “tag” the CRP haplotypes observed to be associated with SLE risk in our cohort. Two SLE susceptibility genes that also mapped to 1q23, FcγRIIA and FcγRIIIA, are genes encoding low-affinity receptors for IgG. Recent meta-analyses revealed that the FcγRIIA-R/H131 polymorphism was associated with a 1.3-fold greater risk of development of lupus, and that the FcγRIIA-V/F158 polymorphism conferred 1.4-fold risk for developing lupus nephritis
The interaction of immunoglobulin (Ig)G Fc receptors containing an activation motif (ITAM) with immune complexes and cytotoxic autoantibodies can initiate an inflammatory response leading to tissue damage (59). It has been also demonstrated that \( Fc_γRIIA-R/H131 \), working in conjunction with CRP, has the unique ability to alter the cytokine profile of the host (60) by mediating phagocytosis (61), and contributing to the impaired removal of circulating immune complexes (62), resulting in the antibody-triggered inflammation and disease pathogenesis of SLE and nephritis. Given the overlapping chromosomal position of the human CRP, \( Fc_γRIIA \) and \( Fc_γRIIIA \) genes and their unique ability to modify SLE phenotype when working together, it is likely that genetic interaction between these three loci (epistasis) may modify SLE susceptibility.

Using 237 SLE women, mean logCRP levels were significantly higher in homozygotes of T allele at the triallelic promoter SNP-390 (1.305 ± 1.128) and heterozygotes with an A allele (CA) (1.356 ± 0.947) when compared to homozygotes of the wild type (CC) (0.519 ± 0.947, \( p = 0.12 \)). Homozygotes of the less common allele (T) at +90 were also associated with more than two-fold increased in logCRP (1.544 ± 1.048) compared to homozygotes of the wild type allele (0.639 ± 0.976) and heterozygotes (0.623 ± 1.054, \( p = 0.0032 \))(Table 4). In fact, these increased CRP level association observed with these two individual SNPs have also been reported by studies of healthy population (63-65).

Polymorphisms located in gene promoters likely play an important role in gene function by altering transcription factor identification and binding, which in turn can influence gene expression and affect biological impacts. Promoter SNP -390 is a tri-allelic SNP that forms an E-box element (66) in the promoter, and its minor alleles (T and A) were found to be associated with increased CRP levels in the non-SLE populations (63, 64, 66, 67). Similarly, SNP +90,
which is located on the intron/exon boundary and has the potential to affect alternative splicing in the gene, also showed association with increased CRP levels (63). Previously Russell et al. (39) reported significant association between SNPs +1059, +2147 and decreased CRP level in British SLE cohort. Similarly, a more recent study by Miller et al. (63) also reported the same association of these two SNPs in three large cohorts of healthy general population. However, we did not observe the same association with decreased CRP in our SLE sample. The lack of association in our SLE women may be attributed to the limited sample size of the minor allele carriers in our study, or it may be confounded by the effects from anti-inflammatory medications SLE patients take on the regular basis, like corticosteroids.

To further understand the gene effects on CRP levels, we performed both promoter region (3-SNP) haplotype and global (5-SNP) haplotype analyses. Our 3-site haplotype analysis included two promoter SNPs and the intron/exon boundary SNP, revealed again and confirmed the importance of the minor alleles -390T and +90T. Both haplotypes containing -390T and +90T, H01 (T-861C X C-390TA X A+90T) and H4/5 (T-861C X C-390TA X A+90T), was associated with significantly increased logCRP levels compared to the referent 3-site haplotype (T-861C X C-390TA X A+90T). When all 5 tagSNPs were examined together as haplotypes, the association to CRP levels became non-significant. This “diluted-out” association with circulating CRP using the global haplotype approach, is however not entirely surprising.

Determining the true association between genetic variation and CRP levels is inherently difficult due to the complex mechanism of CRP production, which is activated by cytokines IL-6 and IL-1 and influenced by multiple other genes and environmental factors (68). The difficulty is compounded in SLE cohort because not only are the inflammatory cytokines found to be increased in SLE patients (69), but the strong correlation observed between CRP and IL-6 levels
in healthy subjects may be absent in SLE (70). SLE is a chronic-inflammatory disease with abnormal expression of CRP during both the presence and absence of acute infections. Multiple studies have also found inconsistent correlations between CRP levels and SLE disease activity, indicating that the mechanism influencing CRP expression in SLE may be different from those in the general population. Our data showed that even though the individual SNP and promoter region haplotype analysis revealed similar pattern of association as shown in general population, no evidence of gene-wide haplotype association with circulating CRP levels could be determined as previously showed using large healthy cohorts (64). This finding further supports the hypothesis that mechanism influencing CRP expression in SLE may be different from those in the general population.

Russell et al.’s family-based study proposed that low levels of basal CRP may predispose to antinuclear autoantibody production, which in turn contributes to the development of human lupus. Our results show that individually, certain SNPs correlated with CRP levels, their association with SLE risk was not significant. Conversely, we found multiple global haplotypes that correlated highly with SLE risk, but had no direct associations with CRP level in SLE subjects. Although we did not find strong evidence that any of the CRP global haplotypes influence CRP level and thereby predict SLE risk, the significant haplotype results suggest that variation in the CRP gene modifies SLE risk via as yet unidentified mechanisms. Determination of how CRP variation influences SLE risk is expected to further our understanding of SLE etiology and may have direct clinical relevance.
2.6 ACKNOWLEDGEMENTS

The authors thank Dr. Christopher Carlson from the University of Washington for his assistance in the triallelic SNP -390 genotyping protocol. This study was supported in part by NHLBI grants HL54900 and HL74165.
### 2.7 MANUSCRIPT 1 TABLES AND FIGURES

**Table 2.1.** Genotype and allele frequencies of CRP SNPs

<table>
<thead>
<tr>
<th>SNP</th>
<th>Genotype</th>
<th>SLE cases n (%)</th>
<th>Controls n (%)</th>
<th>p-value</th>
<th>Allele</th>
<th>SLE cases n (%)</th>
<th>Controls n (%)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>-861 (rs3093059)</td>
<td>TT</td>
<td>287 (85.93)</td>
<td>388 (86.8)</td>
<td>0.813</td>
<td>T</td>
<td>621 (93)</td>
<td>834 (93.3)</td>
<td>0.802</td>
</tr>
<tr>
<td></td>
<td>TC</td>
<td>47 (14.07)</td>
<td>58 (12.98)</td>
<td></td>
<td>C</td>
<td>47 (7)</td>
<td>60 (6.7)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CC</td>
<td>0 (0)</td>
<td>1 (0.22)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-860 (rs3093060)</td>
<td>GG</td>
<td>333 (99.7)</td>
<td>447 (100)</td>
<td>0.428</td>
<td>G</td>
<td>667 (99.9)</td>
<td>894 (100)</td>
<td>0.428</td>
</tr>
<tr>
<td></td>
<td>GA</td>
<td>1 (0.3)</td>
<td>0 (0)</td>
<td></td>
<td>A</td>
<td>1 (0.1)</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AA</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-390 (rs3091244)</td>
<td>CC</td>
<td>92 (38)</td>
<td>135 (41)</td>
<td>0.849</td>
<td>C</td>
<td>289 (59)</td>
<td>416 (61)</td>
<td>0.565</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>88 (36)</td>
<td>118 (36)</td>
<td></td>
<td>T</td>
<td>161 (33)</td>
<td>196 (31)</td>
<td>0.451</td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>29 (12)</td>
<td>30 (9)</td>
<td></td>
<td>A</td>
<td>36 (7)</td>
<td>52 (8)</td>
<td>0.789</td>
</tr>
<tr>
<td></td>
<td>CA</td>
<td>17 (7)</td>
<td>28 (8)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TA</td>
<td>15 (6)</td>
<td>18 (5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AA</td>
<td>2 (1)</td>
<td>3 (1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+90 (rs1417938)</td>
<td>AA</td>
<td>117 (48.75)</td>
<td>159 (47.89)</td>
<td>0.973</td>
<td>A</td>
<td>335 (69.8)</td>
<td>461 (69.4)</td>
<td>0.895</td>
</tr>
<tr>
<td></td>
<td>AT</td>
<td>101 (42.08)</td>
<td>143 (43.07)</td>
<td></td>
<td>T</td>
<td>145 (30.2)</td>
<td>203 (30.6)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>22 (9.17)</td>
<td>30 (9.04)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+838 (rs1800947)</td>
<td>GG</td>
<td>283 (83.98)</td>
<td>395 (88.17)</td>
<td>0.125</td>
<td>G</td>
<td>619 (91.8)</td>
<td>840 (93.8)</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>GC</td>
<td>53 (15.73)</td>
<td>50 (11.16)</td>
<td></td>
<td>C</td>
<td>55 (8.2)</td>
<td>56 (6.3)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CC</td>
<td>1 (0.3)</td>
<td>3 (0.67)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+2043 (rs1205)</td>
<td>GG</td>
<td>142 (42.51)</td>
<td>207 (46.31)</td>
<td>0.538</td>
<td>G</td>
<td>441 (66)</td>
<td>607 (67.9)</td>
<td>0.434</td>
</tr>
<tr>
<td></td>
<td>GA</td>
<td>157 (47.01)</td>
<td>193 (43.18)</td>
<td></td>
<td>A</td>
<td>227 (34)</td>
<td>287 (32.1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AA</td>
<td>35 (10.48)</td>
<td>47 (10.51)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Table 2.2. Pairwise Linkage Disequilibrium between CRP SNPs

<table>
<thead>
<tr>
<th></th>
<th>+90</th>
<th>+838</th>
<th>+2043</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pairwise Linkage Disequilibrium - SLE</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-861</td>
<td>0.018</td>
<td>0.993</td>
<td>0.880</td>
</tr>
<tr>
<td></td>
<td>(0.945)</td>
<td>(0.037)</td>
<td>(&lt; 0.001)</td>
</tr>
<tr>
<td>+90</td>
<td>0.066</td>
<td>0.234</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.511)</td>
<td>(0.016)</td>
<td></td>
</tr>
<tr>
<td>+838</td>
<td>0.846</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(&lt; 0.001)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

|              |         |         |         |
| **Pairwise Linkage Disequilibrium - Controls** |         |         |         |
| -861         | 0.996   | 0.482   | 0.996   |
|              | (< 0.001)| (0.319)| (< 0.001)|     |
| +90          | 0.687   | 0.952   |         |
|              | (0.002) |         |         |
| +838         | 0.753   |         |         |
|              | (< 0.001)|         |         |
### Table 2.3. CRP Haplotype Case-Control Comparison

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>-861 (T&gt;C)</th>
<th>-390 (C&gt;T&gt;A)</th>
<th>+90 (A&gt;T)</th>
<th>+838 (G&gt;C)</th>
<th>+2043 (G&gt;A)</th>
<th>SLE Frequency (n = 222)</th>
<th>Control Frequency (n = 313)</th>
<th>Frequency Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>T</td>
<td>C</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>0.330</td>
<td>0.304</td>
<td>0.026</td>
</tr>
<tr>
<td>H2</td>
<td>T</td>
<td>C</td>
<td>A</td>
<td>G</td>
<td>A</td>
<td>0.183</td>
<td>0.259</td>
<td>-0.076</td>
</tr>
<tr>
<td>H3</td>
<td>T</td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>A</td>
<td>0.047</td>
<td>0.050</td>
<td>-0.004</td>
</tr>
<tr>
<td>H4</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>G</td>
<td>0.201</td>
<td>0.286</td>
<td>-0.085</td>
</tr>
<tr>
<td>H5</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>G</td>
<td>A</td>
<td>0.061</td>
<td>0.002</td>
<td>0.058</td>
</tr>
<tr>
<td>H6</td>
<td>T</td>
<td>A</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>0.038</td>
<td>0.011</td>
<td>0.027</td>
</tr>
<tr>
<td>H7</td>
<td>C</td>
<td>C</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>0.032</td>
<td>0.004</td>
<td>0.028</td>
</tr>
<tr>
<td>H8</td>
<td>C</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>G</td>
<td>0.007</td>
<td>0.061</td>
<td>-0.054</td>
</tr>
</tbody>
</table>

Overall p < 0.000001
\( \chi^2 = 138.86 \)
Table 2.4. Association of CRP Polymorphisms and Mean logCRP level (± SD)

<table>
<thead>
<tr>
<th>SNP</th>
<th>Genotype</th>
<th>n (%)</th>
<th>Mean ± S.D.</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>-861*</td>
<td>TT</td>
<td>191 (84.14)</td>
<td>0.718 ± 1.013</td>
<td>0.16</td>
</tr>
<tr>
<td>(rs3093059)</td>
<td>TC</td>
<td>36 (15.86)</td>
<td>0.977 ± 1.103</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CC</td>
<td>0 (0)</td>
<td>…</td>
<td></td>
</tr>
<tr>
<td>-861* **</td>
<td>TT</td>
<td>191 (84.14)</td>
<td>0.685 ± 1.072</td>
<td>0.021</td>
</tr>
<tr>
<td>(rs3093059)</td>
<td>TC</td>
<td>36 (15.86)</td>
<td>1.141 ± 1.115</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CC</td>
<td>0 (0)</td>
<td>…</td>
<td></td>
</tr>
<tr>
<td>-390*</td>
<td>CC</td>
<td>50 (34.48)</td>
<td>0.519 ± 0.947</td>
<td>0.012</td>
</tr>
<tr>
<td>(rs3091244)</td>
<td>CT</td>
<td>55 (37.93)</td>
<td>0.589 ± 0.899</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>19 (13.10)</td>
<td>1.305 ± 1.128</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CA</td>
<td>11 (7.59)</td>
<td>1.356 ± 0.947</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TA</td>
<td>8 (5.52)</td>
<td>0.696 ± 0.992</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AA</td>
<td>2 (1.38)</td>
<td>0.515 ± 0.843</td>
<td></td>
</tr>
<tr>
<td>+90*</td>
<td>AA</td>
<td>66 (45.83)</td>
<td>0.639 ± 0.976</td>
<td>0.0032</td>
</tr>
<tr>
<td>(rs1417938)</td>
<td>AT</td>
<td>63 (43.75)</td>
<td>0.623 ± 0.932</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>15 (10.42)</td>
<td>1.544 ± 1.048</td>
<td></td>
</tr>
<tr>
<td>+838*</td>
<td>GG</td>
<td>193 (83.91)</td>
<td>0.703 ± 1.043</td>
<td>0.373</td>
</tr>
<tr>
<td>(rs1800947)</td>
<td>GC</td>
<td>37 (16.09)</td>
<td>0.869 ± 0.979</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CC</td>
<td>0 (0)</td>
<td>…</td>
<td></td>
</tr>
<tr>
<td>+2043*</td>
<td>GG</td>
<td>100 (43.67)</td>
<td>0.845 ± 1.117</td>
<td>0.207</td>
</tr>
<tr>
<td>(rs1205)</td>
<td>GA</td>
<td>107 (46.72)</td>
<td>0.605 ± 0.949</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AA</td>
<td>22 (9.61)</td>
<td>0.857 ± 0.999</td>
<td></td>
</tr>
</tbody>
</table>

