Investigation of Candidate Genes and HLA-Related Risk Factors in a Genetic Study of Autoimmune Disease

By

Paola Grasso Bronson

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Epidemiology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Lisa F. Barcellos, Chair
Professor Patricia A. Buffler
Professor Alan E. Hubbard
Professor Montgomery Slatkin

Fall 2010

Investigation of Candidate Genes and HLA-Related Risk Factors in a Genetic Study of Autoimmune Disease

by

Paola Grasso Bronson

ABSTRACT

Investigation of Candidate Genes and HLA-Related Risk Factors in a Genetic Study of Autoimmune Disease

by

Paola Grasso Bronson

Doctor of Philosophy in Epidemiology

University of California, Berkeley

Professor Lisa F. Barcellos, Chair

Collectively autoimmune diseases constitute a major burden to society. Though the etiology of autoimmune diseases remain largely unknown, evidence supports a substantial genetic component. For many autoimmune diseases, twin studies demonstrate a dramatically higher disease concordance rate in monozygotic twins than in dizygotic twins. Genes in the major histocompatibility complex (MHC) region on the short arm of chromosome 6, particularly the human leukocyte antigen (HLA) class II genes, are strongly associated with risk of developing rheumatoid arthritis (RA), systemic lupus erythematosus (SLE), multiple sclerosis (MS) and type 1 diabetes (T1D).

The MHC class II transactivator gene (*CIITA*, also called *MHC2TA*), located on the short arm of chromosome 16, encodes an important transcription factor (CIITA) regulating the genes required for HLA class II MHC-restricted antigen presentation. Thus *CIITA* is a strong biological candidate for studies of autoimmune disease. Directly adjacent to *CIITA* lies the C-type lectin domain family 16, member A gene (*CLEC16A*, previously called *KIAA0350*). *CLEC16A* is a sugar binding receptor containing a putative immunoreceptor and was recently identified as a novel T1D and MS susceptibility locus through genomewide association (GWA) studies.

HLA may also influence susceptibility to autoimmune disease through other inherited and noninherited mechanisms, in addition to genetic transmission of risk alleles. Evidence for increased maternal-offspring HLA compatibility and differences in both maternal vs. paternal transmission rates (parent-of-origin effects) and nontransmission rates (noninherited maternal antigen (NIMA) effects) in autoimmune diseases have been reported.

The investigation described in this dissertation tested hypotheses that (1) the *CIITA* -168A/G promoter polymorphism (rs3087456) influences susceptibility to RA (Chapter 2); (2) common genetic variation in *CIITA* influences susceptibility to RA in a case-control study (Chapter 3); (3) common genetic variation in *CIITA* influences susceptibility to SLE or specific secondary SLE phenotypes (Chapter 4); (4) common genetic variation in *CIITA* influences susceptibility to MS (Chapter 5); (5) common genetic variation in *CLEC16A* influences susceptibility to RA (Chapter 6); (6) the HLA class II *DRB1* locus influences susceptibility to SLE through maternal-offspring

HLA compatibility, parent-of-origin and NIMA effects (Chapter 7); and (7) the HLA classical loci influence susceptibility to T1D through maternal-offspring HLA compatibility, parent-of-origin and NIMA effects (Chapter 8).

This dissertation includes the first study to fully characterize common genetic variation in *CIITA* and *CLEC16A*, including assesment of haplotypes, sex-specific effects, secondary clinical phenotypes and HLA risk alleles. Results do not provide evidence for association between *CIITA* and RA or SLE or for association between *CLEC16A* and RA. Interestingly, this study revealed evidence for an association between the *CIITA* missense mutation rs4774 and increased risk for MS in the presence of the *HLA-DRB1*1501* risk allele. There was no linkage disequilibrium between *CIITA* and *CLEC16A*, and the observed association between *CIITA* and MS in the presence of *HLA-DRB1*1501* was independent of the association between *CLEC16A* and MS.

The first studies to examine maternal-offspring HLA compatibility in T1D and *HLA-DRB1* parent-of-origin and NIMA effects in SLE, and the largest study to examine maternal-offspring HLA compatibility in SLE and HLA parent-of-origin and NIMA effects in T1D were also performed. No evidence that the *HLA-DRB1* locus influences risk for SLE or that the classical HLA loci influence risk for T1D through these novel biological phenomena was revealed.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENTS

I would like to sincerely thank my outstanding advisor, Professor Lisa Barcellos, and members of my dissertation committee, Professors Patricia Buffler, Alan Hubbard and Montgomery Slatkin. They made this work possible. It has also been an immense pleasure to collaborate with Professors Lindsey Criswell and Glenys Thomson. I am grateful to Professors Barcellos, Buffler and Criswell for sponsoring my Predoctoral training grant from the National Institutes of Health (NIH), and to Professors Ira Tager and Art Reingold for their guidance.

Special thanks to my wonderful husband, Nathan Bronson, for his deep understanding and patience. His loving support and encouragement have been very meaningful to me.

I am grateful to my parents (Angelo and Soffy Grasso), brothers (Angelo and Giovanni Grasso), relatives in the extended Grasso, Bronson and Sage families, and friends for their vital support; members of the Barcellos Genetic Epidemiology and Genomics Laboratory for their helpful assistance (especially Patricia Ramsay for graciously sharing her expertise, Gary Artim for running a tight IT ship, Suzanne May for coordinating genotyping, and fellow doctoral students Dr. Farren Briggs, Benjamin Goldstein and Xiaorong Shao for interesting discussions); my classmates for brightening my experience (especially Dr. Melinda Aldrich and Wendy Verret); rock climber extraordinaire Lynn Hill for her encouragement; and Professors Margaret Pericak-Vance, Eden Martin, Yi-Ju Li, William Scott, Marcy Speer and Silke Schmidt for introducing me to the field of genetic epidemiology.

I would like to thank my research collaborators for their invaluable support and for allowing me to use their patient and control data: Dr. Kenneth Beckman, Dr. Timothy Behrens, Stacy Caillier, Dr. Frans Claas, Dr. Alastair Compston, Dr. Criswell, Dr. Philip De Jager, Dr. Patrick Gaffney, Dr. Peter Gregersen, Dr. David Hafler, Dr. Jonathan Haines, Dr. John Harley, Dr. Stephen Hauser, Dr. Adrian Ivinson, Jennifer Kelly, Dr. Leanne Komorowski, Julie Lane, Dr. Jacob McCauley, Dr. Kathy Moser, Joanne Nititham, Dr. Janelle Noble, Dr. Jorge Oksenberg, Dr. Pericak-Vance, Dr. John Rioux, Dr. Stephen Sawcer, Dr. Michael Seldin, Dr. Kim Taylor, Dr. Thomson, Dr. Rebecca Zuvich, the Wellcome Trust Case Control Consortium, the International Multiple Sclerosis Genetics Consortium and the International Type 1 Diabetes Genetics Consortium.

This dissertation work was supported by the NIH National Institute of Allergy and Infectious Diseases (NIAID) (F31AI075609), American College of Rheumatology Research and Education Foundation (REF)/Abbott Graduate Student Achievement Award, UC Berkeley Graduate Division Conference Travel Grant and Russell M. Grossman Medical Research Endowment. The content is solely the responsibility of the author and does not necessarily represent the official views of the NIH or NIAID.

Parts of this dissertation work were funded by the NIH/NIAID (R01 AI065841, R01 AI059829, U19 AI067152, R37 AI24717, P01 AI083194, 5R01 A1063274-06), the NIH National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS) (RO1 AR44422, NO1 AR22263, R01 AR050267, K24 AR02175, R01 AR22804, R01 AR052300, K24 AR02175, P01 AR049084, R01 AR42460, R01 AR043274-14, N01 AR62277), NCRR (P20-RR020143), the National MS Society (RG201-A-1, RG4201, RG2901), the Medical Research Council

(G0700061), the Cambridge National Institute for Health Research (NIHR) Biomedical Research Centre, the NIH National Institute of Neurological Disorders and Stroke (NINDS) (R01 NS049477), the Kirkland Scholar Award from the Mary Kirkland Center for Lupus Research, the Alliance for Lupus Research, the U.S. Department of Veterans Affairs, the Lupus Foundation Minnesota.

This study was performed in part in the General Clinical Research Center, Moffitt Hospital, University of California San Francisco, with funds provided by the National Center for Research Resources, 5-M01-RR-00079, U.S. Public Health Service.

This dissertation work used resources provided by the Type 1 Diabetes Genetics Consortium, a collaborative clinical study sponsored by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), NIAID, National Human Genome Research Institute (NHGRI), National Institute of Child Health and Human Development (NICHD), and Juvenile Diabetes Research Foundation International (JDRF) and supported by U01 DK062418.

This dissertation work makes use of data generated by the WTCCC; a full list of the investigators who contributed to the generation of the data is available from www.wtccc.org.uk, and funding for the project was provided by the Wellcome Trust under award 076113.

I would like to acknowledge the Nature Publishing Group, John Wiley and Sons, Inc., Oxford University Press and BMJ Publishing Group Ltd. for granting me permission to use their published material.

CHAPTER ONE

Introduction

AUTOIMMUNE DISEASES

Public health significance of autoimmune diseases

Autoimmune diseases include more than 80 chronic disorders affecting around 5 to 8 percent of Americans (1). The prevalence has recently been estimated to be 7.6 to 9.4 percent (2). In general, autoimmune diseases tend to afflict females more often than males (3). Females have 2.7 times the risk of developing an autoimmune disease compared to males (4). Indeed, autoimmune diseases are one of the 10 leading causes of death among every age category for females younger than 65 years old in the United States (U.S.) (5). Though many autoimmune disorders are very rare, collectively they affect 14.7 to 23.5 million Americans, and can be severely debilitating (1). Some have medical treatments available but at this time, these conditions cannot be cured.

Neuropsychiatric involvement is common in autoimmune diseases, including multiple sclerosis (MS), systemic lupus erythematosus (SLE), rheumatoid arthritis (RA) and primary Sjögren's syndrome (6-8). For example, almost a third of SLE patients suffer from neuropsychiatric events attributable to SLE, including: seizures (21%), mood disorders such as depression (18%), cognitive dysfunction (9%) and psychosis (5%) (9, 10). In addition, limitations in function resulting from autoimmune disease can lead to severe depression. For example, depression is much more common in individuals suffering from arthritis than those without (attributable risk (AR) = 18.1%, 95% confidence interval (95% CI) = 9.9 to 25.6) (11). Because RA patients suffering from depression have worse health outcomes, investigators are examining predictors of depression to help guide treatment of RA that also targets depressive symptoms (12). The World Health Organization (WHO) estimates that depression is the number one leading cause of burden of disease (measured by disability-adjusted life years (DALYs)) in middle-income (29 million DALYs) and high-income countries (10 million DALYs) (eighth leading cause in low-income countries (26.5 million DALYs), third leading cause worldwide (total 65.5 million DALYs)) (13). Thus, autoimmune diseases represent a major public health burden to society due to decreased quality of life, lost productivity, co-morbid mental illnesses, medical care utilization and direct and indirect economic costs (3).

The common pathological thread underlying autoimmune disease is dysregulation of the immune system, whereby a loss of tolerance to self antigens leads the body to attack its own cells, tissues and organs through inappropriate immune responses. Autoimmune diseases can be classified as organ-specific or systemic. Organ-specific autoimmune diseases lead to localized damage. Examples of organ-specific autoimmune diseases include: autoimmune adrenal insuffiency disease, characterized by the presence of autoantibodies to cells of the adrenal gland, Graves' disease, characterized by the presence of autoantibodies to thyroid-stimulating hormone receptors, and Hashimoto's thyroiditis, characterized by the presence of T helper cells and autoantibodies responding to thyroid proteins and cells (14). Systemic autoimmune diseases are not organ-specific, and can result in widespread damage throughout the body. Examples of systemic autoimmune diseases include: scleroderma, characterized by autoantibodies to nuclei, heart, lungs, gastrointestinal tract and kidneys, Sjögren's syndrome (primary or secondary), characterized by autoantibodies to the salivary gland, liver, kidney, thyroid, and ankylosing spondylitis, characterized by immune complexes responding to vertebrae (14).

Role of the major histocompatibility complex (MHC) in autoimmune diseases

Current research into the etiology of autoimmune disease examines hypotheses involving environmental, infectious and genetic factors (1). In particular, a strong genetic component is indicated in the etiology of autoimmune diseases, with the strongest evidence observed between susceptibility to autoimmune disease and genes (loci) in the major histocompatibility complex (MHC) region. The MHC on chromosome 6p21.3 is a gene-dense region of the human genome spanning approximately 4.5 mega base pairs (Mb) of DNA and encoding more than 180 expressed genes (15, 16). Forty percent of the expressed loci have functions related to immune activation and response. These include the highly polymorphic class I and II human leukocyte antigen (HLA) membrane glycoproteins (classical HLA loci) that present peptides for recognition by T lymphocytes (a type of white blood cell that forms in the thymus and has a special receptor on its cell surface (T cell receptor (TCR)). The MHC region exhibits the strongest linkage disequilibrium (LD) observed in the human genome (17). LD refers to alleles being inherited together on a chromosome over generations in a population (18).

Evidence for the genetic contribution of the MHC to autoimmune diseases is substantial. Genetic association tests of diseased cases and healthy controls and/or families with one or more diseased cases can be used to investigate whether a specific allele (a variant of a gene) occurs more often in individuals with a specific disease than in healthy individuals (18). The most dramatic example illustrating association between the MHC and autoimmunity is the association between ankylosing spondylitis and the class I *HLA-B*27* allele, which has an odds ratio (OR) of around 171 (19). Interestingly, association between disease susceptibility and genes in the HLA region and other genes with immune-related functions can illuminate whether a disease has an autoimmune etiology, as demonstrated by the strong association between narcolepsy and both the class II *HLA-DQB1*0602* allele and the T-cell alpha receptor locus (20). In the current genetic study, we focused on four of the most common autoimmune diseases known to have a strong genetic component: RA, SLE, MS and type 1 diabetes (T1D) (Table 1).

Public health significance of RA

RA is the most common systemic autoimmune disease with a worldwide prevalence approaching one percent (21, 22). It is a chronic inflammatory disease with the potential to cause substantial disability, primarily as a result of the erosive and deforming processes in joints, and is associated with increased mortality, particularly among individuals who develop extra-articular manifestations (outside of the joints) (23, 24). Prevalence estimates of RA (in 100,000 persons) range from 381 in Denmark; 310 to 810 in France, Hungary, Spain, Turkey, Greece and the United Kingdom (U.K.); 120-280 in Thailand, Philippines, Vietnam and China; 510-550 in India and Pakistan; and 197 in Argentina (as reviewed by Cooper *et al*. (2009)) (2, 25-37). RA is rare in Africa (38). No difference in prevalence has been observed between Americans of European and African descent (39).

About 1.5 million American adults suffer from RA, a decrease in the estimate of 2.05 million from the 1980s, due to a decline in prevalence of RA (39-41). In the U.K. the prevalence of RA declined in women after the 1950s (32). In the U.S. and northern Europe, a decline in RA prevalence appears to have occurred primarily during the 1970s and early 1980s (40, 42-45). The cause of this decline is unknown and may possibly be environmental (44). On the other hand,

recent research indicates that incidence and prevalence of RA has actually increased in American women between 1995 and 2007, due to unknown reasons (41).

In the U.S., about 80 percent of RA patients suffer from limitations in function, and the economic impact for medical and surgical treatment and lost wages due to disability adds up to billions of dollars per year (1). Unfortunately, mortality rates have not improved over the past several decades: mortality rates for female and male RA patients are 2.4 and 2.5 per 100 person-years, respectively, compared to 0.2 and 0.3 per 100 person-years in the general female and male populations, respectively (46). The relatively high prevalence of RA, in conjunction with the potentially debilitating impact on the health status of patients, results in tremendous associated costs to patients, their families and society (47, 48). RA costs the US an estimated total of $39.2 billion a year due to excess health care ($8.4 billion), other RA consequences ($10.9 billion), quality-of-life deterioration ($10.3 billion) and premature mortality ($9.6 billion) (49).

Clinical characteristics of RA

Late-adult onset RA usually occurs in middle age and is more common in older people, with a mean age at onset of 58 years (standard deviation (s.d.) ± 16 years) (3, 46, 50). Juvenile RA (JRA) appears by the age of 16 years, with a mean age at onset of 8 (s.d. ± 5 years) (3, 51, 52). In the U.S., about 30,000 to 50,000 RA patients are children (1). The current genetic study was restricted to late-adult onset RA. Females are more commonly affected than males, with a female to male ratio of up to three to one.

RA is characterized by a range of clinical manifestations that result in variations in the RA phenotype expressed by affected individuals. In addition, the risk of lymphoma is increased, and studies show that this increased risk is not due to anti-tumor necrosis factor (anti-TNF) or methotrexate therapy (53). The characteristic erosive process in joints develops in most, but not all, RA patients. Patients with erosive disease experience more disability and poorer outcomes generally compared to RA patients without erosive disease (54, 55). Greater disability and pain have been observed in African American and Latino American RA patients compared to European American RA patients (56). The major autoantibodies in RA are called rheumatoid factor (RF) and react with the immunoglobulin G (IgG) antibody. RF is present in about 80 percent of RA patients. RF-positive patients experience increased disability, joint damage and mortality compared to RF-negative patients (57, 58). Recent data suggest that the presence of anti-cyclic citrullinated peptide (anti-CCP) autoantibodies is also associated with increased joint destruction in RA (59). Presumably, this clinical heterogeneity reflects differences in underlying disease mechanisms, which may be based partly on genetic differences. A growing body of evidence supports this hypothesis (60-62). RF-positivity, anti-CCP positivity and erosive disease are common secondary phenotypes that are strongly associated with disease outcome and have evidence of specific genetic associations.

RA is diagnosed by the American College of Rheumatology (ACR) criteria, which consists of four or more of the following: 1) morning stiffness in and around joints lasting at least one hour before maximal improvement; 2) soft tissue swelling (arthritis) of three or more joint areas observed by a physician; 3) arthritis of the proximal interphalangeal, metacarpophalangeal, or wrist joints; 4) symmetric arthritis; 5) rheumatoid nodules; 6) the presence of rheumatoid factor; and 7) radiographic erosions and/or periarticular osteopenia in

hand and/or wrist joints (63). In order for criteria one through four to count toward the diagnosis they must have been present for ≥6 weeks (63). The sensitivity and specificity of the ACR criteria for RA is 91 to 94 percent and 89 percent, respectively (63).

Evidence for a genetic component in RA

Although the etiology of RA remains unknown, it is clearly complex with important contributions from both genetic and non-genetic factors. A significant genetic contribution to RA development is well established and accounts for an estimated 60 percent of disease risk (21). For example, in a quantitative analysis of Finnish and English twins, 50 to 60 percent of RA in twins was explained by shared genetic effects (47). Several full genome screens to identify disease genes have been performed in families with multiple RA cases (42, 64-67). Their results underscore the importance of the MHC region as harboring the major genetic risk factor(s) for RA susceptibility. MHC genes, particularly those in the class II HLA region, account for an estimated 30 to 50 percent of the genetic component in northern Europeans, suggesting that non-MHC genes also contribute to disease risk (43, 68, 69). The MHC class II gene *HLA-DRB1* demonstrates the strongest association with RA, highlighting antigen presentation and subsequent T cell activation as a potential pathway in RA pathogenesis (43). All RA-associated *HLA-DRB1* alleles (**0101, *0102, *0104, *0401, *0404, *0405, *0408, *0413, *0416, *1001*) encode a shared epitope (SE) not present on non-RA associated alleles (68). The highest-risk SE alleles include *DRB1*0401*, **0404*, **0405*, **0408* and **0409* alleles.

There is evidence for other RA susceptibility loci in the MHC, besides *HLA-DRB1* (43). Conditional analysis of the MHC and the SE has identified three additional SNPs in the MHC associated with RA: one SNP in the HLA class II region between genes butyrophilin-like 2 (MHC class II associated) (*BTNL2*) and *HLA-DRA*, and two SNPs in the HLA class I region near *HLA-C* and the tripartite motif-containing 39 gene (*TRIM 39*) (70). Previous RA studies have implicated *HLA-C* and *BTNL2* (whose association may be due to the linkage disequilibrium displayed with predisposing *HLA DQB1-DRB1* haplotypes) but Taylor et al. are the first to report evidence of association with *TRIM39* (42, 70-72). A haplotype is a set of alleles inherited together on a chromosome. Additional genetic risk loci not located in the MHC have been identified for RA, including protein tyrosine phosphatase, non-receptor type 22 (*PTPN22*, 1p13), cytotoxic T-lymphocyte-associated protein 4 (*CTLA4*, 2q33), peptidyl arginine deiminase, type IV (*PADI4*, 1p36), signal transducer and activator of transcription 4 (*STAT4*, 2q32), TNF receptor-associated factor 1 (*TRAF1-C5*, 9q33), CD40 molecule, TNF receptor superfamily member 5 (*CD40*, 20q12-13), v-rel reticuloendotheliosis viral oncogene homolog (avian) (*REL*, 2p13), tumor necrosis factor, α-induced protein 3 (*TNFAIP3*, 6q23), and Fc receptor-like 3 (*FCRL3*, 1q21-22) (73-79). Interestingly, evidence for association between *HLA-DRB1* and *PTPN22* interaction and susceptibility to anti-CCP positive RA has been reported (80).

Potential environmental risk factors still under investigation include diet, adverse pregnancy outcomes, obesity and recent infections (55, 56). It has been over a decade since evidence for association between exposure to tobacco smoke and RA was first reported, and recent studies support this association and suggest association with interaction between exposure to tobacco smoke and genetic risk factors (57-62). In the Nurses' Health Study, cigarette smoking (past or present) was estimated to elevate risk for RA in women by about 40 percent (63).