* Impact of CRP Polymorphisms on Mean logCRP level (± SD) - Adjusted for Age, BMI, and Smoking
** Impact of CRP Polymorphisms on Mean logCRP level (± SD) - Adjusted for Age and Smoking
Table 2.5. Association of CRP Haplotype with Serum logCRP Levels in SLE

### 3-Loci Promotor Haplotype

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Frequency</th>
<th>Coefficient</th>
<th>SE</th>
<th>t-stat</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>...</td>
<td>-1.651</td>
<td>0.364</td>
<td>-4.536</td>
<td>0.000</td>
</tr>
<tr>
<td>Age</td>
<td>...</td>
<td>0.014</td>
<td>0.007</td>
<td>2.064</td>
<td>0.040</td>
</tr>
<tr>
<td>BMI</td>
<td>...</td>
<td>0.048</td>
<td>0.010</td>
<td>4.887</td>
<td>0.000</td>
</tr>
<tr>
<td>Smoke</td>
<td>...</td>
<td>0.394</td>
<td>0.137</td>
<td>2.878</td>
<td>0.004</td>
</tr>
<tr>
<td>H7</td>
<td>C</td>
<td>0.050</td>
<td>0.028</td>
<td>0.273</td>
<td>0.103</td>
</tr>
<tr>
<td>H01</td>
<td>C</td>
<td>0.019</td>
<td>1.171</td>
<td>0.483</td>
<td>2.425</td>
</tr>
<tr>
<td>H6</td>
<td>T</td>
<td>0.074</td>
<td>0.263</td>
<td>0.246</td>
<td>1.072</td>
</tr>
<tr>
<td>H02</td>
<td>T</td>
<td>0.024</td>
<td>-0.158</td>
<td>0.384</td>
<td>-0.411</td>
</tr>
<tr>
<td>H4/5</td>
<td>T</td>
<td>0.301</td>
<td>0.293</td>
<td>0.143</td>
<td>2.042</td>
</tr>
<tr>
<td>H_other*</td>
<td>*</td>
<td>0.009</td>
<td>0.648</td>
<td>0.682</td>
<td>0.950</td>
</tr>
<tr>
<td>H1</td>
<td>T</td>
<td>0.522</td>
<td>Referent</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

**NOTE.** -- t statistics and p values were calculated from the coefficients and SEs within the best-fit multivariate model by the haplo.glm function in the haplo.stats R package.

* Haplotypes with frequency <2% were pooled as "H_other."

### Global Haplotype

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Frequency</th>
<th>Coefficient</th>
<th>SE</th>
<th>t-stat</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>...</td>
<td>-1.611</td>
<td>0.402</td>
<td>-4.010</td>
<td>0.000</td>
</tr>
<tr>
<td>Age</td>
<td>...</td>
<td>0.013</td>
<td>0.007</td>
<td>1.925</td>
<td>0.056</td>
</tr>
<tr>
<td>BMI</td>
<td>...</td>
<td>0.049</td>
<td>0.010</td>
<td>4.952</td>
<td>0.000</td>
</tr>
<tr>
<td>Smoke</td>
<td>...</td>
<td>0.356</td>
<td>0.140</td>
<td>2.540</td>
<td>0.012</td>
</tr>
<tr>
<td>H7</td>
<td>C</td>
<td>0.0404</td>
<td>0.144</td>
<td>0.352</td>
<td>0.409</td>
</tr>
<tr>
<td>H6</td>
<td>T</td>
<td>0.0386</td>
<td>0.309</td>
<td>0.359</td>
<td>0.860</td>
</tr>
<tr>
<td>H3</td>
<td>T</td>
<td>0.0407</td>
<td>0.117</td>
<td>0.349</td>
<td>0.334</td>
</tr>
<tr>
<td>H2</td>
<td>T</td>
<td>0.1656</td>
<td>-0.094</td>
<td>0.189</td>
<td>-0.497</td>
</tr>
<tr>
<td>H5</td>
<td>T</td>
<td>0.0661</td>
<td>0.228</td>
<td>0.257</td>
<td>0.890</td>
</tr>
<tr>
<td>H4</td>
<td>T</td>
<td>0.2152</td>
<td>0.319</td>
<td>0.216</td>
<td>1.478</td>
</tr>
<tr>
<td>H_other*</td>
<td>*</td>
<td>0.1157</td>
<td>0.349</td>
<td>0.217</td>
<td>1.610</td>
</tr>
<tr>
<td>H1</td>
<td>T</td>
<td>0.3178</td>
<td>Referent</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

**NOTE.** -- t statistics and p values were calculated from the coefficients and SEs within the best-fit multivariate model by the haplo.glm function in the haplo.stats R package.

* Haplotypes with frequency <2% were pooled as "H_other."
2.8 MANUSCRIPT 1 REFERENCES


18. Kaplan MH, Volanakis JE. Interaction of C-reactive protein complexes with the complement system. I. Consumption of human complement associated with the reaction of C-reactive protein with pneumococcal C-polysaccharide and with the choline phosphatides, lecithin and sphingomyelin. Journal of Immunology 1974;112(6):2135-47.


53


3.0 MANUSCRIPT 2. EVIDENCE FOR HAPLOTYPE-BASED ASSOCIATION IN SLE AT THE C-REACTIVE PROTEIN LOCUS: A FAMILY-BASED ASSOCIATION STUDY

Manuscript in Preparation

P. Betty Shih¹,², Hui Wu², Candace Kammerer¹, Marisa C. Mizus², Christopher S. Carlson³, Jennifer M. Grossman², Bevra H. Hahn², M. Ilyas Kamboh¹ Betty P. Tsao²
¹ Graduate School of Public Health, University of Pittsburgh, PA; ² Department of Medicine, Division of Rheumatology, University of California David Geffen School of Medicine, Los Angeles, CA; ³ Department of Genome Sciences, University of Washington, Seattle, WA
3.1 ABSTRACT

C-reactive protein gene (CRP) maps on 1q23, a genomic region shown strong evidence in harboring lupus susceptibility genes on distal mouse chromosome 1 and its syntenic human counterpart. Recently we have found evidence of CRP as a susceptibility gene for systemic lupus erythematosus (SLE) using a population-based case-control study. Here we aim to use a family-based population to test and confirm the evidence of association. Five single nucleotide polymorphisms (SNPs) in CRP were genotyped in 205 UCLA Genetic Study trio families. Of the five SNPs, none individually was significantly associated with SLE risk. However, minor alleles of promoter SNP-390 (T) and intron/exon boundary SNP+90 (T) showed marginal association with the disease risk ($P = 0.05$ and 0.06, respectively). Haplotype analyses further demonstrated that any combination of CRP haplotypes containing +390T and +90T showed evidence of significant overtransmission to the SLE probands in our 205 trios ($P = 0.01 – 0.06$). Our findings, in conjunction with previous case-control study, suggest CRP to be a likely susceptibility gene for human systemic lupus erythematosus. Future molecular studies of CRP gene and sequencing of additional loci in the 1q23 region will help identify the likely genetic epistasis effect contributing to this complex polygenetic disease.
3.2 INTRODUCTION

Systemic lupus erythematosus (SLE) is a complex polygenic autoimmune disease characterized by production of autoantibodies against different autoantigens. The chronic inflammation resulting from the hyperactive immune system coupled with immune complex deposition results in a wide range of symptoms, from mild arthritis to life-threatening multiple organ systems involvement. Modern therapeutic strategies have improved the disease prognosis significantly compared to 20 years ago. However, treatments are generally immunosuppressive, which can result in devastating side effects such as serious infections, and some patients are not responsive to these aggressive therapies. Given the well established genetic- and environmental-contributions to the development of SLE (1, 2), early identification of individuals with genetic susceptibility of SLE may enable health care professionals to provide timely preventative intervention such as lifestyle management to prevent or delay the onset of disease and the subsequent organ damage.

The genetic basis for SLE susceptibility is complex, possibly involving multiple genes and their interactions (epistasis) to contribute to SLE or its many subsets defined by clinical and laboratory features (3). Individuals carrying the susceptibility loci increase the probability of disease onset upon further environmental exposures that modify these genetic effects. Several independent genome scans in multiplex SLE families have identified a number of genomic regions that may harbor susceptibility genes (2). Amongst the various candidate genes mapping within intervals of these genomic regions, the well studied genes to date include HLA-DR, Fc receptor cluster (FCGR2A, FCGR3A), PARP and PDCD1. The gene encoding C-reactive protein (CRP) maps on 1q23, one of the multiple susceptibility loci that are syntenic between
human SLE and murine SLE models in linkage studies (2), making it a logical positional candidate gene for SLE association study.

CRP is a prototypic acute phase protein in humans that increases rapidly in circulating concentration after an acute stimulus, such as infection, tissue damage or inflammation, making CRP serve a gold-standard biomarker of systemic inflammation. Therefore CRP also serves as an important functional candidate gene for SLE for inflammatory cascade is believed to be chronically “activated” in patients with active disease. However, CRP has been found to be non-remarkable in patients with active disease (4, 5) and non-responsive in patients with evidence of infection (6, 7). Therefore, the reliability of circulating CRP as a diagnostic biomarker of disease activity and infection in SLE patients remains questionable. Given the abnormal behavior of this acute phase protein in SLE patients and its important ability to modulate immunity and inflammation cascade, we aim to determine if the genetic variation in the CRP gene may contribute to SLE susceptibility. Here, we use 205 Caucasian SLE trios from the Southern California region to assess whether polymorphisms in CRP, a highly sensitive inflammatory marker, are associated with the risk of SLE development.

3.3 SUBJECTS AND METHODS

3.3.1 Human Subjects and Data Collection.

Two hundred and five Caucasian SLE complete trios consisting of mother-father and affected child from the University of California, Los Angeles (UCLA) Lupus Genetic Study (8) were enrolled in this study. Of the 205 trios, 138 were recruited from the metropolitan areas of Los
Angeles; and 67 from Columbus, OH. The mean age for the 205 SLE probands was 43.16. Ninety percent of SLE probands were female (n = 186) and ten percent were male (n = 19). All probands met the American College of Rheumatology criteria for the SLE classification (9). Of the 205 probands, thirty six percent (n = 74) had evidence of glomerulonephritis. Of the 410 parents of affected probands in this study, 5.3 percent (n = 22 (19 female/ 3 male)) also had the diagnosis of SLE. All participants gave written informed consent. Approval for human study protocols were obtained from the human subjects review boards at University of California, Los Angeles. Genomic DNA was purified from peripheral blood mononuclear cells.

3.3.2 SNP selection and Genotyping.

CRP tagSNPs were chosen from the SeattleSNPs Program for Genomic Applications (http://pga.gs.washington.edu/) European population, using the LDSelect algorithm (10) using a linkage disequilibrium threshold of $r^2 = 0.64$ and a minor allele frequency threshold of five percent. Five tagSNPs were selected relative to the ATG codon of the CRP translation site in the FASTA database (Genbank NC_000001.9). The positions of these SNPs are -861, -390, +90, +838 and +2043. SNPs -861 and -390 positioned in the promoter region, +90 located in intron/exon boundary, +838 positioned within exon 2, and + 2043 in the 3’ untranslated region. For clarification, reference numbers from NCBI single nucleotide polymorphism are provided for each of the five SNPs: -861 is rs3093059, -390 is rs3091244, +90 is rs1417938, +838 is rs1800947 and +2043 is rs1205. Genotyping for -861 was carried out using Pyrosequencing assay on PSQ 96MA system. SNPs -390, +90, +838 and +2043 were genotyped using TaqMan SNP genotyping assays under standard conditions (http://www.appliedbiosystems.com/), except for the triallelic tagSNP -390, which was genotyped using methods previously described (11).
3.3.3 **Statistical and In Silico Analysis.**

The family-based association test (FBAT, ver.1.7.2) software was used to determine evidence of association between the *CRP* polymorphisms and SLE risk, using both individual SNPs and haplotypes (12). The FBAT statistic is an extension of the original transmission disequilibrium Test (TDT) (13). Statistics for individual SNPs were calculated under an additive risk model, and for both the bi-allelic (each allele against all others) and multi-allelic (all alleles at a marker were compared simultaneously in one test) mode of testing. Haplotype statistics were tested using individual haplotype test and global haplotype test. Genehunter was utilized to confirm the FBAT statistics and to estimate copies of transmitted versus untransmitted alleles for each individual marker and haplotypes (GENEHUNTER, ver.2.0). Cohort summary statistics, linkage disequilibrium statistics of the four bi-allele SNPs, and Hardy-Weinberg equilibrium (HWE) tests were carried out using the R statistics program ver.2.0.1. The genotype and allele distribution of the tagSNPs in probands with and without glomerulonephritis were compared using Fisher’s Exact test. Mendelian inconsistencies were evaluated by FBAT. Mendelian error rates were 13 families for SNP -861, 28 for SNP -390, 8 for SNP +90, 6 for SNP +838 and 8 for SNP +2043. These Mendelian errors were excluded in the analysis. Promoter SNP analysis for transcription factor element binding was performed using TFSEARCH ver.1.3 (http://www.cbrc.jp/research/db/TFSEARCH.html), while splice site prediction was performed using Berkeley Drosophila Genome Project (BDGP) (http://www.fruitfly.org/seq_tools/splice.html).
3.4 RESULTS

3.4.1 Gene structure, Hardy-Weinberg equilibrium, and linkage disequilibrium in CRP

CRP is a relatively small gene with approximately 6,800 base pairs. The gene contains two exons (black boxes in Figure 1) with a 104 amino acid peptide coding regions. The gene contains a short 5'-untranslated region (89 bp, grey box), with two exons separated by a short intron region (286 bp), and an unexpectedly longer (1.2 kb) 3'-untranslated region (grey box). Two of the 5 SNPs locate on the promoter regions (SNPs -861 and -390), SNP +90 locates on the intron and exon border, SNP +838 is an synonymous coding SNP in exon 2, and +2043 loates on 3'-untranslated region (Figre 1). All 5 SNPs examined were in HWE in probands alone, parents alone or combined. Linkage disequilibrium test of the founder (parents) showed that all SNPs were in strong LD with each other (D' >0.96) (Table 3.1). CRP polymorphisms genotype distributions for 205 probands in this study and 337 caucasain SLE cases from our previous study for comparison purpose (Table 3.2).