Public health significance of SLE

SLE is the second most common autoimmune disease affecting women in their childbearing years (next only to autoimmune thyroid disease) (81, 82). As the prototypic systemic autoimmune disease, lupus may involve virtually every organ system (81). Prevalence is an estimated 52.2 per 100,000 individuals worldwide and an estimated 40 per 100,000 individuals in the U.S. (83, 84). Prevalence estimates of RA (in 100,000 persons) range from 32 in Denmark; 34 to 150 in the U.S., Spain and Greece; 42 in Canada; 19 in Saudi Arabia; 45 in Australia; and 93 in Australia (Indigenous) (as reviewed by Cooper *et al*. (2009)) (3, 25, 85-91). Prevalence and incidence rate estimates vary by race and ethnicity and are highest in minority populations (92, 93). Although survival has improved over the past five decades, SLE patients are still three to five times more likely to die compared to the general population (81, 82). There are about 322,000 Americans suffering from SLE: 161,000 Americans with confirmed SLE (6.8% males and 49.7% females of European ancestry; 4.3% males and 34.8% females of African ancestry; and 4.3% persons of other ancestry), and another 161,000 Americans with suspected SLE (40). The costs associated with treating SLE complications are tremendous. For example, the economic impact for children with childhood-onset SLE in the U.S. is an estimated $146 to $650 million per year (94).

Clinical characteristics of SLE

SLE usually occurs in middle age, with a mean age at onset of 40 years (s.d. ± 10 years) (3, 93, 95-97). Women are affected more often than men (about 9 to one); the difference is most striking in the childbearing years, with a female to male ratio of up to 12 to one (83, 94, 98). Diagnostic tests are based on the presence of four out of the following 11 ACR criteria and achieve specificity and sensitivity greater than 95 percent: malar rash, discoid rash, photosensitivity, oral ulcers, arthritis, serositis, renal disorder, neurologic disorder, hematologic disorder, immune disorder and an abnormal antinuclear antibody (ANA) titer (99, 100). Multiple autoantibodies are present in SLE patients (ANA, anti-Ro, anti-La, antiphospholipid, anti-double-stranded DNA (dsDNA), anti-Sm, and anti-nuclear ribonucleoprotein), and autoantibodies that are less specific for SLE (ANA, anti-Ro, anti-La, antiphospholipid) can appear years before symptoms begin (101). Anti-dsDNA and anti-Sm are usually seen only in SLE patients. Complications include renal disease, musculoskeletal complications, osteoporosis, osteonecrosis, malignancy, cardiovascular disease, psychiatric disease and serious infections.

Lupus nephritis (LN) is the most serious SLE complication and becomes clinically evident in up to half of patients (102, 103). Even among those treated with combinations of cytotoxic drugs and corticosteroids, one-fifth of patients with LN progress to end stage renal disease (104). Thus, LN is a marker of severe disease, a strong predictor of adverse outcomes, and a leading cause of damage associated with SLE (105-108).

Evidence for a genetic component in SLE

The results of twin and family studies implicate genetic factors in disease etiology. There is a higher concordance rate for SLE between monozygotic twins (24 to 69 percent) compared to dizygotic twins (2 to 9 percent) (109, 110). About 10 to 12 percent of SLE patients have a first-degree relative with SLE; familial aggregation studies indicate that SLE risk for the sibling of an affected individual is at least 10-fold greater than the general population (111). The class II HLA genes *DRB1* and/or *DQB1* demonstrate the strongest association. The most consistent

associations are with the HLA class I and II genes and include *HLA-A*01, B*08* and *DRB1*0301, *1501* and **08* (112, 113). Genes within the HLA class III and extended MHC regions have also been strongly implicated, and may be independent of HLA class II. Some studies suggest that the tumor necrosis factor α (TNF-α) and complement *C4A* genes are independent of the *DR3* haplotype (*DRB1*0301*/*DQB1*0201*) or interact with the *DR3* haplotype (114, 115). Recently, a high-density SNP screening of the MHC replicated association between SLE risk and *DRB1*0301* (OR = 2.21) and **1501* (OR = 1.3), and also reported a protective effect of *DRB1*1401* (OR = 0.5) (116). Barcellos et al. also reported associations independent of *DRB1* for the *OR2H2* (extended class I), *CREBL1* (class III), *DQB2* (class II) and *MICB* (class I) genes (116). Exposure to tobacco smoke has been implicated with both increased SLE risk (OR = 1.5) and poorer clinical course (117-121). Sex hormones also appear to play a role in SLE (122). Other environmental risk factors under investigation are occupational exposures to silica, exposure to infection in childhood and viral infections (as reviewed by Molokhia *et al*. (2006)) (123-126).

Evidence for a genetic component in LN

Like SLE, family studies implicate an important role for genetic factors in LN and other forms of renal disease (127, 128). Furthermore, studies in both humans and animal models indicate that the MHC is a risk factor for both SLE and LN, and suggest that genes mediating LN and end-organ damage are distinct from genes influencing susceptibility to SLE (127, 129-132). For example, at least one gene in the *Sle1d* mouse locus critical for LN is specific to autoimmune disease response and end-organ damage (132). Another animal study highlights loci linked to LN and mortality, but not to autoantibody production; this also suggests that discrete sets of genes regulate the development or progression of renal disease as opposed to disease initiation (131). Studies in humans also support a role for specific genetic polymorphisms, particularly Fcγ receptor polymorphisms, in susceptibility to LN (130, 133-136).

Public health significance of MS

MS is a chronic inflammatory and organ-specific autoimmune disorder of the central nervous system (CNS). Worldwide there are 2.5 million MS cases (137). Prevalence estimates of MS (in 100,000 persons) range from 182 in Denmark; 177 to 358 in the U.S. and Canada; 100 in Canada (Indigenous); 121 to 200 in Italy, Greece, France and Ireland; 46 to 50 in Norway, Portugal and New Zealand; four to 20 in Colombia, Brazil and Argentina; 13 in Japan; 11 to 62 in Israel, Kuwait, Jordan and Iran; and 101 in Turkey (as reviewed by Cooper *et al*. (2009)) (2, 25, 138-160). There are an estimated 400,000 Americans suffering from MS, with an estimated total lifetime cost per MS patient of $2.1 million dollars (161, 162). MS is relatively common in individuals of Northern European ancestry, with a prevalence of one in a 1,000, and a sex-specific incidence of 3.6 per 100,000 person-years in females and two per 100,000 person-years in males. There is a north-south risk gradient in the U.S. and Europe (163). MS is quite rare in individuals of African and Asian ancestry. For example, prevalence in Asians is two in 100,000.

Clinical characteristics of MS

MS is characterized by inflammation and demyelination of the CNS (especially the spinal cord, optic nerves, brain stem and cerebrum), astrogliosis and varying degrees of axonal pathology (163, 164). Myelin is a fatty protective layer around the outside of an axon, which is a

long, thin, wire-like part of a neuron (nerve cell) that can transmit signals. MS exhibits differing clinical types of disease progression, measured by increasing disability over time. Typically, patients present with a relapsing-remitting course, which consists of neurological attacks followed by partial or complete recovery; half of these patients continue on to develop a secondary progressive disease course, where the disease progresses without relapses or remissions (137, 163). The other type of MS disease course is primary progressive, where the neurologic function of patients slowly deteriorate over time from the beginning without any relapses or remissions; this presents in about 10 percent of patients (137, 163). Progressive-relapsing MS is the rarest type, occurring in 5 percent of patients; the disease steadily worsens from onset but has distinct relapses of deterioration, which may or may not be followed by partial recovery (137). Common symptoms in MS include sensory, coordination and mobility disturbances; bowel, bladder and sexual dysfunction; and episodes of impaired vision and eye pain due to optic neuritis (163). Optic neuritis is inflammation of the optic nerves, which connects the eye to the brain. The female to male ratio is around two or three to one. The typical age at onset ranges from 20 to 40 years, with a mean age at onset of 29 years in non-primary progressive patients and 39 years in primary-progressive MS (165).

Evidence for a genetic component in MS

Familial aggregation studies MS risk for the sibling of an affected individual is 20 to 40-fold greater than the general population. Disease concordance in monozygotic (identical) twins (25 to 30 percent) is much greater than in dizygotic (fraternal or nonidentical) twins (two and a half to four percent). The risk for full siblings is greater than the risk for half-siblings, and the risk for biological children is greater than the risk for adopted children. This indicates that genetics plays a primary role, but it is believed that environmental influences also affect MS risk. Known risk factors for MS are age, gender, race/ethnicity and exposure to tobacco smoke (OR = 1.5). It has been hypothesized that viral infections and vitamin D insufficiency may contribute to MS susceptibility but these risk factors have not been well established. The largest known risk factor for MS is the *DRB1*1501* allele (frequency 0.25, OR = 2). Genomewide association studies have also established the cytokine receptors Interleukin 2 receptor α chain (*IL2RA,* 10p15) rs2104286 *T* (frequency 0.75, OR = 1.2) and Interleukin 7 receptor (*IL7RA,* 5p13) rs6897932 *C* (frequency 0.75, OR = 1.2) as well as the pattern recognition sugar-binding receptor (*CLEC16A,* 16p13*)* rs6498169 *G* (frequency 0.37, OR = 1.14) as MS risk loci. Like RA and SLE, exposure to tobacco smoke has been implicated in MS. Environmental risk factors also under investigation include decreased sunlight exposure and viral infections (166-169).

Public health significance of T1D

T1D (or insulin-dependent diabetes mellitus (IDDM)) is an organ-specific autoimmune disease characterized by chronic T-cell mediated destruction of pancreatic insulin-producing β-cells (170). Prevalence estimates of T1D (in 100,000 persons) range from 946 in Denmark; 340 to 570 in the U.K., Sweden and Australia; and 118 in Lithuania for all ages; and 87 to 120 in Spain and Germany; 227 to 355 in the U.S. and New Zealand; 70 in American Indians; 31 in the Bahamas; and 110 to 270 in Kuwait and Saudi Arabia for individuals with an age of onset younger than 20 years old (as reviewed by Cooper *et al*. (2009)) (2, 25, 171-182). There are 300,000 to 500,000 T1D patients in the U.S., including 123,000 patients that are younger than 20 years old (1). Incidence in the U.S. is estimated to be ~15 in 100,000 children per year; however, it varies widely around the world and has been increasing over the past decade (183).

Clinical characteristics of T1D

In T1D, T helper lymphocytes and autoantibodies react to and destroy pancreatic β-cells which are responsible for producing the insulin hormone needed for regulating blood glucose levels (14). Decreased insulin levels result in increased blood glucose levels. This disorder is distinct from the more common Type 2, non-insulin-dependent diabetes (NIDDM) which typically presents in adulthood and is associated with obesity. Onset of T1D usually occurs in childhood (mean age at onset of 10 years (s.d. ± 4 years)) (3, 184, 185). Childhood-onset T1D is not more common in females, unlike most autoimmune diseases (one to one female to male ratio), and adult-onset T1D is more common in males (one to two female to male ratio). T1D accounts for 5 to 10 percent of the diabetes in the U.S. Complications include metabolic dysfunction (e.g. ketoacidosis (toxic levels of ketone in the body caused by breaking down fat instead of sugar for energy that can lead to coma and death), increased urine production), renal disease, vascular disease (stroke, heart disease, high blood pressure), nervous system disease, gangrene, lower-limb amputations, periodontal (gum) disease and blindness (14).

Evidence for a genetic component in T1D

Although the etiology of T1D remains unknown, evidence for genetic susceptibility is well established (186, 187). Concordance for T1D in monozygotic twins is 70 percent compared to just 13 percent in dizygotic twins; the relative risk for sibs (λ_s) is approximately 15 in the U.S. Caucasian population (188). The HLA class II genes *HLA-DRB1, DQA1* and *DQB1* in the MHC region appear to be directly involved; the HLA region accounts for 40 to 50 percent of the genetic susceptibility in individuals of Northern European descent (189). The majority of individuals of European descent with T1D carry the *HLA-DR3* (*DRB1*0301-DQA1*0501-DQB1*0201*) or *DR4* (*DRB1*04-DQA1*0301-DQB1*0302*) class II haplotype, and approximately 30 to 50 percent of individuals are *DR3/DR4* heterozygotes (190). *DR3/DR4* heterozygosity confers the highest diabetes risk (191). Different class II HLA associations with T1D are present in non-European populations (192). Class I *HLA-B* has also been associated with T1D risk, specifically the *B*39* and *B*18* alleles (193, 194). Interestingly, the class II *HLA-DR2 (DRB1*1501-DQB1*0602*) haplotype is protective in all populations studied to date (195). Additional non-MHC genetic risk factors for T1D include *PTPN22* (1p13), *CTLA4* (2q33) and *IDDM2* (11p15) (196-198). Environmental factors under investigation include high intake of nitrites during childhood, viral infections, timing of the introduction of solid cereal into the infant diet and consumption of cow's milk during early childhood (as reviewed by Adeghate *et al*. (2006)) (199-206).

CIITA ENCODES AN IMPORTANT TRANSCRIPTION FACTOR FOR HLA CLASS II GENE EXPRESSION

CIITA

The MHC class II transactivator (*CIITA* or *MHC2TA*) gene, located on chromosome 16p13, is a global regulator of expression of proteins involved in antigen presentation and processing, including MHC class II molecules (207-209). *CIITA* has 20 exons spanning 47.8 Mb (10,878 to 10,926 kilo base pairs (kb)). The CIITA protein encoded by the *CIITA* does not bind DNA, but instead serves as a chaperone for assembly of several transcription factors at MHC promoters (210). In addition, it may also have a modulatory role in regulating expression of other

genes, including some phosphatases, kinases and genes involved in cell signaling. Defects in *CIITA* lead to a rare immunodeficiency called bare lymphocyte syndrome (BLS) (211). *CIITA* activity is regulated at the transcriptional level under the control of four different promoters in humans, which span as 12 kb region. Due to the prominent role of the CIITA protein in immunoregulation, the *CIITA* locus is an attractive candidate for genetic studies of autoimmune diseases and other inflammatory conditions for which HLA associations have been well established. Class II gene expression is regulated almost exclusively by *CIITA*, and differential expression of class II genes has shown evidence for association with RA susceptibility and progressive disease in a study with direct implications for *CIITA* (212). Recent genetic linkage studies in RA families provide evidence for a disease susceptibility locus on chromosome 16p. Fisher et al. (2003) conducted a meta-analysis of four RA genome-wide linkage studies (64, 65) and reported strong evidence for linkage on chromosome 16p (P=0.004), in addition to the MHC region (P = 0.00002). This result has been supported by a genome-wide meta-analysis (213) of additional linkage studies (64-67) (chromosome 16p (P<0.01). A recent linkage analysis of RA in American and British families yielded evidence for interaction between *HLA-DRB1* and loci on chromosome 16p (P = 0.0002) (214). Swanberg et al. (2005) reported that the rs3087456 (-168 A/G) polymorphism, located in the type III promoter, was associated with RA and MS, and functional studies in animals and humans provided evidence that it determined expression in antigen-presenting cells, including both B and activated T lymphocytes (215).

CLEC16A

Adjacent to CIITA on chromosome 16p13 lies the C-type lectin domain family 16, member A gene (*CLEC16A*, previously called *KIAA0350*). *CLEC16A* spans 237.7 kb and encodes a sugar-binding receptor that contains a putative immunoreceptor tyrosine-based activation motif (216). C-type lectin receptors can be expressed on dendritic cells to distinguish between self and non-self glycoproteins, and may be involved in immune activation and peripheral tolerance (217, 218). Recently, GWA studies have identified the sugar binding receptor gene *CLEC16A* as a novel risk locus for both T1D and MS, and this association has since been replicated in independent samples (216, 219-224).

HLA COMPATIBILITY, PARENT-OF-ORIGIN AND NONINHERITED MATERNAL ANTIGEN EFFECTS

HLA Compatibility

In addition to specific genetic effects described above, HLA loci within the MHC may also influence autoimmune disease through other inherited or noninherited mechanisms. Further, HLA compatibility between a mother and her offspring may contribute to susceptibility to autoimmunity, possibly because HLA similarity between mother and fetus may promote the persistence of maternal cells in the fetus, also known as ‘maternal microchimerism’ (MMc), following pregnancy. MMc and a role for HLA alleles mediating this effect has been implicated in several autoimmune diseases (225-227). Maternal-offspring cell trafficking is common and bidirectional; maternal nucleated cell and plasma DNA transfers into fetal circulation in 24 and 30 percent of offspring, respectively (228). Maternal-offspring effects can present as excess HLA compatibility between the mother and affected offspring or excess maternal homozygosity. Non-host exposure during fetal development and MMc may play a role in autoimmune disease pathogenesis (229-232). Several biological hypotheses have been proposed, where increased

compatibility could result in a small number of non-host cells that could ultimately 1) cause dysregulation among host cells, 2) lead to presentation of non-host peptides by host cells to other host cells, 3) inactivate T lymphocytes upon interaction, or 4) undergo differentiation and become targets of a later immune response (210, 213-215). Evidence for increased maternal-offspring HLA class II compatibility has been previously reported for both SLE and systemic sclerosis (SSc), suggesting that HLA class II loci may be involved through an undefined pathway dependent on maternal-offspring compatibility (226, 233, 234). In addition, female RA patients with RA-associated *HLA-DRB1* alleles exhibited microchimerism more often and at higher levels than healthy women; the presence and level of microchimerism with non-RA-associated *DRB1* alleles did not differ between female RA patients and healthy women (235). Increased *DRB1* bidirectional compatibility in SLE patients (OR = 5, $P = 0.006$) has been previously reported (226).

Parent-of-Origin Effects

One potential mechanism influencing disease susceptibility is 'genomic imprinting', due to epigenetic modification of the genome (236). This modification results in unequal transcription of parental alleles and subsequent allele expression, depending on whether alleles were transmitted maternally or paternally. Differences in maternal and paternal transmission rates, or parent-of-origin effects, have been suggested previously in autoimmune disease. Parent-of-origin effects, potentially operating through imprinting, have been observed previously in MS and T1D with respect to the inheritance of HLA class II alleles. Excess paternal transmission of *DR3* and excess maternal transmission of *DRB1*15* have been reported in MS (237, 238). Excess paternal transmission of *DR4* and excess maternal transmission of *DQA1*0301-DQB1*0302* have been reported in T1D (239, 240). However, other T1D studies have reported no parent-of-origin effects, even for T1D HLA risk alleles (241-243).

Noninherited Maternal Antigen (NIMA) Effects

In addition to the potential role for MMc mediated by HLA alleles in autoimmunity, the developing immune system of the fetus is also directly exposed to noninherited maternal antigens (NIMA) in utero (231, 244, 245). Non-host exposure during fetal development and potential long-term persistence of maternal cells in offspring may play a role in autoimmune disease pathogenesis (229-232). Exposure to NIMA may shape the immune repertoire of the offspring; this lifelong influence on the immune system may tolerize or predispose to future autoimmune reactions (246-248). A tolerogenic effect may explain the longer survival of renal transplants from sibling donors expressing NIMA vs. noninherited paternal HLA (249). In the precyclosporine era, breastfeeding exposure was associated with improved graft survival in recipients of maternal kidney transplants (250, 251). A highly immunogenic heart allograft mouse model in which both in utero exposure and milk feeding were required for the NIMA effect confirmed the role of breast milk in this observation (252). The current study tested the hypothesis that both cells and antigens of the mother may modulate the antigen-specific reactivity of the fetal immune system. In addition to maternal-offspring cell trafficking and oral exposure through breast milk, NIMA effects may occur through MMc. Maternal cells have been detected in offspring several decades following birth (253). Compared to healthy women, female SSc patients have increased frequencies of maternal cells in their peripheral blood cells (227). Both risk and protective NIMA effects have been previously reported for RA (254-256).

RESEARCH AIMS TO BE ADDRESSED IN THIS DISSERTATION

This dissertation tested hypotheses that:

(1) The *CIITA* -168A/G promoter polymorphism influences RA risk;

(2) Genetic variation in *CIITA* influences RA risk;

(3) Genetic variation in *CLEC16A* influences RA risk;

(4) Genetic variation in *CIITA* influences SLE risk, lupus nephritis, and other secondary clinical phenotypes;

(5) Genetic variation in *CIITA* influences MS risk, and this risk is dependent on the presence of MS risk allele *DRB1*1501*;

(6) *HLA-DRB1* influences SLE through maternal-offspring HLA compatibility, parent-of-origin and noninherited maternal antigen effects in addition to genetic transmission of particular risk alleles; and

(7) Classical HLA loci influence T1D through maternal-offspring HLA compatibility, parent-of-origin and noninherited maternal antigen effects in addition to genetic transmission of particular risk alleles.

My objective, to examine the role of *CIITA* variation in SLE, RA and MS, the role of *CLEC16A* variation in RA, and HLA-related risk factors in SLE and T1D is motivated by the desire to increase our understanding of the biological basis of disease susceptibility and pathogenesis. Such understanding may possibly help contribute to disease prevention and therapy and improved disease prognosis and outcomes in the future.