3.4.2 Marginal over-transmission of -390T and +90T alleles

Of the five SNPs examined in this family-based analysis, we found evidence of marginal over-transmission to SLE probands on the minor allele (T) of the triallelic SNP -390 (Z = 1.947, P = 0.051) and the minor allele (T) of the intron/exon border SNP +90 (Z = 1.819, P = 0.068) (Table 3.3). SNP -390 (T) was transmitted 88 times from a heterozygous parent to the index probands, and remained untransmitted 64 times. SNP +90 (T) was transmitted 99 times from a heterozygous parent to the index probands, and remained untransmitted 75 times. Multiallelic
analysis produced a global chi-square value of 3.96 ($P = 0.138$) for SNP -390, and chi-square value of 3.31 ($P = 0.068$) for SNP +90. FBAT analysis failed to show evidence of significant over-transmission on minor alleles of SNP-861 ($P = 0.896$), +838 ($P = 768$), and +2043 ($P = 1$) to our SLE probands (Table 3.3).

### 3.4.3 Significant “Mini-Haplotype” Associations with SLE risk

*CRP* polymorphisms were analyzed in multiple combinations to assess effects of “mini-haplotypes” in *CRP* gene on SLE risk. We termed these combinations “mini-haplotype” to differentiate them from the gene-wide haplotypes consisting of all five tagSNPs. Our results showed evidence of overtransmission of 2-SNP haplotypes -390T → +90T (72 transmitted/48 untransmitted, $p = 0.028$) and +90T → +838G (80 transmitted/53 untransmitted, $p = 0.043$). Of the 3-SNP haplotypes, two haplotypes containing minor alleles at -390T and +90T showed overtransmission to SLE probands: -861T → -390T → +90T (64 transmitted/44 untransmitted, $p = 0.029$) and -390T → +90T → +838G (72 transmitted/45 untransmitted, $p = 0.013$). The 4-SNP haplotype analysis revealed the same pattern of haplotype-tag effect from the combination of -390T and +90T alleles: -861T → 390T → +90T → +838G ($p = 0.042$) and 390T → +90T → +838G → +2043G ($p = 0.038$). In summary, all mini-haplotypes containing minor alleles at +390 and +90 showed evidence of overtransmission to the SLE probands in our 205 trios.

### 3.4.4 Marginal Global Haplotype association with SLE risk

When all 5 SNPs were included for gene-wide haplotype analysis, they formed five common haplotypes ($\geq 1\%$), accounting for 96.2% of the founder chromosomes in the population studied.
(Table 3.4). The combination of -390T and +90T alleles occurred only on the most frequent haplotype (H1), same as the mini-haplotype results, this haplotype also showed an increased transmission rate to SLE proband (54 transmitted/39 untransmitted, p = 0.068) in the individual haplotype analysis. The global haplotype analysis revealed similar finding of $\chi^2$ statistic of 10.764 with the p-value of 0.056. These results are consistent with the single SNP analyses, as the high risk haplotype (h1) is tagged by the high risk alleles -390T and +90T.

### 3.4.5 Association of CRP SNPs with SLE glomerulonephritis

We also examined genotype and allele distribution between SLE with (n = 74) and without (n = 131) the diagnosis of glomerulonephritis. The allele and genotype frequencies between these two groups did not differ significantly (data not shown).

### 3.5 DISCUSSION

C-reactive protein serves several important roles in the human innate immune system, including agglutination, bacterial capsular swelling, phagocytosis, and complement activation/regulation (14). These varied biologic functions of CRP as an acute phase protein distinguish it as a remarkable marker of non-specific systemic inflammation in healthy individuals. In SLE, CRP appears to serve an intriguing dual role, both as a contributor of inflammation which results in premature coronary artery disease in patients (15), and as a protective factor for SLE associated renal involvement in mice models (16, 17).
CRP has been shown to alter the clearance of the SLE autoantigen chromatin (18) and histones (19), to bind to the apoptotic cells surface (20), and to suppress acute inflammation via possible cytokine cascade regulation (21) and/or IL-10 synthesis (22), all of which are consistent with the protective role of CRP. These activities are thought to be necessary for clearing of the apoptotic debris, a process that is believed to be defective in patients with SLE. However, we still do not understand how circulating CRP impacts SLE pathogenesis. The fascinating opposing roles CRP plays in patients with SLE may be better understood by investigating the blueprint coding this protein – the \textit{CRP} gene. Here we examined association of \textit{CRP} polymorphisms with SLE, and our results suggest that two \textit{CRP} SNPs might contribute to the risk of SLE development.

We have previously identified evidence supporting linkage to SLE in chromosome region 1q23 using 238 individuals from 62 multiplex, multiethnic SLE families (23). The \textit{CRP} gene maps directly to the 1q23 region, which also contains a cluster of four genes that encode low-affinity receptors for IgG (\textit{FCGR2A}, \textit{FCGR3A}, \textit{FCGR3B}, and \textit{FCGR2B}). Genetic polymorphisms in these four genes, especially alleles of \textit{FCGR2A} and \textit{FCGR3A}, have been associated with SLE in multiple cohorts (24-27). Various studies have demonstrated the ligand- and allele-dependent differential interaction effects between CRP and various Fc gamma receptors (28-31). Therefore the relative importance of these Fc gamma alleles found to be associated with SLE may depend on not only the IgG subclass of pathogenic autoantibodies in each patient, but also the integrity of the circulating CRP that binds to these receptors. Hence it is reasonable to assume that interactions between \textit{CRP} gene and these neighboring Fc gamma receptor genes (epistasis) play an important role in SLE susceptibility. Before any epistatic
effects can be identified, we must first identify the associations between each candidate locus with the disease risk and patterns of linkage disequilibrium between candidate loci.

In our recent SLE population-based study (32) examining the same five CRP SNPs using 337 Caucasian American patients and 448 sex- and race-matched controls, we found no statistically significant associations between each of the individual SNP with SLE risk \( (P = 0.12 - 0.97) \) but discovered significantly different global haplotype frequencies comparing cases to controls \( (P < 0.000001) \) (32). Here, we also found no evidence of significant associations on the individual SNP analysis \( (P = 0.051 - 1) \). In the global haplotype analysis comparing number of haplotypes transmitted to probands versus number of haplotypes not transmitted, the evidence of SLE association is only marginal \( (P = 0.056) \). However, using a mini-haplotype analysis, we have additionally shown here that the SLE risk likely associate with CRP haplotypes bearing two specific alleles \(-390T\) and \(+90T\) in this particular family-based cohort. All mini-haplotypes containing minor alleles at \(+390\) and \(+90\) showed evidence of overtransmission to the SLE probands in our 205 trios \( (p = 0.012 - 0.043) \).

When comparing global haplotype results between the two cohorts, we were surprised to see that the high risk global haplotype (H1, Table 3.4) in this study actually appears to be protective in the previous case-control study by having lower estimated frequency in cases compared to controls (32). To rule out the possibility of genotyping or other classification errors, we compared the genotype frequencies on each of the five SNPs between SLE probands in this study \( (n = 205) \) and SLE cases from the case-control \( (n = 337) \). There are no significant differences in genotypic frequency between these two studies, indicating the unlikely chance of misclassification.
Studies of cases and unrelated controls design and is known for its disadvantage of prone to confounding due to unaccounted population admixture, which may affect the validity of the obtained results (33). On the other hand, family-based study designs offer the advantages of sharing common genetic background among the family members and more homogeneous environmental exposures associated with the disease risk. Thus, the problem of population stratification and admixture is bypassed (34). The finding that high risk global H1 haplotype in this study appears to be protective is likely to be a result of the population stratification unaccounted for in our previous case-control designs. SLE is known for its complex influence from multiple genes and environmental factors, indicating that genetic effect from each susceptibility gene is modest at the best. Furthermore, phenotype variation in SLE is almost always product of many pleiotropic genes and epistatic genes, further increasing the difficulty of detecting significant effect from one individual gene in multiple independent cohorts.

In fact, the combined results of our two studies support the complex CRP genetic influence on SLE risk. Unlike the recent British SLE family-based study by Russell et al. (35), neither one of our US-based studies found evidence of individual SNP associations, which supports the differential admixture effect on SLE risk. In the absence of individual SNP association, global haplotype analyses revealed evidence of association between CRP and SLE in both study designs \((p = < 0.000001 \text{ and } 0.056)\), in which case-control design showed higher level of statistical significance than the family-based design. Observing a higher degree of association in a more genetically heterogeneous population (case-control design) provides evidence on presence of additional susceptibility genes near by CRP, and supports the epistatic effects of CRP has on SLE. The less significant but still suggestive global haplotype association in the
family-based design suggest a significant role CRP variation plays in the polygenic nature of SLE.

Russell et al. have previously identified over-transmission of the A allele at +2043 locus to SLE probands in their British family-based study of CRP (35). In this study consisting of North American trios, the expected and observed transmission of alleles at +2043 are identical, resulting in a p-value of 1 (Table 3.3). Taken into context together with our previous population-based study, we found no evidence of association to support +2043 as a susceptibility polymorphism in North American Caucasian SLE. Russell et al. had also suggested that since SNP+2043 was found to be associate with a lower base-line CRP levels, ANA production, and SLE risk in their cohort, relative deficiency of CRP (indicated by low CRP level) may predispose to development of SLE (35). It is believed that CRP binds to ribonucleoproteins and polar phospholipid materials released during apoptosis, therefore decreased circulating CRP may interfere with effective clearing of the immune complexes and result in SLE development. This theory, however, dose not support the epidemiological finding in African Americans who have higher circulating levels of CRP (36), and have up to 3 times the risk of developing SLE compared to other racial categories (37). Additionally, CRP level is found to be higher in women compared to men in the general population (38), and this may contribute to the high female to male ratio (10-15:1) of the SLE prevalence (39).

The hypothesis that decreased circulating CRP levels may predispose to increased risk of SLE is further contradicted by a recent study conducted by Miller et al. (14). In their association study of CRP polymorphisms and CRP levels in healthy individuals, minor allele of +2043 was also associated with significantly decreased circulating CRP levels in all three of their large cohorts (Women's Health Study (n = 717), Pravastatin Inflammation/CRP Evaluation trial (n =
and Physicians' Health Study (n = 509)). This strongly suggests that the association between +2043 and decreased CRP levels in Russell’s SLE families is a normal genotype-phenotype association, not likely to be attributed to SLE-related risk factors, or has potential to increase the development of SLE. Perhaps the non-remarkable nature of CRP in SLE patients is a result of some form of yet to be identified defect in the protein integrity, a defect which may not necessarily be reflected in circulating CRP levels.

Our lack of significant association with SNP +2043 does not rule CRP out as a SLE susceptibility gene. In fact, our mini-haplotype analysis found evidence of CRP association with SLE risk on two potentially functional SNPs: a promoter triallelic SNP at -390, and an intron/exon boundary SNP at +90 lying at the 29th base pair downstream of exon 1, both of which have been shown to alter CRP levels(11, 14, 35, 40, 41). We showed that all individual SNPs/mini-haplotypes/global-haplotype that are specific to -390T and -90T alleles are associated with an increase in SLE risk.

We (11, 32) and others (14, 41) have previously found significant association between the -390T allele and increased circulating CRP levels. Using the transcription factor–motif analysis, -390T allele in the CRP promoter showed to form a potential USF binding motif (41). Functional study indeed demonstrated that the -390 SNP resides within the hexameric core of transcription factor binding E-box elements and was associated with increases promoter activity (41). An SLE association with a high-CRP associated SNP may sound contradictory to the theory that CRP “protects” against SLE risk, however the observation that African American women associate with both significantly elevated risk of SLE and increased basal CRP levels should not be ignored. In fact we have previously found that the -390T allele which is associated with increased CRP levels is more frequent in African Americans than Caucasian Americans.
(11). It remains unclear whether increased circulating CRP is a risk factor for SLE risk, or rather an indicator of the chronic systemic inflammatory processes underlying SLE etiology.

SNP +90 resides at 27 basepairs downstream from the end of exon 1. Using the splice site prediction software, we identified a possible alternative splicing site (GT) \textit{in silico} which is 10 basepairs downstream from +90. It is possible that transcription of CRP may not splice out the entire intron between exon 1 and 2, hence leaving SNP +90 as part of the coding sequence in the final protein product which in turn interferes with the integrity of the final protein product. The significance and mechanisms of this potential alternative splicing in SLE remain to be determined.

It remains a strong possibility that either SNPs -390 and +90, or both are in significant LD with one or more functional SLE-risk polymorphism(s), possibly in the Fc gamma receptor genes which were previously found to confer risk for SLE. The interplay between these risk polymorphisms may result in haplotype-specific transmission to individuals who then develop SLE upon exposure to additional environmental injury or trigger. This hypothesis is supported in our Pittsburgh population-based study in which the haplotype distribution was significantly different between cases and controls (32). We are in the process of performing additional characterization of candidate genes in 1q23 as well as linkage disequilibrium analyses between these genes in order to better understand the epistasis effects in the 1q23 region. Better understanding of the gene-gene interactions may shed light on novel pathways in which CRP participates to modulate the risk of SLE.
3.6 ACKNOWLEDGEMENTS

This work was supported in part by the ACR REF/Abbott Health Professional Graduate Student Research Preceptorship Award.
3.7 MANUSCRIPT 2 TABLES AND FIGURES

<table>
<thead>
<tr>
<th></th>
<th>+90</th>
<th>+838</th>
<th>+2043</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>D'</strong></td>
<td>0.9978</td>
<td>0.9869</td>
<td>0.9976</td>
</tr>
<tr>
<td><strong>-861</strong> Corr.</td>
<td>-0.2227</td>
<td>-0.0782</td>
<td>-0.2086</td>
</tr>
<tr>
<td><strong>p-value</strong></td>
<td>&lt;0.0001</td>
<td>0.0328</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>D'</strong></td>
<td></td>
<td>0.9968</td>
<td>0.9704</td>
</tr>
<tr>
<td><strong>+90</strong> Corr.</td>
<td>-0.1773</td>
<td>-0.4557</td>
<td></td>
</tr>
<tr>
<td><strong>p-value</strong></td>
<td>&lt;0.0001</td>
<td>&lt; 0.0001</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>D'</strong></td>
<td></td>
<td></td>
<td>0.9569</td>
</tr>
<tr>
<td><strong>+838</strong> Corr.</td>
<td></td>
<td>0.3624</td>
<td></td>
</tr>
<tr>
<td><strong>p-value</strong></td>
<td></td>
<td>&lt; 0.0001</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.1. Pairwise Linkage Disequilibrium between *CRP* SNPs in founders.
Table 3.2. Genotype frequencies of CRP SNPs – comparison between UCLA cohort and Pitt Cohort

<table>
<thead>
<tr>
<th>SNP</th>
<th>Genotype</th>
<th>UCLA Probands</th>
<th>Pitt SLE Cases</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>-861</td>
<td>TT</td>
<td>158 (0.82)</td>
<td>287 (0.86)</td>
<td>0.2154</td>
</tr>
<tr>
<td></td>
<td>TC</td>
<td>35 (0.18)</td>
<td>47 (0.14)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CC</td>
<td>0 (0.00)</td>
<td>0 (0.00)</td>
<td></td>
</tr>
<tr>
<td>-390</td>
<td>CC</td>
<td>62 (0.33)</td>
<td>92 (0.38)</td>
<td>0.6042</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>68 (0.36)</td>
<td>88 (0.36)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>25 (0.13)</td>
<td>29 (0.12)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CA</td>
<td>20 (0.11)</td>
<td>17 (0.07)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TA</td>
<td>12 (0.06)</td>
<td>15 (0.06)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AA</td>
<td>0 (0.00)</td>
<td>2 (0.01)</td>
<td></td>
</tr>
<tr>
<td>+90</td>
<td>AA</td>
<td>88 (0.44)</td>
<td>117 (0.49)</td>
<td>0.3085</td>
</tr>
<tr>
<td></td>
<td>AT</td>
<td>86 (0.43)</td>
<td>101 (0.42)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>27 (0.13)</td>
<td>22 (0.09)</td>
<td></td>
</tr>
<tr>
<td>+838</td>
<td>GG</td>
<td>176 (0.88)</td>
<td>283 (0.84)</td>
<td>0.1965</td>
</tr>
<tr>
<td></td>
<td>GC</td>
<td>23 (0.11)</td>
<td>53 (0.16)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CC</td>
<td>2 (0.01)</td>
<td>0 (0.00)</td>
<td></td>
</tr>
<tr>
<td>+2043</td>
<td>GG</td>
<td>97 (0.49)</td>
<td>142 (0.43)</td>
<td>0.356</td>
</tr>
<tr>
<td></td>
<td>GA</td>
<td>82 (0.41)</td>
<td>157 (0.47)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AA</td>
<td>20 (0.10)</td>
<td>35 (0.10)</td>
<td></td>
</tr>
</tbody>
</table>
Table 3.3. Single-marker association analysis in the CRP gene

**Bi-allelic association analysis result**

<table>
<thead>
<tr>
<th>Marker</th>
<th>Allele</th>
<th>afreq</th>
<th>Fam&lt;sup&gt;a&lt;/sup&gt;</th>
<th>S&lt;sup&gt;b&lt;/sup&gt;</th>
<th>E(S)&lt;sup&gt;c&lt;/sup&gt;</th>
<th>Var(S)&lt;sup&gt;d&lt;/sup&gt;</th>
<th>Z</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>-861</td>
<td>T</td>
<td>0.909</td>
<td>53</td>
<td>83</td>
<td>82.5</td>
<td>14.75</td>
<td>0.13</td>
<td>0.8964</td>
</tr>
<tr>
<td>-861</td>
<td>C</td>
<td>0.091</td>
<td>53</td>
<td>29</td>
<td>29.5</td>
<td>14.75</td>
<td>-0.13</td>
<td>0.8964</td>
</tr>
<tr>
<td>-390</td>
<td>C</td>
<td>0.59</td>
<td>130</td>
<td>138</td>
<td>150</td>
<td>41.5</td>
<td>-1.863</td>
<td>0.0625</td>
</tr>
<tr>
<td>-390</td>
<td>T</td>
<td>0.323</td>
<td>117</td>
<td>88</td>
<td>83</td>
<td>1.947</td>
<td>0.13</td>
<td>0.8964</td>
</tr>
<tr>
<td>-390</td>
<td>A</td>
<td>0.087</td>
<td>45</td>
<td>25</td>
<td>25</td>
<td>12.5</td>
<td>0</td>
<td>1.0000</td>
</tr>
<tr>
<td>+90</td>
<td>A</td>
<td>0.678</td>
<td>131</td>
<td>155</td>
<td>167</td>
<td>43.5</td>
<td>-1.819</td>
<td>0.0688</td>
</tr>
<tr>
<td>+90</td>
<td>T</td>
<td>0.322</td>
<td>131</td>
<td>113</td>
<td>101</td>
<td>43.5</td>
<td>1.819</td>
<td>0.0688</td>
</tr>
<tr>
<td>+838</td>
<td>G</td>
<td>0.937</td>
<td>44</td>
<td>66</td>
<td>65</td>
<td>11.5</td>
<td>0.295</td>
<td>0.7681</td>
</tr>
<tr>
<td>+838</td>
<td>C</td>
<td>0.063</td>
<td>44</td>
<td>22</td>
<td>23</td>
<td>11.5</td>
<td>-0.295</td>
<td>0.7681</td>
</tr>
<tr>
<td>+2043</td>
<td>G</td>
<td>0.696</td>
<td>134</td>
<td>171</td>
<td>171</td>
<td>45</td>
<td>0</td>
<td>1.0000</td>
</tr>
<tr>
<td>+2043</td>
<td>A</td>
<td>0.304</td>
<td>134</td>
<td>103</td>
<td>103</td>
<td>45</td>
<td>0</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

**Multi-allelic association analysis result**

<table>
<thead>
<tr>
<th>Marker</th>
<th>Allele#</th>
<th>DF</th>
<th>Chisq</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>-861</td>
<td>2</td>
<td>1</td>
<td>0.017</td>
<td>0.8964</td>
</tr>
<tr>
<td>-390</td>
<td>3</td>
<td>2</td>
<td>3.958</td>
<td>0.1382</td>
</tr>
<tr>
<td>+90</td>
<td>2</td>
<td>1</td>
<td>3.31</td>
<td>0.0688</td>
</tr>
<tr>
<td>+838</td>
<td>2</td>
<td>1</td>
<td>0.087</td>
<td>0.7681</td>
</tr>
<tr>
<td>+2043</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

<sup>a</sup> Number of informative families (i.e. families with at least one heterozygous parent).

<sup>b</sup> Statistic test from family–based association test for the observed number of transmitted alleles.

<sup>c</sup> Expected value of S under the null hypothesis (ie no linkage or association).

<sup>d</sup> Empirical variance.
### Table 3.4. Haplotype association analysis of CRP gene

#### Individual haplotype association analysis result

<table>
<thead>
<tr>
<th>Haplotypes</th>
<th>-861</th>
<th>-390</th>
<th>+90</th>
<th>+838</th>
<th>+2043</th>
<th>Haplotype frequency</th>
<th>Fam</th>
<th>S</th>
<th>E(S)</th>
<th>Var(S)</th>
<th>Z</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>h1</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>G</td>
<td>G</td>
<td>0.305</td>
<td>92.800</td>
<td>100.763</td>
<td>90.493</td>
<td>31.863</td>
<td>1.819</td>
<td>0.069</td>
</tr>
<tr>
<td>h2</td>
<td>T</td>
<td>C</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>0.288</td>
<td>101.000</td>
<td>79.237</td>
<td>89.007</td>
<td>33.802</td>
<td>-1.680</td>
<td>0.093</td>
</tr>
<tr>
<td>h3</td>
<td>T</td>
<td>C</td>
<td>A</td>
<td>G</td>
<td>A</td>
<td>0.234</td>
<td>93.900</td>
<td>79.822</td>
<td>76.885</td>
<td>32.463</td>
<td>0.515</td>
<td>0.606</td>
</tr>
<tr>
<td>h4</td>
<td>C</td>
<td>A</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>0.075</td>
<td>38.000</td>
<td>26.000</td>
<td>25.500</td>
<td>11.083</td>
<td>0.150</td>
<td>0.881</td>
</tr>
<tr>
<td>h5</td>
<td>T</td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>A</td>
<td>0.060</td>
<td>36.900</td>
<td>21.941</td>
<td>21.941</td>
<td>9.943</td>
<td>0.000</td>
<td>1.000</td>
</tr>
</tbody>
</table>

* Haplotypes with frequencies < 0.01 were not listed

#### Global haplotype association analysis result

<table>
<thead>
<tr>
<th>Allele#</th>
<th>DF</th>
<th>Chisq</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>5</td>
<td>10.764</td>
<td>0.056</td>
</tr>
</tbody>
</table>


Figure 3.1. CRP gene and position of five TagSNPs
3.8 MANUSCRIPT 2 REFERENCES


18. 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. Clinical Immunology & Immunopathology 1994;70(1):22-7.


35. Russell AI, Cunninghamhame Graham DS, Shepherd C, Robertson CA, Whittaker J, Meeks J, et al. Polymorphism at the C-reactive protein locus influences gene expression and


4.0 MANUSCRIPT 3. PLEIOTROPIC EFFECTS OF C-REACTIVE PROTEIN POLYMORPHISMS INFLUENCING INTERMEDIATE TRAITS OF CARDIOVASCULAR DISEASE RISK IN SYSTEMIC LUPUS ERYTHEMATOSUS

Alternative Title:
Can C-reactive Protein Genetic Variation Serve as Biomarker to Assess Cardiovascular Disease Risk in Systemic Lupus Erythematosus?

Manuscript in Preparation

P. Betty Shih$^{1,2}$, Susan Manzi$^{1,2}$, Penny Shaw$^2$, Amy Kao$^2$, Candace Kammerer$^1$, M. Ilyas Kamboh$^1$

$^1$ Graduate School of Public Health, University of Pittsburgh, PA; $^2$Lupus Center of Excellence, University of Pittsburgh School of Medicine, PA.
4.1 ABSTRACT

Cardiovascular disease (CVD), a multi-factorial trait with complex pathophysiology, is one of the major causes of morbidity and mortality in all westernized populations. Elevated level of CRP has emerged as a sensitive predictor of cardiovascular disease (CVD) in the general population. However, the value of circulating CRP as biomarker for CVD in systemic lupus erythematous (SLE) remains uncertain. Given the positional and functional candidacy of the CRP gene in SLE, we hypothesized that variation in the CRP gene may serve as a potential biomarker to assess CVD risk in patients with SLE. We also hypothesized that association between CRP polymorphism and CVD risk are independent of the effects from circulating CRP levels. Two hundred thirty seven white women who met the ACR criteria for definite or probable SLE were genotyped for five CRP tagSNPs (-861, -390, +90, +838 & +2043). Genotyping was performed using PCR-RFLP, Pyrosequencing or TaqMan assays. B-mode ultrasound was used to measure carotid plaque and carotid intima-media wall thickness (IMT). Association studies were performed using the Fisher’s exact test, multivariable ANOVA and multivariable logistic regression models. Our data showed significant associations between CRP polymorphisms with several important intermediate phenotypes of CVD, including high blood pressure, high BMI, high waist-hip ratio, increased acute phase proteins CRP and fibrinogen, levels of cholesterol, triglycerides, low-density lipoprotein (LDL), high-density lipoprotein (HDL), IMT and stroke risk. Our data revealed that CRP gene modifies the risk of CVD intermediate phenotypes in SLE patients independent of the gene expression levels. Additionally, pleiotropic effects of these CRP polymorphisms reinforced the important role CRP plays in the development of CVD. CVD risk assessment for SLE patients could potentially be improved by inclusion of CRP genetic variation.
Systemic lupus erythematosus (SLE) is a prototypic autoimmune disease characterized by the production of various autoantibodies and involvement of multiple organs. The increased morbidity and mortality in SLE patients can be explained largely from the 5 to 6-fold increased risk of coronary heart disease (CHD). This excess risk is especially pronounced in younger women (age 35-44) where the risk was found to be > 50-fold (1). While traditional cardiovascular disease (CVD) risk factors are found to be more prevalent in patients with SLE, they alone do not fully explain the significantly increased risk for atherosclerosis and CVD in SLE population (2, 3).