REFERENCES

1. NIH. Progress in Autoimmune Diseases Research. In: Committee ADC, editor.; 2005.
2. Cooper GS, Bynum ML, Somers EC. Recent insights in the epidemiology of autoimmune diseases: improved prevalence estimates and understanding of clustering of diseases. J Autoimmun 2009;33(3-4):197-207.
3. Cooper GS, Stroehla BC. The epidemiology of autoimmune diseases. Autoimmun Rev 2003;2(3):119-25.
4. Jacobson DL, Gange SJ, Rose NR, Graham NMH. Epidemiology and Estimated Population Burden of Selected Autoimmune Diseases in the United States. Clinical Immunology and Immunopathology 1997;84(3):223-243.
5. Walsh SJ, Rau LM. Autoimmune diseases: a leading cause of death among young and middle-aged women in the United States. Am J Public Health 2000;90(9):1463-6.
6. Feinstein A. The neuropsychiatry of multiple sclerosis. Can J Psychiatry 2004;49(3):157-63.
7. Hanly JG, Fisk JD, McCurdy G, Fougere L, Douglas JA. Neuropsychiatric syndromes in patients with systemic lupus erythematosus and rheumatoid arthritis. J Rheumatol 2005;32(8):1459-6.
8. Harboe E, Tjensvoll AB, Maroni S, Goransson LG, Greve OJ, Beyer MK, et al. Neuropsychiatric syndromes in patients with systemic lupus erythematosus and primary Sjogren syndrome: a comparative population-based study. Ann Rheum Dis 2009;68(10):1541-6.
9. Committee ACR. The American College of Rheumatology nomenclature and case definitions for neuropsychiatric lupus syndromes. Arthritis Rheum 1999;42(4):599-608.
10. Hanly JG, Urowitz MB, Su L, Bae SC, Gordon C, Wallace DJ, et al. Prospective analysis of neuropsychiatric events in an international disease inception cohort of patients with systemic lupus erythematosus. Ann Rheum Dis;69(3):529-35.
11. Dunlop DD, Lyons JS, Manheim LM, Song J, Chang RW. Arthritis and Heart Disease as Risk Factors for Major Depression: The Role of Functional Limitation. Medical Care 2004;42(6):502-511.
12. Margaretten M, Yelin E, Imboden J, Graf J, Barton J, Katz P, et al. Predictors of depression in a multiethnic cohort of patients with rheumatoid arthritis. Arthritis Rheum 2009;61(11):1586-91.
13. WHO. The global burden of disease: 2004 update. Geneva; 2008.
14. Kindt T.J., Goldsby RA, Osborne BA. Kuby Immunology. Sixth ed. New York: W.H. Freeman and Company; 2007.
15. Complete sequence and gene map of a human major histocompatibility complex. The MHC sequencing consortium. Nature 1999;401(6756):921-3.
16. Stewart CA, Horton R, Allcock RJ, Ashurst JL, Atrazhev AM, Coggill P, et al. Complete MHC Haplotype Sequencing for Common Disease Gene Mapping. Genome Res 2004;14(6):1176-87.
17. Kelly AP, Trowsdale J. The MHC paradigm: genetic variation and complex disease. In: Wright A, Hastie N, editors. Genes and Common Diseases. Cambridge: Oxford University Press; 2007. p. 142-151.
18. King RA, Rotter JI, Motulsky AG. Approach to Genetic Basis of Common Diseases. In: King RA, Rotter JI, Motulsky AG, editors. The Genetic Basis of Common Diseases. 2nd ed. New York: Oxford University Press, Inc.; 2002.

19. Brown MA, Pile KD, Kennedy LG, Calin A, Darke C, Bell J, et al. HLA class I associations of ankylosing spondylitis in the white population in the United Kingdom. Ann Rheum Dis 1996;55(4):268-70.
20. Hallmayer J, Faraco J, Lin L, Hesselson S, Winkelmann J, Kawashima M, et al. Narcolepsy is strongly associated with the T-cell receptor alpha locus. Nat Genet 2009;41(6):708-11.
21. Gabriel SE. The epidemiology of rheumatoid arthritis. Rheum Dis Clin North Am 2001;27(2):269-81.
22. Silman AJ, Pearson JE. Epidemiology and genetics of rheumatoid arthritis. Arthritis Res 2002;4 Suppl 3:S265-72.
23. Gabriel SE, Crowson CS, Kremers HM, Doran MF, Turesson C, O'Fallon WM, et al. Survival in rheumatoid arthritis: a population-based analysis of trends over 40 years. Arthritis Rheum 2003;48(1):54-8.
24. Turesson C, O'Fallon WM, Crowson CS, Gabriel SE, Matteson EL. Occurrence of extraarticular disease manifestations is associated with excess mortality in a community-based cohort of patients with rheumatoid arthritis. J Rheumatol 2002;29(1):62-7.
25. Eaton WW, Rose NR, Kalaydjian A, Pedersen MG, Mortensen PB. Epidemiology of autoimmune diseases in Denmark. J Autoimmun 2007;29(1):1-9.
26. Andrianakos A, Trontzas P, Christoyannis F, Dantis P, Voudouris C, Georgountzos A, et al. Prevalence of rheumatic diseases in Greece: a cross-sectional population based epidemiological study. The ESORDIG Study. J Rheumatol 2003;30(7):1589-601.
27. Guillemin F, Saraux A, Guggenbuhl P, Roux CH, Fardellone P, Le Bihan E, et al. Prevalence of rheumatoid arthritis in France: 2001. Ann Rheum Dis 2005;64(10):1427-30.
28. Kiss CG, Lovei C, Suto G, Varju C, Nagy Z, Fuzesi Z, et al. Prevalence of rheumatoid arthritis in the South-Transdanubian region of Hungary based on a representative survey of 10,000 inhabitants. J Rheumatol 2005;32(9):1688-90.
29. Carmona L, Villaverde V, Hernandez-Garcia C, Ballina J, Gabriel R, Laffon A. The prevalence of rheumatoid arthritis in the general population of Spain. Rheumatology (Oxford) 2002;41(1):88-95.
30. Akar S, Birlik M, Gurler O, Sari I, Onen F, Manisali M, et al. The prevalence of rheumatoid arthritis in an urban population of Izmir-Turkey. Clin Exp Rheumatol 2004;22(4):416-20.
31. Kacar C, Gilgil E, Tuncer T, Butun B, Urhan S, Arikan V, et al. Prevalence of rheumatoid arthritis in Antalya, Turkey. Clin Rheumatol 2005;24(3):212-4.
32. Symmons D, Turner G, Webb R, Asten P, Barrett E, Lunt M, et al. The prevalence of rheumatoid arthritis in the United Kingdom: new estimates for a new century. Rheumatology (Oxford) 2002;41(7):793-800.
33. Chaiamnuay P, Darmawan J, Muirden KD, Assawatanabodee P. Epidemiology of rheumatic disease in rural Thailand: a WHO-ILAR COPCORD study. Community Oriented Programme for the Control of Rheumatic Disease. J Rheumatol 1998;25(7):1382-7.
34. Dans LF, Tankeh-Torres S, Amante CM, Penserga EG. The prevalence of rheumatic diseases in a Filipino urban population: a WHO-ILAR COPCORD Study. World Health Organization. International League of Associations for Rheumatology. Community Oriented Programme for the Control of the Rheumatic Diseases. J Rheumatol 1997;24(9):1814-9.

35. Minh Hoa TT, Darmawan J, Chen SL, Van Hung N, Thi Nhi C, Ngoc An T. Prevalence of the rheumatic diseases in urban Vietnam: a WHO-ILAR COPCORD study. J Rheumatol 2003;30(10):2252-6.
36. Dai SM, Han XH, Zhao DB, Shi YQ, Liu Y, Meng JM. Prevalence of rheumatic symptoms, rheumatoid arthritis, ankylosing spondylitis, and gout in Shanghai, China: a COPCORD study. J Rheumatol 2003;30(10):2245-51.
37. Chopra A, Patil J, Billempelly V, Relwani J, Tandle HS. Prevalence of rheumatic diseases in a rural population in western India: a WHO-ILAR COPCORD Study. J Assoc Physicians India 2001;49:240-6.
38. Silman AJ, Ollier W, Holligan S, Birrell F, Adebajo A, Asuzu MC, et al. Absence of rheumatoid arthritis in a rural Nigerian population. J Rheumatol 1993;20(4):618-22.
39. Lawrence RC, Helmick CG, Arnett FC, Deyo RA, Felson DT, Giannini EH, et al. Estimates of the prevalence of arthritis and selected musculoskeletal disorders in the United States. Arthritis Rheum 1998;41(5):778-99.
40. Helmick CG, Felson DT, Lawrence RC, Gabriel S, Hirsch R, Kwoh CK, et al. Estimates of the prevalence of arthritis and other rheumatic conditions in the United States. Part I. Arthritis Rheum 2008;58(1):15-25.
41. Myasoedova E, Crowson CS, Kremers HM, Therneau TM, Gabriel SE. Is the incidence of rheumatoid arthritis rising? Results from olmsted county, minnesota, 1955-2007. Arthritis Rheum.
42. Jawaheer D, Seldin MF, Amos CI, Chen WV, Shigeta R, Etzel C, et al. Screening the genome for rheumatoid arthritis susceptibility genes: a replication study and combined analysis of 512 multicase families. Arthritis Rheum 2003;48(4):906-16.
43. Jawaheer D, Li W, Graham RR, Chen W, Damle A, Xiao X, et al. Dissecting the genetic complexity of the association between human leukocyte antigens and rheumatoid arthritis. Am J Hum Genet 2002;71(3):585-94.
44. Uhlig T, Kvien TK. Is rheumatoid arthritis disappearing? Ann Rheum Dis 2005;64(1):7-10.
45. Alamanos Y, Drosos AA. Epidemiology of adult rheumatoid arthritis. Autoimmun Rev 2005;4(3):130-6.
46. Doran MF, Pond GR, Crowson CS, O'Fallon WM, Gabriel SE. Trends in incidence and mortality in rheumatoid arthritis in Rochester, Minnesota, over a forty-year period. Arthritis Rheum 2002;46(3):625-31.
47. Lajas C, Abasolo L, Bellajdel B, Hernandez-Garcia C, Carmona L, Vargas E, et al. Costs and predictors of costs in rheumatoid arthritis: a prevalence-based study. Arthritis Rheum 2003;49(1):64-70.
48. Yelin E, Wanke LA. An assessment of the annual and long-term direct costs of rheumatoid arthritis: the impact of poor function and functional decline. Arthritis Rheum 1999;42(6):1209-18.
49. Birnbaum H, Pike C, Kaufman R, Marynchenko M, Kidolezi Y, Cifaldi M. Societal cost of rheumatoid arthritis patients in the US. Current Medical Research and Opinion;26(1):77-90.
50. Kaipiainen-Seppanen O, Aho K, Isomaki H, Laakso M. Shift in the incidence of rheumatoid arthritis toward elderly patients in Finland during 1975-1990. Clin Exp Rheumatol 1996;14(5):537-42.