Chronic inflammation has been identified as significant part of the underlying mechanism for atherosclerosis and SLE pathogenesis, hence the gene coding for acute phase protein, such as C-reactive protein (CRP), may play an important role in the pathogenesis of CVD in SLE. CRP is a stable and sensitive acute phase protein which has been used as a gold standard of inflammation biomarker. Recent data has demonstrated that CRP not only reflects but also participates in inflammatory cascades (4), and is associated with major CVD risk factors, including inflammation (5), metabolic syndrome (6), and hypertension (HTN) (7) in the general population. These specific characteristics of CRP mark the gene coding for this acute phase protein as an ideal functional candidate gene to investigate as a susceptibility locus for CVD risk.

Genetic studies of CRP have identified several single nucleotide polymorphisms (SNPs) to be associated with the basal CRP levels in both healthy and SLE populations (8-10). However, SNPs that were found to be associated with increased circulating CRP levels have not been found to consistently correlate with increased CVD risk (5, 10, 11). We hypothesize that genetic variation in CRP significantly impact the risk of CVD by promoting the accelerated
development of Framingham CVD risk factors at pre-clinical stage of the CVD. Here the risk factors for disease model are characterized as “intermediate phenotypes”; intermediate in time/mechanism between gene action and the ultimate disease trait. The CVD risk factors we investigate in this study include: inflammatory markers (CRP, fibrinogen, homocysteine, albumin), hemodynamic function (SBP, DBP, hypertension), lipids (total cholesterol, HDL LDL, triglycerides), metabolism (BMI, waist-hip ratio, fasting glucose), and subclinical atherosclerosis (carotid artery intima-media thickness (IMT), carotid plaque).

The objective of this study was to utilize a cohort of SLE women, a population well-known for its inherent increased risk of accelerated atherosclerosis and CVD morbidity and mortality, to examine the role of CRP genetic variation (tagSNPs) in relation to CVD risk factors (intermediate phenotypes). Additionally, we hypothesize that not only are the intermediate phenotypes of CVD in SLE affected by CRP SNPs, but may also be independent from the effect of circulating CRP levels, modified by other epistasis genes, and are likely to share genetic determination (pleiotropy).

4.3 SUBJECTS AND METHODS

4.3.1 Subjects

A total of 237 white female SLE cases currently enrolled in the Pittsburgh Lupus Registry were included in this study. All eligible women who were 18 years of age or older were invited to participate, regardless of their history of cardiovascular event. The 237 cases included in this study had the mean age of 44.26 ± 10.9 (SD), and met the 1982 and 1997 American College of
Rheumatology criteria for definite or probable SLE (12, 13). All subjects have been seen either at the University of Pittsburgh Medical Center or by practicing rheumatologists in the Pittsburgh metropolitan area. This study was approved by the University of Pittsburgh Institutional Review Board, and all subjects provided written informed consent prior to participation. Each participant also provided an authorization for release of medical information so that pertinent hospital and outpatient records could be reviewed to confirm aforementioned events.

4.3.2 Cardiovascular disease intermediate phenotypes measurement

Patients’ clinical information on age, race, smoking habits (having ever smoked), history of stroke were obtained using a standard questionnaire. Body mass index (calculated from height and weight) and waist-to-hip ratio were obtained using a formula of waist measurement divided by the hips measurement of each patient. Current blood pressure status was determined using an average of 2 consecutive sitting blood pressure readings. Levels of total cholesterol, low-density lipoprotein (LDL) cholesterol, high-density lipoprotein (HDL) cholesterol, triglyceride, glucose, and homocysteine levels were measured in fasting blood samples using standardized laboratory tests. Hypertension (HTN) was defined as an average systolic blood pressure = 140 mmHg or an average diastolic blood pressure = 90 mmHg or the use of antihypertensive agents. History of stroke were confirmed by medical records ascertainment.

4.3.3 Inflammatory markers measurement

Serum albumin (dye binding assay), high sensitivity C-reactive protein (ultrasensitive colometric ELISA), and fibrinogen (modified clot-rate assay) were measured. Methods for these
measurements have been previously described (14, 15). Each laboratory test was done in the same lab and same vascular disease measurement session for all the study participants.

4.3.4 Vascular disease measurement

Carotid ultrasound was performed at the University of Pittsburgh Epidemiology Ultrasound Research Laboratory. Briefly, a Toshiba SSA-270A scanner (Tustin, CA) equipped with a 5-MHz linear array imaging probe was used to image the right and left common artery, carotid bulb, and the first 1.5 cm of the internal and external carotid arteries.

Plaque was defined as a distinct focal area protruding into the vessel lumen, with at least 50% greater thickness than that found in surrounding areas. For each segment scanned, the degree of plaque was graded as follows: 0 = no observable plaque; 1 = one small plaque (less than 30% of the vessel diameter); grade 2 = one medium plaque (between 30% and 50% of the vessel diameter) or multiple small plaques; grade 3 = one large plaque (greater than 50% of the vessel diameter) or multiple plaques with at least one medium plaque. The grades were summed across the right and the left carotid arteries to create the combined plaque index (possible range 0-30), the overall measure of the extent of focal plaque.

Intima-media thickness (IMT) was measured across 1-cm segments of both the right and left sides of the near and far walls of the distal common carotid artery and the far wall of the carotid bulb and the internal carotid artery. Values from each location were averaged to produce an overall measure of IMT.
4.3.5 Statistical analysis

Allele and genotype frequencies were calculated by the allele counting method. Goodness of fit to Hardy–Weinberg expected proportions was examined by $x^2$ test. Pair-wise linkage disequilibrium (LD) between markers was estimated using the D’ method (16). The descriptive analyses were summarized as means and standard deviation for continuous variables and as percentages for categorical variables. Correlations between serum CRP and various quantitative intermediate phenotypes were measured using pearson correlation tests. Statistical genotype-phenotype association tests included multivariate one way analysis of variance (ANOVA) and multivariate general linear models for continuous variables, and Fisher’s exact test and multivariate logistic regression for categorical data. Effects of age, BMI, and smoking were controlled in ANOVA and multivariate regression models. CRP level was then included as an additional covariate to assess the genotype-phenotype association independent of the CRP levels. Quantitative risk factors were log-transformed to normal distribution. Carotid plaque was categorized as plaque positive (degree of plaque of equal to or greater than 1) and plaque negative (degree of plaque of zero). R.2.0.1 was used to perform all statistical procedures. All p-value reported in this study is the nominal p-value.

4.3.6 Genotyping

CRP tagSNPs were selected as previously described (17). Reference numbers from NCBI single nucleotide polymorphism are provided for each of the five SNPs examined: -861 is rs3093059, -390 is rs3091244, +90 is rs1417938, +838 is rs1800947 and +2043 is rs1205. Genotyping for +838 and +2043 SNPs was obtained using polymerase chain reaction-restriction fragment
length polymorphism (PCR-RFLP). SNPs -861 was genotyped using pyrosequencing assays. SNPs -390 and +90 were genotyped using TaqMan assays.

4.4 RESULTS

4.4.1 Clinical characteristics of SLE women

A total of 273 women from our original cardiovascular disease study (18) were genotyped and phenotyped on intermediate traits of CVD, carotid artery intima-media thickness (IMT), carotid plaque, and history of ischemic stroke. All women are self-reported to be Caucasians in ethnicity. 106 (44.72%) of the 273 women have had a history of smoking and 95 (40%) are post-menopausal. Their mean age at the time of the phenotype characterization was 44.26 ± 10.9 (mean ± SD) years. Their mean duration of SLE was 10.12 ± 7.09 years with disease activity score (SLAM) of 6.71 ± 3.51 and disease damage score (SLICC) of 1.38 ± 1.74. Of the 273 women, 77 (32.5%) had evidence of carotid plaque, 47 (19.83%) showed renal involvement, and 77 (32.5%) had diagnosis of HTN (Table 4.1).

4.4.2 Allele and genotype frequencies of CRP SNPs and corresponding CRP levels

All five tagSNPs examined showed no significant deviations from Hardy-Weinberg equilibrium. Genotype and allele frequencies of each SNP and their associated log-transformed serum CRP levels (mean ± SD) are presented in Table 4.2. Having adjusted for effects of age, BMI, and smoking, minor alleles of two SNPs revealed significant associations with increased CRP (+90, p
Homozygotes of the minor allele (T) at +90 had the highest logCRP level (1.544 ± 1.048) compared to homozygotes of the common allele (A) (0.639 ± 0.976) and heterozygotes (0.623 ± 1.054). Mean CRP levels were significantly higher in homozygotes of the T allele at the triallelic promoter SNP-390 (1.305 ± 1.128) and heterozygotes with an A allele (CA) (1.356 ± 0.947) when compared to homozygotes of the common allele (C) (0.519 ± 0.947). Promoter SNP -861 was marginally associated with CRP level after adjusting for effects of age, BMI, and smoking ($P = 0.159$)(data not shown). However BMI was found to be significantly associated with this SNP, therefore BMI was treated as a potential effect modifier and removed from the ANOVA test. Without BMI as a covariate in the model, association between -861 and CRP became statistically significant ($p = 0.02$). Subjects having one copy of the minor allele C at –861 were associated with increased CRP levels compared to homozygotes of the common allele T (1.141 ± 1.115 vs. 0.685 ± 1.072) (Table 4.2).

### 4.4.3 Correlation between circulating CRP levels and cardiovascular disease intermediate phenotypes.

Correlation coefficient tests were applied to determine correlation between serum CRP levels and quantitative intermediate phenotypes for CVD shown in Table 4.3. CRP was significantly correlated with BMI ($p < 0.0001$), waist-hip ratio ($p < 0.0001$), glucose ($p =0.003$), HDL ($p = 0.04$), triglycerides ($p < 0.001$), SBP ($p = 0.002$), DBP ($p = 0.037$), fibrinogen ($p < 0.0001$), albumin ($p < 0.0001$), and IMT ($p < 0.001$) (Table 4.3).
4.4.4 Pleiotropic effects of CRP with CVD intermediate phenotype.

Association tests for effects of individual CRP polymorphisms (SNPs) on intermediate phenotypes (listed in Table 4.3) were performed first with age, BMI, and smoking as covariates to adjust their potential confounding effects and the results are presented in Table 4.4A. Four SNPs (-861, -390, +90, and +838) exhibited significant associations with CVD intermediate phenotypes, while three showed pleiotropic effects (one allele → multiple traits) on these intermediate phenotypes. SNP -861 was pleiotropically associated with DBP ($p = 0.05$), HTN ($p = 0.01$), BMI ($p = 0.004$), and waist-hip ratio ($p = 0.037$). SNP -390 was associated with circulating CRP levels ($p = 0.015$). SNP +90 was pleiotropically associated with CRP ($p = 0.003$), fibrinogen ($p = 0.027$), total cholesterol ($p = 0.015$), and LDL ($p = 0.047$). SNP +838 was also pleiotropically associated with HDL ($p < 0.001$), glucose ($p = 0.01$), and IMT ($p = 0.026$) (Table 4.4A).

Serum CRP levels were then added as covariates in the model to determine genotype-phenotype associations independent of the influence from CRP expression. The original associations between these SNPs with multiple intermediate traits remained statistically significant. SNP -861 was associated with DBP ($p = 0.038$), HTN ($p = 0.008$), BMI ($p = 0.04$), and waist-hip ratio ($p = 0.013$). SNP +90 was associated with total cholesterol ($p = 0.009$), and LDL ($p = 0.042$). SNP +838 was associated with HDL ($p < 0.001$), glucose ($p = 0.016$), and IMT ($p = 0.028$) (Table 4.4B). Table 4.5 shows genotype-specific mean ± SD of these intermediate phenotypes.

Pleiotropic effects of the three SNPs were observed in Table 4.5. Homozygotes of the major allele (T) at -861, had BMI of 3.282 ± 0.203 (log-transformed mean ± SD), waist-hip ratio of -0.173 ± 0.124, and DBP of 4.355 ± 0.131. Heterozygotes at -861 showed association with
BMI of 3.361 ± 0.257 ($p = 0.04$), waist-hip ratio of -0.227 ± 0.104 ($p = 0.013$), and DBP of 4.307 ± 0.115 ($p = 0.038$). Homozygotes of the major allele (A) at +90 had total cholesterol level of 5.333 ± 0.189 (log-transformed mean ± SD), and LDL of 4.766 ± 0.229; heterozygotes had a total cholesterol level of 5.237 ± 0.205, and LDL of 4.633 ± 0.332, while homozygotes of the minor allele (T) had a total cholesterol level of 5.209 ± 0.194, ($p = 0.009$) and LDL of 4.645 ± 0.355 ($p = 0.042$). Homozygotes of the major allele (G) at +838 were associated with HDL of 4.03 ± 0.268 (log-transformed mean ± SD), glucose of 4.581 ± 0.143, and IMT of -0.367 ± 0.133, while heterozygotes were associated with HDL of 3.842 ± 0.236 ($p < 0.001$), glucose of 4.652 ± 0.246 ($p = 0.016$), and IMT of -0.313± 0.141 ($p = 0.028$) (Table 4.5). Additionally, heterozygotes at SNP -861 were significantly associated with protection against high blood pressure (OR = 0.231 (95% CI: 0.07-0.68); $p = 0.008$) (data not shown).

### 4.4.5 Associations of SNP -861 with BMI and SNP +838 with Ischemic Stroke

Individual ANOVA tests revealed significant associations between minor allele (T) at -861 with both elevated CRP levels and increased BMI ($p = 0.021$ and 0.004, respectively). We therefore performed a multivariate regression model for BMI in 237 SLE patients including age, smoking, CRP levels, and all 5 SNPs as independent variables. Our data showed that CRP levels and SNP -861 remain significantly associated with BMI (Table 4.6).

A logistic regression model that controlled for the effects of age, BMI, and smoking revealed that carriers of the C allele at SNP +838 had > 4 times the risk of past ischemic stroke (OR, 4.28 [95% CI, 1.12-16.36]; $P = 0.033$). Therefore we performed a multivariate regression model to include additional well-known stroke risk factors to determine the significant contributing variables. The specific risk factors included were LDL, triglycerides, HDL, and
glucose. This analysis revealed that SNP+838 remains the most significant independent risk factor for stroke in SLE patients (OR = 5.61 (1.24-25.27); p = 0.015) (Table 4.7).

4.5 DISCUSSION

An increased CRP level serves as a powerful inflammatory marker and is a risk factor in predicting cardiovascular events in patients with coronary heart disease (19). It is also associated with major cardiovascular disease risk factors including inflammation (5), metabolic syndrome (6), and HTN (7). Surprisingly, genetic polymorphisms within the CRP gene associated with increased CRP levels do not always correspond with increased risk of cardiovascular events (8, 11, 20), casting doubts on the causal role CRP plays in the CVD pathogenesis. Using the cohort of 237 women with SLE we find CRP SNPS altering quantitative variation in levels of several important CVD intermediate phenotypes. We also find HTN and ischemic stroke risks to be significantly affected by SNPs in the CRP gene. We observed pleiotropic effects of 3 of the 4 CRP SNPs and more strikingly, the genotype-phenotype associations were shown to be independent from the effects of circulating CRP levels.

In light of the recent contradictory in vitro roles circulating CRP plays in the development of atherosclerosis (21-23), it is important to determine the association between CRP genetic variation and the development of CVD. The multifactorial nature of CVD development implies that genes affect the risk of CVD in the early pre-clinical stage of the disease. The human CRP gene, which maps to 1q21-q23, is not only a functional candidate gene for inflammation but also a positional candidate gene for SLE (24). Several studies have shown that women with SLE have a high incidence of coronary heart disease, where the highest increased
risk was observed in women under the age of 45 (1, 25). We have reported that women with SLE aged 35-44 were over 50 times more likely to have a myocardial infarction (MI) than were women of similar age from a population based sample (rate ratio = 52.43, 95% CI = 21.6 to 98.5) (1). The unique nature of accelerated atherosclerosis and increased CVD risk makes a cohort of SLE women an ideal population to examine the genetic factors contributing to the development of intermediate phenotypes of CVD.

In addition to successfully replicating the associations for increased CRP levels with SNPs -861, -390, and +90 that have been previously reported in healthy general populations (5, 8) (Table 4.2), we also noted that homozygotes for the rare allele (T) at SNP +90 showed significant associations with lower LDL ($p = 0.042$) and total cholesterol levels ($p = 0.009$) compared to homozygotes of the major allele (A). Carriers of the rare allele (C) at the synonymous SNP +838 were significantly associated with increased IMT ($p = 0.028$), increased glucose level ($p = 0.016$), decreased HDL ($p < 0.001$) (Table 4.5), and increased stroke risk in SLE patients (OR = 5.61; $p = 0.015$) (Table 4.7). These significant genotype-phenotype effects were tested both with and without controlling for the effects of circulating CRP levels. Interestingly, the evidence of association remains significant in both types of multivariate regression models. This strongly suggest that variation in the $CRP$ gene affect risk of these intermediate traits not simply by promoting increased circulating CRP levels, but more likely to behave as a pleiotropic gene (one gene affecting multiple traits) and possibly an epistatic gene (gene interacting with other genes to confer risk of trait) (26).

One of the main advantages of studying intermediate phenotypes instead of the disease itself (CVD events) is their significantly higher population prevalence (prevalence of HTN [~22%] > prevalence of stroke [~2.5%]) (27). More importantly, the number of genetic factors
influencing each intermediate phenotype is presumably smaller than the number of factors affecting the ultimate disease trait. Therefore, the proportion of variance in a disease risk factor explained by a given genetic locus will be greater than the proportion of variance explained in the disease trait, which helps in the identification of genetic biomarkers for both intermediate phenotypes and the disease itself. Studies of multiple disease intermediate phenotypes enable the understanding of the genetic contribution to components of CVD development, which may shed light on novel mechanisms of disease pathogenesis.

We observed an interesting pleiotropic effect of the minor allele (C) of promoter SNP -861. Although the carriers of this minor allele were associated with increased BMI and elevated CRP levels, they were also associated with protection against high blood pressure \((p = 0.008)\) and high waist-hip ratio \((p = 0.038)\) (Table 4.4). To confirm the association between SNP -861 with BMI, we performed a multivariate regression model including CRP, age, smoking, and 5 \(CRP\) tagSNPs as covariates. We again found that CRP levels and -861 significantly and independently contribute to increased BMI \((p = 0.001\) and 0.053, respectively) (Table 4.6). High BMI is a strong risk factor for increased CRP and human CRP has been shown to directly inhibited the binding of leptin to its receptor and blocked cellular signaling \textit{in vitro}, which may play a role in obesity (28). It is therefore not surprising that susceptibility genes for obesity and circulating CRP levels should be shared and that SNP that predisposes to susceptibility of increased weight may also contribute to elevated baseline CRP. It is likely that CRP gene has pleiotropic effects which independently impact obesity and CRP expression via mechanisms unrelated to risk of traditional metabolic syndrome.

We have shown here compelling evidence of association between several \(CRP\) SNPs and well-known intermediate phenotypes of CVD using a cohort of SLE women at high risk of
accelerated atherosclerosis. CVD is a polygenic disease having its phenotypes as product of pleiotropic genes and epistatic genes. In the last decade a number of epistatic genes and pleiotropic genes have been defined based on the CVD event risk (29). The contributing pathophysiologic consequences from each susceptibility gene of CVD likely affect the disease risk at pre-clinical stage by promoting the development of CVD risk factors. Identifying susceptibility loci which promote the accelerated development of these intermediate phenotypes (risk factors) may shed light on the CVD pathogenesis and allow detection of high risk individuals for timely therapeutic interventions. Here we have shown pleiotropic effects of CRP loci influencing several intermediate phenotypes of CVD in SLE women. CVD risk assessment for SLE patients and understanding of the disease mechanisms could potentially be improved by inclusion of CRP genetic variants involved in the different pathophysiological steps leading to CVD.
### Table 4.1. Descriptive statistics of White SLE Women (mean ± sd)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Caucasian SLE Women (n=237)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>44.26 ± 10.9</td>
</tr>
<tr>
<td>Renal</td>
<td>47 (19.83%)</td>
</tr>
<tr>
<td>Smoking History</td>
<td>106 (44.72%)</td>
</tr>
<tr>
<td>BMI</td>
<td>27.69 ± 7.28</td>
</tr>
<tr>
<td>Post Menopausal</td>
<td>95 (40%)</td>
</tr>
<tr>
<td>Antinuclear Antibody +</td>
<td>92%</td>
</tr>
<tr>
<td>Total Cholesterol</td>
<td>195.57 ± 42.1</td>
</tr>
<tr>
<td>Total HDL</td>
<td>56.84 ± 16.64</td>
</tr>
<tr>
<td>LDL</td>
<td>110.22 ± 36.59</td>
</tr>
<tr>
<td>Triglyc.</td>
<td>130.41 ± 91.66</td>
</tr>
<tr>
<td>W-H Ratio</td>
<td>0.839 ± 0.123</td>
</tr>
<tr>
<td>IMT</td>
<td>0.703 ± 0.18</td>
</tr>
<tr>
<td>Disease Duration</td>
<td>10.12 ± 7.09</td>
</tr>
<tr>
<td>C3</td>
<td>95.33 ± 24.07</td>
</tr>
<tr>
<td>C4</td>
<td>21.4 ± 7.45</td>
</tr>
<tr>
<td>High BP</td>
<td>77 (32.4%)</td>
</tr>
<tr>
<td>Plaque Positive</td>
<td>77 (32.5%)</td>
</tr>
<tr>
<td>Current Plaquenil Use</td>
<td>108 (45.6%)</td>
</tr>
<tr>
<td># Years on Steroids (0 - 36)</td>
<td>mean = 5.54, median = 3</td>
</tr>
<tr>
<td>Cholesterol Med. Use</td>
<td>7/229' (3.1%)</td>
</tr>
<tr>
<td>Pulse Pressure</td>
<td>42.07 ± 13.22</td>
</tr>
<tr>
<td>SLICC</td>
<td>1.3787 ± 1.7484</td>
</tr>
<tr>
<td>SLAM</td>
<td>6.7106 ± 3.5145</td>
</tr>
</tbody>
</table>
Table 4.2. Genotype and allele frequencies of CRP SNPs and associated mean logCRP levels (± SD) in 237 SLE women

<table>
<thead>
<tr>
<th>SNP</th>
<th>Genotype</th>
<th>n (%)</th>
<th>Allele</th>
<th>n (%)</th>
<th>LogCRP (Mean ± p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-861 (rs3093059)</td>
<td>TT</td>
<td>192 (84)</td>
<td>T</td>
<td>420 (92)</td>
<td>0.685 ± 1.072 0.021**</td>
</tr>
<tr>
<td></td>
<td>TC</td>
<td>36 (16)</td>
<td>C</td>
<td>36 (8)</td>
<td>1.141 ± 1.115</td>
</tr>
<tr>
<td></td>
<td>CC</td>
<td>0 (0)</td>
<td></td>
<td></td>
<td>...</td>
</tr>
<tr>
<td>-390 (rs3091244)</td>
<td>CC</td>
<td>50 (34)</td>
<td>C</td>
<td>166 (57)</td>
<td>0.519 ± 0.947 0.012*</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>55 (38)</td>
<td>T</td>
<td>101 (35)</td>
<td>0.589 ± 0.899</td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>19 (13)</td>
<td>A</td>
<td>23 (8)</td>
<td>1.305 ± 1.128</td>
</tr>
<tr>
<td></td>
<td>CA</td>
<td>11 (8)</td>
<td></td>
<td></td>
<td>1.356 ± 0.947</td>
</tr>
<tr>
<td></td>
<td>TA</td>
<td>8 (6)</td>
<td></td>
<td></td>
<td>0.696 ± 0.992</td>
</tr>
<tr>
<td></td>
<td>AA</td>
<td>2 (1)</td>
<td></td>
<td></td>
<td>0.515 ± 0.843</td>
</tr>
<tr>
<td>+90 (rs1417938)</td>
<td>AA</td>
<td>66 (46)</td>
<td>A</td>
<td>195 (68)</td>
<td>0.639 ± 0.976 0.0032*</td>
</tr>
<tr>
<td></td>
<td>AT</td>
<td>63 (44)</td>
<td>T</td>
<td>93 (32)</td>
<td>0.623 ± 0.932</td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>15 (10)</td>
<td></td>
<td></td>
<td>1.544 ± 1.048</td>
</tr>
<tr>
<td>+838 (rs1800947)</td>
<td>GG</td>
<td>194 (84)</td>
<td>G</td>
<td>425 (92)</td>
<td>0.703 ± 1.043 0.3728*</td>
</tr>
<tr>
<td></td>
<td>GC</td>
<td>37 (16)</td>
<td>C</td>
<td>37 (8)</td>
<td>0.869 ± 0.979</td>
</tr>
<tr>
<td></td>
<td>CC</td>
<td>0 (0)</td>
<td></td>
<td></td>
<td>...</td>
</tr>
<tr>
<td>+2043 (rs1205)</td>
<td>GG</td>
<td>101 (44)</td>
<td>G</td>
<td>309 (67)</td>
<td>0.845 ± 1.117 0.2068*</td>
</tr>
<tr>
<td></td>
<td>GA</td>
<td>107 (47)</td>
<td>A</td>
<td>151 (33)</td>
<td>0.605 ± 0.949</td>
</tr>
<tr>
<td></td>
<td>AA</td>
<td>22 (10)</td>
<td></td>
<td></td>
<td>0.857 ± 0.999</td>
</tr>
</tbody>
</table>

* Inferential statistic adjusted for Age, BMI, and Smoking
** Inferential statistics adjusted for Age and Smoking
Table 4.3. Pearson's product-moment correlation between CRP and various intermediate phenotypes of CVD

<table>
<thead>
<tr>
<th>Intermediate Traits</th>
<th>Phenotype</th>
<th>Correlation</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Inflammatory markers</strong></td>
<td>C-reactive protein, mg/dL</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>Fibrinogen, mg/dL</td>
<td>0.50147</td>
<td>&lt;2.2e-16</td>
</tr>
<tr>
<td></td>
<td>Homocysteine, mol/L</td>
<td>0.298</td>
<td>0.0774</td>
</tr>
<tr>
<td></td>
<td>Albumin, mg/dL</td>
<td>-0.2818</td>
<td>3.14E-06</td>
</tr>
<tr>
<td><strong>Hemodynamics</strong></td>
<td>Systolic blood pressure, mmHg</td>
<td>0.1885911</td>
<td>0.002048</td>
</tr>
<tr>
<td></td>
<td>Diastolic blood pressure, mmHg</td>
<td>0.1279071</td>
<td>0.03744</td>
</tr>
<tr>
<td></td>
<td>Hypertensive*</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td><strong>Lipids</strong></td>
<td>Total cholesterol, mg/dl</td>
<td>-0.0292864</td>
<td>0.6351</td>
</tr>
<tr>
<td></td>
<td>HDL cholesterol, mg/dl</td>
<td>-0.126</td>
<td>0.04039</td>
</tr>
<tr>
<td></td>
<td>LDL cholesterol, mg/dl</td>
<td>-0.067</td>
<td>0.2808</td>
</tr>
<tr>
<td></td>
<td>Triglycerides, mg/dl</td>
<td>0.21241</td>
<td>0.00049</td>
</tr>
<tr>
<td><strong>Metabolic</strong></td>
<td>Body mass index, kg/m²</td>
<td>0.3600031</td>
<td>1.82E-09</td>
</tr>
<tr>
<td></td>
<td>Waist-hip ratio</td>
<td>0.3197</td>
<td>1.23E-07</td>
</tr>
<tr>
<td></td>
<td>Glucose, mg/dL</td>
<td>0.18303</td>
<td>0.00278</td>
</tr>
<tr>
<td><strong>Subclinical atherosclerosis</strong></td>
<td>Carotid artery intima-media thickness</td>
<td>0.2331</td>
<td>0.000132</td>
</tr>
<tr>
<td></td>
<td>Carotid plaque*</td>
<td>---</td>
<td>---</td>
</tr>
</tbody>
</table>
Table 4.4. Inferential statistics for association between CRP polymorphisms and various intermediate phenotypes