51. Peterson LS, Mason T, Nelson AM, O'Fallon WM, Gabriel SE. Juvenile rheumatoid arthritis in Rochester, Minnesota 1960-1993. Is the epidemiology changing? Arthritis Rheum 1996;39(8):1385-90.
52. Moe N, Rygg M. Epidemiology of juvenile chronic arthritis in northern Norway: a ten-year retrospective study. Clin Exp Rheumatol 1998;16(1):99-101.
53. Wolfe F, Michaud K. The effect of methotrexate and anti-tumor necrosis factor therapy on the risk of lymphoma in rheumatoid arthritis in 19,562 patients during 89,710 person-years of observation. Arthritis Rheum 2007;56(5):1433-9.
54. Suarez-Almazor ME, Soskolne CL, Saunders LD, Russell AS. Outcome in rheumatoid arthritis. A 1985 inception cohort study. J Rheumatol 1994;21(8):1438-46.
55. Young A, Bielawska C, Corbett M, Roitt I. A prospective study of early onset rheumatoid arthritis over fifteen years: prognostic features and outcome. Clin Rheumatol 1987;6 Suppl 2:12-9.
56. Bruce B, Fries JF, Murtagh KN. Health status disparities in ethnic minority patients with rheumatoid arthritis: a cross-sectional study. J Rheumatol 2007;34(7):1475-9.
57. van Schaardenburg D, Hazes JMW, de Boer A, Zwinderman AH, Meijers KA, Breedveld FC. Outcome of rheumatoid arthritis in relation to age and rheumatoid factor at diagnosis. J Rheumatol 1993;20:45-52.
58. Combe B, Dougados M, Goupille P, Cantagrel A, Eliaou JF, Sibilia J, et al. Prognostic factors for radiographic damage in early rheumatoid arthritis: a multiparameter prospective study. Arthritis Rheum 2001;44(8):1736-43.
59. Van Gaalen FA, Van Aken J, Huizinga TW, Schreuder GM, Breedveld FC, Zanelli E, et al. Association between HLA class II genes and autoantibodies to cyclic citrullinated peptides (CCPs) influences the severity of rheumatoid arthritis. Arthritis Rheum 2004;50(7):2113-21.
60. Jawaheer D, Lum RF, Amos CI, Gregersen PK, Criswell LA. Clustering of disease features within 512 multicase rheumatoid arthritis families. Arthritis Rheum 2004;50(3):736-41.
61. Gorman JD, Lum RF, Chen JJ, Suarez-Almazor ME, Thomson G, Criswell LA. Impact of shared epitope genotype and ethnicity on erosive disease: a meta-analysis of 3,240 rheumatoid arthritis patients. Arthritis Rheum 2004;50(2):400-12.
62. Huizinga TWJ, Amos CI, van der Helm-van Mil AHM, Chen W, Van Gaalen FA, Jawaheer D, et al. Refining the complex rheumatoid arthritis phenotype based on specificity of the HLA-DRB1 shared epitope for antibodies to citrullinated proteins. Arthritis Rheum 2005;52(11):3433-8.
63. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988;31(3):315-324.
64. Cornelis F, Faure S, Martinez M, Prud'homme J-F, Fritz P, Dib C, et al. New susceptibility locus for rheumatoid arthritis suggested by a genome-wide linkage study. Proc Natl Acad Sci USA 1998;95:10746-50.
65. Jawaheer D, Seldin MF, Amos CI, Chen WV, Monteiro J, Kern M, et al. A genome-wide screen in multiplex rheumatoid arthritis families suggests genetic overlap with other autoimmune diseases. Am J Hum Genet 2001;68:927-936.
66. MacKay K, Eyre S, Myerscough A, Milicic A, Barton A, Laval S, et al. Whole-genome linkage analysis of rheumatoid arthritis susceptibility loci in 252 affected sibling pairs in the United Kingdom. Arthritis Rheum 2002;46(3):632-9.

67. John S, Shephard N, Liu G, Zeggini E, Cao M, Chen W, et al. Whole-genome scan, in a complex disease, using 11,245 single-nucleotide polymorphisms: comparison with microsatellites. Am J Hum Genet 2004;75(1):54-64.
68. Gregersen PK, Silver J, Winchester RJ. The shared epitope hypothesis: an approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis. Arthritis Rheum 1987;30:1205-1213.
69. MacGregor AJ, Snieder H, Rigby AS, Koskenvuo M, Kaprio J, Aho K, et al. Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins. Arthritis Rheum 2000;43(1):30-7.
70. Taylor KE, Criswell LA. Conditional analysis of the major histocompatibility complex in rheumatoid arthritis. BMC Proc 2009;3 Suppl 7:S36.
71. Lee HS, Lee AT, Criswell LA, Seldin MF, Amos CI, Carulli JP, et al. Several regions in the major histocompatibility complex confer risk for anti-CCP-antibody positive rheumatoid arthritis, independent of the DRB1 locus. Mol Med 2008;14(5-6):293-300.
72. Orozco G, Eerligh P, Sanchez E, Zhernakova S, Roep BO, Gonzalez-Gay MA, et al. Analysis of a functional BTNL2 polymorphism in type 1 diabetes, rheumatoid arthritis, and systemic lupus erythematosus. Hum Immunol 2005;66(12):1235-41.
73. Begovich AB, Carlton VE, Honigberg LA, Schrodi SJ, Chokkalingam AP, Alexander HC, et al. A missense single-nucleotide polymorphism in a gene encoding a protein tyrosine phosphatase (PTPN22) is associated with rheumatoid arthritis. Am J Hum Genet 2004;75(2):330-7.
74. Rodriguez MR, Nunez-Roldan A, Aguilar F, Valenzuela A, Garcia A, Gonzalez-Escribano MF. Association of the CTLA4 3' untranslated region polymorphism with the susceptibility to rheumatoid arthritis. Hum Immunol 2002;63(1):76-81.
75. Suzuki A, Yamada R, Chang X, Tokuhiro S, Sawada T, Suzuki M, et al. Functional haplotypes of PADI4, encoding citrullinating enzyme peptidylarginine deiminase 4, are associated with rheumatoid arthritis. Nat Genet 2003;34(4):395-402.
76. Remmers EF, Plenge RM, Lee AT, Graham RR, Hom G, Behrens TW, et al. STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus. N Engl J Med 2007;357(10):977-86.
77. Plenge RM, Seielstad M, Padyukov L, Lee AT, Remmers EF, Ding B, et al. TRAF1-C5 as a Risk Locus for Rheumatoid Arthritis -- A Genomewide Study. N Engl J Med 2007;357(12):1199-1209.
78. Gregersen PK, Amos CI, Lee AT, Lu Y, Remmers EF, Kastner DL, et al. REL, encoding a member of the NF-kappaB family of transcription factors, is a newly defined risk locus for rheumatoid arthritis. Nat Genet 2009;41(7):820-3.
79. Newman WG, Zhang Q, Liu XD, Walker E, Ternan H, Owen J, et al. Rheumatoid arthritis association with the FCRL3-169C polymorphism is restricted to PTPN22 1858T-homozygous individuals in a Canadian population. Arthritis and Rheumatism 2006;54(12):3820-3827.
80. Kallberg H, Padyukov L, Plenge RM, Ronnelid J, Gregersen PK, van der Helm-van Mil AH, et al. Gene-gene and gene-environment interactions involving HLA-DRB1, PTPN22, and smoking in two subsets of rheumatoid arthritis. Am J Hum Genet 2007;80(5):867-75.
81. Kelley W, Harris E, Jr., Ruddy S, Sledge C. Textbook of Rheumatology. Philadelphia: Harcourt Brace Jovanovich, Inc.; 1993.

82. Bongu A, Chang E, Ramsey-Goldman R. Can morbidity and mortality of SLE be improved? Best Practice & Research Clinical Rheumatology 2002;16(2):313-332.
83. Danchenko N, Satia JA, Anthony MS. Epidemiology of systemic lupus erythematosus: a comparison of worldwide disease burden. Lupus 2006;15:308-318.
84. Hochberg MC. The epidemiology of systemic lupus erythematosus. In: Wallace DJ, Hahn BH, editors. Dubois' Lupus Erythematosus. 5th ed. Baltimore: Williams & Wilkins; 1997. p. 49-65.
85. Chakravarty EF, Bush TM, Manzi S, Clarke AE, Ward MM. Prevalence of adult systemic lupus erythematosus in California and Pennsylvania in 2000: estimates obtained using hospitalization data. Arthritis Rheum 2007;56(6):2092-4.
86. Naleway AL, Davis ME, Greenlee RT, Wilson DA, McCarty DJ. Epidemiology of systemic lupus erythematosus in rural Wisconsin. Lupus 2005;14(10):862-6.
87. Lopez P, Mozo L, Gutierrez C, Suarez A. Epidemiology of systemic lupus erythematosus in a northern Spanish population: gender and age influence on immunological features. Lupus 2003;12(11):860-5.
88. Alamanos Y, Voulgari PV, Siozos C, Katsimpri P, Tsintzos S, Dimou G, et al. Epidemiology of systemic lupus erythematosus in northwest Greece 1982-2001. J Rheumatol 2003;30(4):731-5.
89. Peschken CA, Esdaile JM. Systemic lupus erythematosus in North American Indians: a population based study. J Rheumatol 2000;27(8):1884-91.
90. Al-Arfaj AS, Al-Balla SR, Al-Dalaan AN, Al-Saleh SS, Bahabri SA, Mousa MM, et al. Prevalence of systemic lupus erythematosus in central Saudi Arabia. Saudi Med J 2002;23(1):87-9.
91. Bossingham D. Systemic lupus erythematosus in the far north of Queensland. Lupus 2003;12(4):327-31.
92. Hopkinson ND, Doherty M, Powell RJ. Clinical features and race-specific incidence/prevalence rates of systemic lupus erythematosus in a geographically complete cohort of patients. Ann Rheum Dis 1994;53(10):675-80.
93. Johnson AE, Gordon C, Palmer RG, Bacon PA. The prevalence and incidence of systemic lupus erythematosus in Birmingham, England. Relationship to ethnicity and country of birth. Arthritis Rheum 1995;38(4):551-8.
94. Brunner HI, Sherrard TM, Klein-Gitelman MS. Cost of treatment of childhood-onset systemic lupus erythematosus. Arthritis & Rheumatism-Arthritis Care & Research 2006;55(2):184-188.
95. Hopkinson ND, Doherty M, Powell RJ. The prevalence and incidence of systemic lupus erythematosus in Nottingham, UK, 1989-1990. Br J Rheumatol 1993;32(2):110-5.
96. McCarty DJ, Manzi S, Medsger TA, Jr., Ramsey-Goldman R, LaPorte RE, Kwoh CK. Incidence of systemic lupus erythematosus. Race and gender differences. Arthritis Rheum 1995;38(9):1260-70.
97. Uramoto KM, Michet CJ, Jr., Thumboo J, Sunku J, O'Fallon WM, Gabriel SE. Trends in the incidence and mortality of systemic lupus erythematosus, 1950-1992. Arthritis Rheum 1999;42(1):46-50.
98. McCarty DJ, Manzi S, Medsger TA, Ramseygoldman R, Laporte RE, Kwoh CK. Incidence of Systemic Lupus-Erythematosus - Race and Gender Differences. Arthritis and Rheumatism 1995;38(9):1260-1270.