<table>
<thead>
<tr>
<th>Intermediate Traits</th>
<th>PHenotype</th>
<th>A. Without CRP level as covariate (p-value)</th>
<th>B. With CRP level as covariate (p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>CRP polymorphisms</td>
<td>CRP polymorphisms</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-861</td>
<td>-390</td>
</tr>
<tr>
<td>Inflammatory markers</td>
<td>C-reactive protein, mg/dL</td>
<td>0.150</td>
<td><strong>0.015</strong></td>
</tr>
<tr>
<td></td>
<td>Fibrinogen, mg/dL</td>
<td>0.903</td>
<td>0.393</td>
</tr>
<tr>
<td></td>
<td>Homocysteine, mmol/L</td>
<td>0.275</td>
<td>0.723</td>
</tr>
<tr>
<td></td>
<td>Albumin, mg/dL</td>
<td>0.974</td>
<td>0.197</td>
</tr>
<tr>
<td>Hemodynamics</td>
<td>Systolic blood pressure, mmHg</td>
<td>0.286</td>
<td>0.980</td>
</tr>
<tr>
<td></td>
<td>Diastolic blood pressure, mmHg</td>
<td><strong>0.050</strong></td>
<td>0.545</td>
</tr>
<tr>
<td></td>
<td>HTN</td>
<td><strong>0.010</strong></td>
<td>0.910</td>
</tr>
<tr>
<td>Lipids</td>
<td>Total cholesterol, mg/dl</td>
<td>0.705</td>
<td>0.090</td>
</tr>
<tr>
<td></td>
<td>HDL cholesterol, mg/dl</td>
<td>0.838</td>
<td>0.248</td>
</tr>
<tr>
<td></td>
<td>LDL cholesterol, mg/dl</td>
<td>0.898</td>
<td>0.166</td>
</tr>
<tr>
<td></td>
<td>Triglycerides, mg/dl</td>
<td>0.501</td>
<td>0.481</td>
</tr>
<tr>
<td>Metabolic</td>
<td>Body mass index, kg/m2</td>
<td><strong>0.004</strong></td>
<td>0.359</td>
</tr>
<tr>
<td></td>
<td>Waist-hip ratio</td>
<td><strong>0.037</strong></td>
<td>0.866</td>
</tr>
<tr>
<td></td>
<td>Glucose, mg/dL</td>
<td>0.404</td>
<td>0.686</td>
</tr>
<tr>
<td></td>
<td>Carotid artery intima-media thickness, mm</td>
<td>0.416</td>
<td>0.132</td>
</tr>
<tr>
<td></td>
<td>Carotid plaque</td>
<td>0.846</td>
<td>0.694</td>
</tr>
</tbody>
</table>

Inferential statistics for 273 SLE patients adjusting for age, BMI, smoking as covariates. HDL, high-density lipoprotein; LDL, low-density lipoprotein. All phenotypes are quantitative variables except for Hypertensive and Carotid plaque. Hypertensive is defined by having SBP ≥ 140 or DBP ≥ 90 or BP medication use. Carotid plaque positive is defined by having the degree of plaque score equal or greater than 1. All quantitative phenotypes have been log-transformed to achieve normal distribution. Significant differences (p<0.05) are indicated by bold type.
Table 4.5. Pleiotropic effects of CRP SNPs with cluster of CVD risk factors

<table>
<thead>
<tr>
<th></th>
<th>-861</th>
<th></th>
<th></th>
<th></th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>TT</td>
<td>TC</td>
<td>CC</td>
<td></td>
</tr>
<tr>
<td>BMI</td>
<td>191</td>
<td>36</td>
<td>0</td>
<td></td>
<td>0.040</td>
</tr>
<tr>
<td>mean ± sd</td>
<td>3.282 ± 0.203</td>
<td>3.361 ± 0.257</td>
<td>---</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WHR</td>
<td>36</td>
<td>0</td>
<td></td>
<td></td>
<td>0.013</td>
</tr>
<tr>
<td>mean ± sd</td>
<td>-0.173 ± 0.124</td>
<td>-0.227 ± 0.104</td>
<td>---</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DBP</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td>0.038</td>
</tr>
<tr>
<td>mean ± sd</td>
<td>4.355 ± 0.131</td>
<td>4.307 ± 0.115</td>
<td>---</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>+90</th>
<th>AA</th>
<th>AT</th>
<th>TT</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>66</td>
<td>63</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Cholesterol</td>
<td>mean ± sd</td>
<td>5.333 ± 0.189</td>
<td>5.237 ± 0.205</td>
<td>5.209 ± 0.194</td>
<td>0.009</td>
</tr>
<tr>
<td>LDL</td>
<td>mean ± sd</td>
<td>4.766 ± 0.249</td>
<td>4.633 ± 0.332</td>
<td>4.645 ± 0.355</td>
<td>0.042</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>+838</th>
<th>GG</th>
<th>GC</th>
<th>CC</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>193</td>
<td>37</td>
<td>0</td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>HDL</td>
<td>mean ± sd</td>
<td>4.03 ± 0.268</td>
<td>3.842 ± 0.236</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>Glucose</td>
<td>mean ± sd</td>
<td>4.581 ± 0.143</td>
<td>4.652 ± 0.246</td>
<td>---</td>
<td>0.016</td>
</tr>
<tr>
<td>IMT</td>
<td>mean ± sd</td>
<td>-0.367 ± 0.0133</td>
<td>-0.313 ± 0.141</td>
<td>---</td>
<td>0.028</td>
</tr>
</tbody>
</table>

Results are presented in log-transformed values adjusting for age, BMI (except for BMI as dependent variable), smoking and serum CRP levels.
Table 4.6. Multivariate regression model for Body Mass Index (BMI) in 237 SLE patients

<table>
<thead>
<tr>
<th></th>
<th>Df</th>
<th>Sum of Sq</th>
<th>RSS</th>
<th>AIC</th>
<th>F value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td></td>
<td>5794.7</td>
<td>545</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRP</td>
<td>1</td>
<td>552.1</td>
<td>6346.8</td>
<td>555.5</td>
<td>11.528</td>
<td>0.001</td>
</tr>
<tr>
<td>Age</td>
<td>1</td>
<td>78.9</td>
<td>5873.6</td>
<td>544.9</td>
<td>1.647</td>
<td>0.202</td>
</tr>
<tr>
<td>Smoking</td>
<td>1</td>
<td>69.6</td>
<td>5864.3</td>
<td>544.7</td>
<td>1.453</td>
<td>0.230</td>
</tr>
<tr>
<td>-861</td>
<td>1</td>
<td>183.1</td>
<td>5977.8</td>
<td>547.3</td>
<td>3.823</td>
<td>0.053</td>
</tr>
<tr>
<td>-390</td>
<td>5</td>
<td>107.6</td>
<td>5902.3</td>
<td>537.5</td>
<td>0.449</td>
<td>0.813</td>
</tr>
<tr>
<td>+90</td>
<td>2</td>
<td>11.6</td>
<td>5806.3</td>
<td>541.3</td>
<td>0.121</td>
<td>0.886</td>
</tr>
<tr>
<td>+838</td>
<td>1</td>
<td>22.9</td>
<td>5817.6</td>
<td>543.6</td>
<td>0.479</td>
<td>0.490</td>
</tr>
<tr>
<td>+2043</td>
<td>3</td>
<td>80.3</td>
<td>5875</td>
<td>540.9</td>
<td>0.559</td>
<td>0.643</td>
</tr>
</tbody>
</table>
Table 4.7. Multivariate regression model for Stroke in 11 SLE patients

<table>
<thead>
<tr>
<th></th>
<th>Df</th>
<th>Sum of Sq</th>
<th>RSS</th>
<th>AIC</th>
<th>F value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1</td>
<td>9.21</td>
<td>-703.31</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>1</td>
<td>0.0003951</td>
<td>9.21</td>
<td>-705.3</td>
<td>0.009</td>
<td>0.923</td>
</tr>
<tr>
<td>BMI</td>
<td>1</td>
<td>0.01</td>
<td>9.22</td>
<td>-705.03</td>
<td>0.263</td>
<td>0.608</td>
</tr>
<tr>
<td>Smoking</td>
<td>1</td>
<td>0.01</td>
<td>9.22</td>
<td>-705.12</td>
<td>0.185</td>
<td>0.667</td>
</tr>
<tr>
<td>CRP</td>
<td>1</td>
<td>0.01</td>
<td>9.22</td>
<td>-705.11</td>
<td>0.188</td>
<td>0.665</td>
</tr>
<tr>
<td>LDL</td>
<td>1</td>
<td>0.01</td>
<td>9.22</td>
<td>-705.1</td>
<td>0.197</td>
<td>0.658</td>
</tr>
<tr>
<td>Trig</td>
<td>1</td>
<td>0.01</td>
<td>9.22</td>
<td>-705.09</td>
<td>0.210</td>
<td>0.647</td>
</tr>
<tr>
<td>HDL</td>
<td>1</td>
<td>0.0047983</td>
<td>9.21</td>
<td>-705.19</td>
<td>0.113</td>
<td>0.738</td>
</tr>
<tr>
<td>Glucose</td>
<td>1</td>
<td>0.07</td>
<td>9.28</td>
<td>-703.64</td>
<td>1.606</td>
<td>0.206</td>
</tr>
<tr>
<td>+838</td>
<td>1</td>
<td>0.25</td>
<td>9.46</td>
<td>-699.14</td>
<td>5.981</td>
<td><strong>0.015</strong></td>
</tr>
</tbody>
</table>
4.7 MANUSCRIPT 3 REFERENCES


22. Hirschfield GM, Gallimore JR, Kahan MC, Hutchinson WL, Sabin CA, Benson GM, et al. Transgenic human C-reactive protein is not proatherogenic in apolipoprotein E-


5.0 CONCLUSION CHAPTER

The objective of this dissertation project was to investigate the role of CRP as a potential disease-susceptibility locus or biomarker for SLE. Two independent SLE cohorts of North American Caucasians were applied. The two cohorts consisted of differentially ascertained case-control sample from Pittsburgh and a family-based sample from Los Angeles. In addition, the role of CRP polymorphisms in circulating CRP levels and additional well-established intermediate phenotypes for cardiovascular disease (CVD) in SLE patients were examined. We hypothesize that CRP polymorphisms affect the risk of SLE, and also contribute to accelerated atherosclerosis/cardiovascular disease in SLE women, independent of the effects from circulating CRP.

5.1 SUPPLEMENTARY DATA OF THIS DISSERTATION

5.1.1 Genetic and environmental contributions of serum CRP

To understand the factors contributing to the variation in circulating CRP, we performed multiple linear regression analyses to determine the proportion of variance in CRP levels that are attributed to the effects from: (1) age, BMI and smoking history, (2) five CRP tagSNPs selected in this project, and (3) combined age, BMI, smoking and 5 tagSNPs. Our data showed that
SNP+90 remained associated with CRP levels ($p$-value = 0.08) in the presence of all other variables, while environmental factors contributed most significantly to CRP levels ($P = 0.002$ to 0.003) (Table 5.1).

The proportion of variance (adjusted $R^2$ value) in log-transformed CRP levels explained by the five $CRP$ SNPs together was relatively small (0.062), while the proportion of variance due to age, BMI, and smoking was significantly higher (0.155). Together, age, BMI, smoking and five SNPs explained approximately 0.267 of the total variation in serum CRP levels in our SLE cohort (Table 5.1). While the combined effects from these variables may seem small, they must be considered relative to the many factors that likely influence CRP levels in an inflammatory-based autoimmune disease, and are relatively large for a genetic effect in a complex disease. Moreover, the proportion of variance due to genetic effect ($R^2 = 0.062$) and covariates in the model (age, BMI, smoking, $R^2 = 0.155$) are comparable to those in a recent report by Miller et al. (0.02 and 0.2, respectively) (1). Previously reported heritability of circulating CRP (up to 0.56) (2) supports the presence of additional functional genetic variations with substantial effects that influence CRP levels both in SLE and healthy populations.
Table 5.1. Multivariate regression model for Body Mass Index (BMI) in 237 SLE patients

<table>
<thead>
<tr>
<th></th>
<th>Df</th>
<th>Sum of Sq</th>
<th>RSS</th>
<th>AIC</th>
<th>F value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td></td>
<td>111.377</td>
<td>2.836</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>1</td>
<td>8.35</td>
<td>119.727</td>
<td>10.668</td>
<td>9.072</td>
<td>0.00316</td>
</tr>
<tr>
<td>BMI</td>
<td>1</td>
<td>9.734</td>
<td>121.11</td>
<td>12.23</td>
<td>10.5749</td>
<td>0.00149</td>
</tr>
<tr>
<td>Smoking</td>
<td>1</td>
<td>9.383</td>
<td>120.759</td>
<td>11.835</td>
<td>10.1933</td>
<td>0.0018</td>
</tr>
<tr>
<td>-861</td>
<td>1</td>
<td>1.225</td>
<td>112.601</td>
<td>2.323</td>
<td>1.3305</td>
<td>0.25099</td>
</tr>
<tr>
<td>-390</td>
<td>5</td>
<td>4.709</td>
<td>116.085</td>
<td>-1.533</td>
<td>1.0231</td>
<td>0.40716</td>
</tr>
<tr>
<td>+90</td>
<td>2</td>
<td>4.681</td>
<td>116.058</td>
<td>4.435</td>
<td>2.5428</td>
<td>0.08284</td>
</tr>
<tr>
<td>+838</td>
<td>1</td>
<td>0.841</td>
<td>112.217</td>
<td>1.858</td>
<td>0.9131</td>
<td>0.34119</td>
</tr>
<tr>
<td>+2043</td>
<td>2</td>
<td>2.383</td>
<td>113.76</td>
<td>1.715</td>
<td>1.2946</td>
<td>0.27778</td>
</tr>
</tbody>
</table>

Adjusted R-Square: 0.155 (age, BMI, smoking)
Adjusted R-Square: 0.062 (5 SNPs)
Adjusted R-squared: 0.267 (age, BMI, smoking +5 SNPs)
5.1.2 Correlation between SLE disease activity and CRP levels

Increased SLE disease activity is almost always accompanied with increased systemic inflammation; therefore it was of a particular interest to investigate whether or not there is significant correlation between SLE disease activity (measured by SLAM) and log-transformed CRP levels in the Pittsburgh cohort. As shown in Figure 5.1, there is no distinguishable pattern of correlation between the two quantitative variables. This observation supports the current literature cited in manuscripts 1 and 2 that circulating CRP levels do not accurately predict SLE disease activity.

Figure 5.1. Scatter plot of CRP levels versus SLE disease activity
We furthermore divided quantitative variables of CRP and SLAM scores into “extreme phenotype” categories of three. Based on the “high, middle, and low” category assigned to each patient according to their data, we explored whether genotype distribution of each of the CRP SNPs examined differed significantly between patients who have “high disease activity (SLAM category 3) with high marker of inflammation (CRP category 3)” versus patients who have “high disease activity (SLAM category 3) but with low marker of low inflammation (CRP category 1)”. The result showed no significant differences in genotype distribution between those patients who have the high marker of inflammatory biomarker with high disease activity (as expected) versus those patients who have unexpectedly low marker of inflammatory biomarker with high disease activity (Table 5.2 and 5.3).

Table 5.2. Distribution of SLE patients by disease activity score (SLAM) and CRP levels

<table>
<thead>
<tr>
<th>LogCRP level</th>
<th>SLAM Score</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 0.2624</td>
<td></td>
<td>13</td>
<td>13</td>
<td>37</td>
</tr>
<tr>
<td>0.2624 to 1.1314</td>
<td></td>
<td>14</td>
<td>28</td>
<td>28</td>
</tr>
<tr>
<td>&gt; 1.1314</td>
<td></td>
<td>11</td>
<td>26</td>
<td>45</td>
</tr>
</tbody>
</table>
Table 5.3. No significant association between CRP SNPs and SLAM–CRP correlation

<table>
<thead>
<tr>
<th>SNP</th>
<th>Genotype</th>
<th>lowCRP with highSLAM ( n = 37 )</th>
<th>highCRP with highSLAM ( n = 45 )</th>
<th>( p )-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>-861 (rs3093059)</td>
<td>TT</td>
<td>27</td>
<td>34</td>
<td>0.59</td>
</tr>
<tr>
<td></td>
<td>TC</td>
<td>6</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CC</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>-390 (rs3091244)</td>
<td>CC</td>
<td>10</td>
<td>8</td>
<td>0.31</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>8</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>3</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CA</td>
<td>0</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TA</td>
<td>1</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AA</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>+90 (rs1417938)</td>
<td>AA</td>
<td>11</td>
<td>11</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>AT</td>
<td>10</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>0</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>+838 (rs1800947)</td>
<td>GG</td>
<td>32</td>
<td>35</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td>GC</td>
<td>4</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CC</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>+2043 (rs1205)</td>
<td>GG</td>
<td>19</td>
<td>23</td>
<td>0.52</td>
</tr>
<tr>
<td></td>
<td>GA</td>
<td>15</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AA</td>
<td>2</td>
<td>6</td>
<td></td>
</tr>
</tbody>
</table>
5.2 DISsertation Summary

By studying human subjects with SLE and CVD, we hope to acquire further understanding of mechanisms/pathways contributing to this unique population at marked increased risk of CVD. Our results additionally contribute to a better insight of the inflammatory processes in the context of CVD in the general population. Results addressing specific scientific questions in this study are summarized below:

5.2.1 Does variation in CRP gene confer susceptibility to SLE?

In chapters 2 and 3 of this study, we report that no individual CRP polymorphism is individually associated with the risk of SLE. When we examined the combined effect of polymorphism as haplotypes, we found evidence of associations in both the case-control and family-based study designs involving the -390 and +90 SNPs. We found no association between SNP -2043 and SLE risk as reported recently in a British SLE family-based study (3). The significant haplotype results in this study suggest that the variation in the CRP gene may modify SLE risk. Determination of how CRP variation influences SLE risk is expected to further our understanding of SLE etiology and may have direct clinical relevance.
5.2.2 Does CRP variation affect intermediate phenotypes of CVD in SLE?

In chapter 4 we have presented evidence that several CRP polymorphisms significantly contribute to important intermediate phenotypes of CVD. The CVD risk factors in the general populations affected by CRP SNPs are metabolic syndrome associated traits, BMI, waist-hip ratio, glucose levels, hypertension, and lipids profile. In addition, acute phase proteins CRP and fibrinogen were also found to be effected significantly by several CRP promoter polymorphisms. Together, the degree of association seen here may explain the increased metabolic syndrome observed in SLE patients compared to the general population (4) and chronic systemic inflammation which contribute to accelerated atherosclerosis in SLE patients. The most critical finding are that these significant genotype-phenotype associations remain significant after controlling for the effects of circulating serum CRP, and exhibited pleiotropic effects. This strongly suggests the important role the CRP gene plays in CVD intermediate phenotype development; thereby increasing the risk of CVD events in persons with risk genetic variants.

5.2.3 Does CRP variation affect CVD risk in SLE?

In chapter 4 we have shown for the first time that CRP exon 2 polymorphism +838 significantly alters ischemic stroke risk and increased carotid artery intima-media thickness in SLE patients. Strikingly, effect of +838 on stroke is independent of the well known CVD risk factors: age, BMI, smoking history, serum CRP, low density lipoprotein, high density lipoprotein, triglyceride, and glucose level. Although +838 resides in exon 2, it does not affect the protein structure of CRP. It is likely that SNP is in significant linkage disequilibrium with functional allele in either CRP or nearby loci.
5.2.4 How does environmental factors and CRP variation contribute to circulating CRP in SLE, and is CRP level associated with disease activity?

In chapter 5, we present evidence that age, BMI, and smoking history significantly contribute to variation in circulating serum CRP levels. Together they explained 15.5% of the CRP variation. On the other hand, the combined effect of 5 CRP tagSNP explains only 6.2% of the CRP variation. Together, this model explains 26.7% of the variation in circulating CRP. This finding supports the multi-factorial and polygenic characteristic of circulating CRP. Additionally, the correlation test between CRP and SLE disease activity presented in Figure 5.1 confirmed the previous observation that CRP, being a stable and sensitive inflammation marker, does not correlate with SLE disease activity as one would expect. While the precise mechanism underlying this intriguing phenomenon is still unknown, it appears that the five CRP tagSNPs we examined do not contribute to the discordant correlation pattern.

5.3 FUTURE WORK

Human lupus is a complex genetic trait involving multiple genes and their interactions, leading to heterogeneous phenotype presentations. While many independent studies have identified many susceptibility genes thought to contribute to the development of SLE, success in determining genotypic mechanisms leading to SLE risk and phenotype variation has been rather limited (5-7). It is therefore important to carry out additional genetic studies that may lead to the identification of new biological pathways.
In this study, SNPs -390 and +90, define the CRP haplotypes that are associated with SLE risk. For the promoter triallelic SNP -390, previous work has already shown the T allele correlated with increased binding of transcription factors and thereby increases the promoter activity. It would be of great interest to compare promoter activity between SLE patients and healthy controls given the same genotype. If the promoter activity express differently between SLE and controls given the same genotype, then the evidence of -390 as risk SNP increases. It will suggest that SLE risk associated with this promoter is not solely a result of circulating CRP level (which is shown to increase given -390T). Instead this promoter SNP may have interactive effect with other unidentified genes or proteins in conferring SLE risk.

SNP +90 is an intron/exon SNP that has potential to induce alternative splicing in a gene. Evidence of potential alternative transcript isoform(s) induced by this SNP may be explored by researching the databases of expressed sequence tags (ESTs) (8). If the alternative splicing is seen, effects of this alternative splicing may be further confirmed using gene expression analysis. Differential gene expressions resulting from this alternative splicing may explain in part the mechanism in which CRP participates in the various biological functions in the host (discussed in Chapter 1).

Additional works of interest include determining the associations between CRP polymorphisms and anti-CRP antibody (9) to assess if the abnormally higher prevalence of anti-CRP antibody in SLE may be a result of genetic defect, thereby effect disease severity. Exploration of differential effects native CRP (nCRP) and modified CRP (mCRP) have on different phenotype presentation in SLE may further our understanding of the causes and process of this protein conformation change. Of course, once we understand how the conformational
change in CRP occurs, we should conduct genetic association studies to identify susceptibility loci for this bio-cellular event.

Finally, applying mouse model, including knockout models, genetic hypomorphs and hypermorphs, will be useful in modeling and characterizing the human disease for better understanding of the disease etiology. For example, a non-SLE susceptibility CRP intron repeat polymorphisms has been shown to correlate with baseline CRP levels (10), and may be designed as a transgene to be inserted in lupus-prone mice NZB/NZW model (11). Two CRP haplotypes may be created using the promoter variant GT\textsuperscript{1818} and GT\textsuperscript{2020} which correspond to increased baseline CRP, and GT\textsuperscript{1616} and GT\textsuperscript{2121} which correspond to decreased CRP. Additionally, we will use a non-treated NZB/NZW lacking human CRP expression (CRP in mice does not rise significantly as human CRP does in the presence of infection) as a control model. By comparing the time of onset of SLE-like phenotype and disease severity between these models, insights may be found on whether elevated CRP or decreased CRP directly contribute to the pathogenesis of SLE. If the disease progression among these three differentially treated animal models does not differ significantly, we must re-evaluate the hypothesis that circulating CRP is causally involved in the SLE risk, and turn our attention to the CRP polymorphisms and their interactions with additional susceptibility loci or environmental risk factors. Successful animal models can provide fundamental insights on SLE etiology and help identify more sensitive markers of disease activity and damage. Better understanding of SLE pathophysiology is the fundamental key to improve treatment, and ultimately prevent SLE.
Lupus affects an estimated 239,000 (conservative estimate) people in the United States (US) (12). Due to the lack of highly sensitive diagnostic markers to define and capture cases, the definitive epidemiological information on lupus is also lacking. Therefore the exact number of people affected by lupus is unknown and likely to be significantly higher than the estimate. SLE is a serious rheumatic autoimmune disease with unknown etiology and relatively unpredictable disease manifestations among different patients as well as within a particular patient given her/his lifetime. Even though the advanced medical and scientific research in the past two decades has significantly improved the quality of life for SLE patients, SLE still remains one of the most fatal forms of rheumatic diseases.

Center for Disease Control (CDC) recently reported that the annual number of SLE deaths in the US during year 1979 to 1988 has increased from 879 to 1,406, and the crude death rate has increased from 39 to 52 per 10 million population (13). Of all SLE deaths, 36.4% occurred among persons aged 15 to 44 years. Crude death rates were more than 5 times higher among women than men, and were also more than 3 times higher among blacks than whites (13). The striking disparities and the rising death rate in SLE call for better preventative strategies for this disease on the public health level. Prevention of deaths and disabilities related to SLE requires early recognition and diagnosis in order for physicians to administer timely therapeutic management. Therefore understanding the etiology and pathogenesis of SLE is of a crucial importance.

The results of this project are compelling because they support the hypothesis that variation in the \textit{CRP} gene is involved in the pathogenesis of SLE and accelerated atherosclerosis.
Investigations into the relationship of these genetic variations in a gene coding for such sensitive biomarkers and potential contributor for systemic inflammation, help provide better insights into the biological pathways involved in the SLE pathogenesis. Knowledge of these genetic associations may help elucidate mechanistic pathways amenable to pharmacological intervention. Results of this research further provide the basis for future functional and epidemiological studies to elucidate the pathogenic roles CRP and other inflammatory risk factors play in SLE development.
5.5 REFERENCES FOR CONCLUSION CHAPTER


BIBLIOGRAPHY


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. Clinical Immunology & Immunopathology 1994;70(1):22-7.

Du Clos TW, Zlock LT, Rubin RL. Analysis of the binding of C-reactive protein to histones and chromatin. Journal of Immunology 1988;141(12):4266-70.


Kaplan MH, Volanakis JE. Interaction of C-reactive protein complexes with the complement system. I. Consumption of human complement associated with the reaction of C-reactive protein with pneumococcal C-polysaccharide and with the choline phosphatides, lecithin and sphingomyelin. Journal of Immunology 1974;112(6):2135-47.


Lee PY, Reeves WH. Type I interferon as a target of treatment in SLE. Endocrine, Metabolic & Immune Disorders Drug Targets 2006;6(4):323-30.


Szalai AJ. C-reactive protein (CRP) and autoimmune disease: facts and conjectures. Clinical & Developmental Immunology 2004;11(3-4):221-6.