99. Tan EM, Cohen AS, Fries JF, Masi AT, McShane DJ, Rothfield NF, et al. The 1982 revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum 1982;25:1271-1277.
100. Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum 1997;40(9):1725.
101. Arbuckle MR, McClain MT, Rubertone MV, Scofield RH, Dennis GJ, James JA, et al. Development of autoantibodies before the clinical onset of systemic lupus erythematosus. N Engl J Med 2003;349(16):1526-33.
102. Huong DL, Papo T, Beaufils H, Wechsler B, Bletry O, Baumelou A, et al. Renal involvement in systemic lupus erythematosus. A study of 180 patients from a single center. Medicine (Baltimore) 1999;78(3):148-66.
103. Ponticelli C, Moroni G. Flares in lupus nephritis: incidence, impact on renal survival and management. Lupus 1998;7(9):635-8.
104. Golbus J, McCune WJ. Lupus nephritis: classification, prognosis, immunopathogenesis, and treatment. In; 1994. p. 213-242.
105. Ginzler EM, Diamond HS, Weiner M, Schlesinger M, Fries JF, Wasner C, et al. A multicenter study of outcome in systemic lupus erythematosus. I. Entry variables as predictors of prognosis. Arthritis Rheum 1982;25:601-611.
106. Reveille JD, Bartolucci A, Alarcon GS. Prognosis in systemic lupus erythematosus. Negative impact of increasing age at onset, black race, and thrombocytopenia, as well as causes of death. Arthritis Rheum 1990;33(1):37-48.
107. Gladman DD, Urowitz MB. Systemic lupus erythematosus: clinical manifestations. In: Klippel JH, Dieppe PA, editors. Rheumatology. London: Mosby-Year Book Europe Limited; 1994. p. 6.2.1-6.2.20.
108. Boumpas DT, Austin HA, 3rd., Fessler BJ, Balow JE, Klippel JH, Lockshin MD. Systemic lupus erythematosus: emerging concepts. Part 1: Renal, neuropsychiatric, cardiovascular, pulmonary, and hematologic disease. Ann Intern Med 1995;122(12):940-950.
109. Block SR, Winfield JB, Lockshin MD, D'Angelo WA, Christian CL. Studies of twins with systemic lupus erythematosus. Am J Med 1975;59:533-552.
110. Deapen D, Escalante A, Weinrib L, Horwitz D, Bachman B, Roy-Burman P, et al. A revised estimate of twin concordance in systemic lupus erythematosus. Arthritis Rheum 1992;35:311-317.
111. Vyse TJ, Todd JA. Genetic analysis of autoimmune disease. Cell 1996;85(3):311-8.
112. Graham R, Rodine P, Langefeld C, Williams A, Selby S, Ortmann W, et al. Association of HLA class II risk haplotypes with human SLE: striking evidence for a dose effect in multi-center collection. Arthritis Rheum 2003;48(9 suppl):S701.
113. Graham RR, Ortmann W, Rodine P, Espe K, Langefeld C, Lange E, et al. Specific combinations of HLA-DR2 and DR3 class II haplotypes contribute graded risk for disease susceptibility and autoantibodies in human SLE. Eur J Hum Genet 2007;15(8):823-30.
114. Rood MJ, van Krugten MV, Zanelli E, van der Linden MW, Keijsers V, Schreuder GM, et al. TNF-308A and HLA-DR3 alleles contribute independently to susceptibility to systemic lupus erythematosus. Arthritis Rheum 2000;43(1):129-34.
115. Davies EJ, Steers G, Ollier WER, Grennan DM, Cooper RG, Hay EM, et al. Relative contributions of HLA-DQA and complement C4A loci in determining susceptibility to systemic lupus erythematosus. Br J Rheumatol 1995;34:221-225.

116. Barcellos LF, May SL, Ramsay PP, Quach HL, Lane JA, Nititham J, et al. High-density SNP screening of the major histocompatibility complex in systemic lupus erythematosus demonstrates strong evidence for independent susceptibility regions. PLoS Genet 2009;5(10):e1000696.
117. Majka DS, Holers VM. Cigarette smoking and the risk of systemic lupus erythematosus and rheumatoid arthritis. Ann Rheum Dis 2006;65(5):561-3.
118. Costenbader KH, Kim DJ, Peerzada J, Lockman S, Nobles-Knight D, Petri M, et al. Cigarette smoking and the risk of systemic lupus erythematosus: a meta-analysis. Arthritis Rheum 2004;50(3):849-57.
119. Formica MK, Palmer JR, Rosenberg L, McAlindon TE. Smoking, alcohol consumption, and risk of systemic lupus erythematosus in the Black Women's Health Study. J Rheumatol 2003;30(6):1222-6.
120. Ward MM, Studenski S. Clinical prognostic factors in lupus nephritis. The importance of hypertension and smoking. Arch Intern Med 1992;152(10):2082-8.
121. Ghaussy NO, Sibbitt W, Jr., Bankhurst AD, Qualls CR. Cigarette smoking and disease activity in systemic lupus erythematosus. J Rheumatol 2003;30(6):1215-21.
122. Petri M. Sex hormones and systemic lupus erythematosus. Lupus 2008;17(5):412-5.
123. Molokhia M, McKeigue P. Systemic lupus erythematosus: genes versus environment in high risk populations. Lupus 2006;15(11):827-32.
124. Molokhia M, Hoggart C, Patrick AL, Shriver M, Parra E, Ye J, et al. Relation of risk of systemic lupus erythematosus to west African admixture in a Caribbean population. Hum Genet 2003;112(3):310-8.
125. Parks CG, Cooper GS, Nylander-French LA, Sanderson WT, Dement JM, Cohen PL, et al. Occupational exposure to crystalline silica and risk of systemic lupus erythematosus: a population-based, case-control study in the southeastern United States. Arthritis Rheum 2002;46(7):1840-50.
126. Edwards CJ, Syddall H, Goswami R, Goswami P, Dennison EM, Cooper C. Infections in infancy and the presence of antinuclear antibodies in adult life. Lupus 2006;15(4):213-7.
127. Freedman B. Familial aggregation of end-stage renal failure: aetiological implications. Nephrol. Dial. Transplant. 1999;14(2):295-297.
128. Borch-Johnsen K, Norgaard K, Hommel E, Mathiesen ER, Jensen JS, Deckert T, et al. Is diabetic nephropathy an inherited complication? Kidney Int 1992;41(4):719-22.
129. Schelling JR, Zarif L, Sehgal A, Iyengar S, Sedor JR. Genetic susceptibility to end-stage renal disease. Curr Opin Nephrol Hypertens 1999;8(4):465-72.
130. Tsao BP. Genetic susceptibility to lupus nephritis. Lupus 1998;7(9):585-590.
131. Theofilopoulos AN, Kono DH. Genetics of systemic autoimmunity and glomerulonephritis in mouse models of lupus. Nephrol Dial Transplant 2001;16(Suppl 6):65-7.
132. Morel L, Blenman KR, Croker BP, Wakeland EK. The major murine systemic lupus erythematosus susceptibility locus, Sle1, is a cluster of functionally related genes. Proc Natl Acad Sci USA 2001;98(4):1787-92.
133. Salmon JE, Millard S, Schacter LA, Arnett FC, Ginzler EM, Gourley MF, et al. FcgRIIA alleles are heritable risk factors for lupus nephritis in African Americans. J Clin Invest 1996;97(5):1348-1354.
134. Salmon JE, Ng S, Yoo DH, Kim TH, Kim SY, Song GG. Altered distribution of Fcg receptor IIIA alleles in a cohort of Korean patients with lupus nephritis. Arthritis Rheum 1999;42(4):818-823.

135. Norsworthy P, Theodoridis E, Botto M, Athanassiou P, Beynon H, Gordon C, et al. Overrepresentation of the Fcg receptor type IIA R131/R131 genotype in caucasoid systemic lupus erythematosus patients with autoantibodies to C1q and glomerulonephritis. Arthritis Rheum 1999;42(9):1828-1832.
136. Seligman VA, Suarez C, Lum R, Inda SE, Lin D, Li H, et al. The FcgRIIIA-158F allele is a major risk factor for the development of lupus nephritis among Caucasians but not non-Caucasians. Arthritis Rheum 2001;44(3):618-625.
137. Multiple sclerosis: just the facts; 2009.
138. Mayr WT, Pittock SJ, McClelland RL, Jorgensen NW, Noseworthy JH, Rodriguez M. Incidence and prevalence of multiple sclerosis in Olmsted County, Minnesota, 1985-2000. Neurology 2003;61(10):1373-7.
139. Hader WJ, Yee IM. Incidence and prevalence of multiple sclerosis in Saskatoon, Saskatchewan. Neurology 2007;69(12):1224-9.
140. Warren SA, Svenson LW, Warren KG. Contribution of incidence to increasing prevalence of multiple sclerosis in Alberta, Canada. Mult Scler 2008;14(7):872-9.
141. Svenson LW, Warren S, Warren KG, Metz LM, Patten SB, Schopflocher DP. Prevalence of multiple sclerosis in First Nations people of Alberta. Can J Neurol Sci 2007;34(2):175-80.
142. Granieri E, Monaldini C, De Gennaro R, Guttmann S, Volpini M, Stumpo M, et al. Multiple sclerosis in the Republic of San Marino: a prevalence and incidence study. Mult Scler 2008;14(3):325-9.
143. Grimaldi LM, Palmeri B, Salemi G, Giglia G, D'Amelio M, Grimaldi R, et al. High prevalence and fast rising incidence of multiple sclerosis in Caltanissetta, Sicily, southern Italy. Neuroepidemiology 2007;28(1):28-32.
144. Iuliano G, Napoletano R. Prevalence and incidence of multiple sclerosis in Salerno (southern Italy) and its province. Eur J Neurol 2008;15(1):73-6.
145. Papathanasopoulos P, Gourzoulidou E, Messinis L, Georgiou V, Leotsinidis M. Prevalence and incidence of multiple sclerosis in western Greece: a 23-year survey. Neuroepidemiology 2008;30(3):167-73.
146. Debouverie M, Pittion-Vouyovitch S, Louis S, Roederer T, Guillemin F. Increasing incidence of multiple sclerosis among women in Lorraine, Eastern France. Mult Scler 2007;13(8):962-7.
147. Gray OM, McDonnell GV, Hawkins SA. Factors in the rising prevalence of multiple sclerosis in the north-east of Ireland. Mult Scler 2008;14(7):880-6.
148. McGuigan C, McCarthy A, Quigley C, Bannan L, Hawkins SA, Hutchinson M. Latitudinal variation in the prevalence of multiple sclerosis in Ireland, an effect of genetic diversity. J Neurol Neurosurg Psychiatry 2004;75(4):572-6.
149. Dahl OP, Aarseth JH, Myhr KM, Nyland H, Midgard R. Multiple sclerosis in Nord-Trondelag County, Norway: a prevalence and incidence study. Acta Neurol Scand 2004;109(6):378-84.
150. De Sa J, Paulos A, Mendes H, Becho J, Marques J, Roxo J. The prevalence of multiple sclerosis in the District of Santarem, Portugal. J Neurol 2006;253(7):914-8.
151. Chancellor AM, Addidle M, Dawson K. Multiple sclerosis is more prevalent in northern New Zealand than previously reported. Intern Med J 2003;33(3):79-83.
152. Toro J, Sarmiento OL, Diaz del Castillo A, Satizabal CL, Ramirez JD, Montenegro AC, et al. Prevalence of multiple sclerosis in Bogota, Colombia. Neuroepidemiology 2007;28(1):33-8.

153. Callegaro D, Goldbaum M, Morais L, Tilbery CP, Moreira MA, Gabbai AA, et al. The prevalence of multiple sclerosis in the city of Sao Paulo, Brazil, 1997. Acta Neurol Scand 2001;104(4):208-13.
154. Cristiano E, Patrucco L, Rojas JI, Caceres F, Carra A, Correale J, et al. Prevalence of multiple sclerosis in Buenos Aires, Argentina using the capture-recapture method. Eur J Neurol 2009;16(2):183-7.
155. Houzen H, Niino M, Hata D, Nakano F, Kikuchi S, Fukazawa T, et al. Increasing prevalence and incidence of multiple sclerosis in northern Japan. Mult Scler 2008;14(7):887-92.
156. Alshubaili AF, Alramzy K, Ayyad YM, Gerish Y. Epidemiology of multiple sclerosis in Kuwait: new trends in incidence and prevalence. Eur Neurol 2005;53(3):125-31.
157. Etemadifar M, Janghorbani M, Shaygannejad V, Ashtari F. Prevalence of multiple sclerosis in Isfahan, Iran. Neuroepidemiology 2006;27(1):39-44.
158. El-Salem K, Al-Shimmery E, Horany K, Al-Refai A, Al-Hayk K, Khader Y. Multiple sclerosis in Jordan: A clinical and epidemiological study. J Neurol 2006;253(9):1210-6.
159. Turk Boru U, Alp R, Sur H, Gul L. Prevalence of multiple sclerosis door-to-door survey in Maltepe, Istanbul, Turkey. Neuroepidemiology 2006;27(1):17-21.
160. Alter M, Kahana E, Zilber N, Miller A. Multiple sclerosis frequency in Israel's diverse populations. Neurology 2006;66(7):1061-6.
161. Holland NJ. Overview of Multiple Sclerosis: National MS Society; 2008.
162. Whetten-Goldstein K, Sloan FA, Goldstein LB, Kulas ED. A comprehensive assessment of the cost of multiple sclerosis in the United States. Mult Scler 1998;4(5):419-25.
163. Peltonen L, Saarela J, Kuokkanen S. Multiple sclerosis. In: King RA, Rotter JI, Motulsky AG, editors. The Genetic Basis of Common Diseases. 2nd ed. New York: Oxford University Press, Inc.; 2002. p. 805-817.
164. Hauser SL, Goodin DS. Multiple sclerosis and other demyelinating diseases. In: Kasper DL, Braunwald E, Fauci AD, Hauser SL, Longo DL, Jameson JL, editors. Harrison's Principles of Internal Medicine. 16 ed. NY: McGraw Hill; 2005.
165. Cottrell DA, Kremenchutzky M, Rice GP, Koopman WJ, Hader W, Baskerville J, et al. The natural history of multiple sclerosis: a geographically based study. 5. The clinical features and natural history of primary progressive multiple sclerosis. Brain 1999;122 (Pt 4):625-39.
166. Mehta BK. New hypotheses on sunlight and the geographic variability of multiple sclerosis prevalence. J Neurol Sci;292(1-2):5-10.
167. Dalmay F, Bhalla D, Nicoletti A, Cabrera-Gomez J, Cabre P, Ruiz F, et al. Multiple sclerosis and solar exposure before the age of 15 years: case-control study in Cuba, Martinique and Sicily. Mult Scler.
168. Ji Q, Perchellet A, Goverman JM. Viral infection triggers central nervous system autoimmunity via activation of CD8(+) T cells expressing dual TCRs. Nat Immunol.
169. Levin LI, Munger KL, O'Reilly EJ, Falk KI, Ascherio A. Primary infection with the Epstein-Barr virus and risk of multiple sclerosis. Ann Neurol;67(6):824-30.
170. Eisenbarth GS. Type I diabetes mellitus. A chronic autoimmune disease. N Engl J Med 1986;314(21):1360-8.
171. Ostrauskas R. The prevalence of type 1 diabetes mellitus among adolescents and adults in Lithuania during 1991-2004. Medicina (Kaunas) 2007;43(3):242-50.
172. Forouhi NG, Merrick D, Goyder E, Ferguson BA, Abbas J, Lachowycz K, et al. Diabetes prevalence in England, 2001--estimates from an epidemiological model. Diabet Med 2006;23(2):189-97.

173. Jansson SP, Andersson DK, Svardsudd K. Prevalence and incidence rate of diabetes mellitus in a Swedish community during 30 years of follow-up. Diabetologia 2007;50(4):703-10.
174. Knox SA, Harrison CM, Britt HC, Henderson JV. Estimating prevalence of common chronic morbidities in Australia. Med J Aust 2008;189(2):66-70.
175. Bahillo MP, Hermoso F, Ochoa C, Garcia-Fernandez JA, Rodrigo J, Marugan JM, et al. Incidence and prevalence of type 1 diabetes in children aged <15 yr in Castilla-Leon (Spain). Pediatr Diabetes 2007;8(6):369-73.
176. Rosenbauer J, Icks A, Giani G. Incidence and prevalence of childhood type 1 diabetes mellitus in Germany--model-based national estimates. J Pediatr Endocrinol Metab 2002;15(9):1497-504.
177. Duncan GE. Prevalence of diabetes and impaired fasting glucose levels among US adolescents: National Health and Nutrition Examination Survey, 1999-2002. Arch Pediatr Adolesc Med 2006;160(5):523-8.
178. Wu D, Kendall D, Lunt H, Willis J, Darlow B, Frampton C. Prevalence of Type 1 diabetes in New Zealanders aged 0-24 years. N Z Med J 2005;118(1218):U1557.
179. Moore KR, Harwell TS, McDowall JM, Helgerson SD, Gohdes D. Three-year prevalence and incidence of diabetes among American Indian youth in Montana and Wyoming, 1999 to 2001. J Pediatr 2003;143(3):368-71.
180. Peter SA, Johnson R, Taylor C, Hanna A, Roberts P, McNeil P, et al. The incidence and prevalence of type-1 diabetes mellitus. J Natl Med Assoc 2005;97(2):250-2.
181. Al-Herbish AS, El-Mouzan MI, Al-Salloum AA, Al-Qurachi MM, Al-Omar AA. Prevalence of type 1 diabetes mellitus in Saudi Arabian children and adolescents. Saudi Med J 2008;29(9):1285-8.
182. Moussa MA, Alsaeid M, Abdella N, Refai TM, Al-Sheikh N, Gomez JE. Prevalence of type 1 diabetes among 6- to 18-year-old Kuwaiti children. Med Princ Pract 2005;14(2):87-91.
183. Incidence and trends of childhood Type 1 diabetes worldwide 1990-1999. Diabet Med 2006;23(8):857-866.
184. Kostraba JN, Gay EC, Cai Y, Cruickshanks KJ, Rewers MJ, Klingensmith GJ, et al. Incidence of insulin-dependent diabetes mellitus in Colorado. Epidemiology 1992;3(3):232-8.
185. Rangasami JJ, Greenwood DC, McSporran B, Smail PJ, Patterson CC, Waugh NR. Rising incidence of type 1 diabetes in Scottish children, 1984-93. The Scottish Study Group for the Care of Young Diabetics. Arch Dis Child 1997;77(3):210-3.
186. McCann JA, Xu YQ, Frechette R, Guazzarotti L, Polychronakos C. The Insulin-Like Growth Factor-II Receptor Gene Is Associated with Type 1 Diabetes: Evidence of a Maternal Effect. J Clin Endocrinol Metab 2004;89(11):5700-5706.
187. Lorenzen T, Pociot F, Hougaard P, Nerup J. Long-term risk of IDDM in first-degree relatives of patients with IDDM. Diabetologia 1994;37(3):321-7.
188. Jahromi MM, Eisenbarth GS. Genetic Determinants of Type 1 Diabetes Across Populations. Ann NY Acad Sci 2006;1079(1):289-299.
189. Noble JA, Valdes AM, Cook M, Klitz W, Thomson G, Erlich HA. The role of HLA class II genes in insulin-dependent diabetes mellitus: molecular analysis of 180 Caucasian, multiplex families. Am J Hum Genet 1996;59(5):1134-48.
190. Rotter JI, Anderson CE, Rubin R, Congleton JE, Terasaki PI, Rimoin DL. HLA Genotypic Study of Insulin-Dependent Diabetes - the Excess of Dr3/Dr4 Heterozygotes Allows Rejection of the Recessive Hypothesis. Diabetes 1983;32(2):169-174.

191. Thomson G, Robinson WP, Kuhner MK, Joe S, Macdonald MJ, Gottschall JL, et al. Genetic-Heterogeneity, Modes of Inheritance, and Risk Estimates for a Joint Study of Caucasians with Insulin-Dependent Diabetes-Mellitus. Am J Hum Genet 1988;43(6):799-816.
192. Karvonen M, Tuomilehto J, Libman I, Laporte R. A Review of the Recent Epidemiologic Data on the Worldwide Incidence of Type-1 (Insulin-Dependent) Diabetes-Mellitus. Diabetologia 1993;36(10):883-892.
193. Valdes AM, Erlich HA, Noble JA. Human leukocyte antigen class I B and C loci contribute to Type 1 Diabetes (T1D) susceptibility and age at T1D onset. Hum Immunol 2005;66(3):301-313.
194. Nejentsev S, Howson JMM, Walker NM, Szeszko J, Field SF, Stevens HE, et al. Localization of type 1 diabetes susceptibility to the MHC class I genes HLA-B and HLA-A. Nature 2007;450(7171):887-892.
195. Pugliese A. Genetic protection from insulin-dependent diabetes mellitus. Diabetes Nutr Metab 1997;10(4):169-179.
196. Bottini N, Musumeci L, Alonso A, Rahmouni S, Nika K, Rostamkhani M, et al. A functional variant of lymphoid tyrosine phosphatase is associated with type I diabetes. Nat Genet 2004;36(4):337-8.
197. Nistico L, Buzzetti R, Pritchard LE, Van der Auwera B, Giovannini C, Bosi E, et al. The CTLA-4 gene region of chromosome 2q33 is linked to, and associated with, type 1 diabetes. Belgian Diabetes Registry. Hum Mol Genet 1996;5(7):1075-80.
198. Bennett ST, Lucassen AM, Gough SC, Powell EE, Undlien DE, Pritchard LE, et al. Susceptibility to human type 1 diabetes at IDDM2 is determined by tandem repeat variation at the insulin gene minisatellite locus. Nat Genet 1995;9(3):284-92.
199. Adeghate E, Schattner P, Dunn E. An update on the etiology and epidemiology of diabetes mellitus. Ann NY Acad Sci 2006;1084:1-29.
200. Ziegler AG, Schmid S, Huber D, Hummel M, Bonifacio E. Early infant feeding and risk of developing type 1 diabetes-associated autoantibodies. Jama 2003;290(13):1721-8.
201. Norris JM, Barriga K, Klingensmith G, Hoffman M, Eisenbarth GS, Erlich HA, et al. Timing of initial cereal exposure in infancy and risk of islet autoimmunity. Jama 2003;290(13):1713-20.
202. Roivainen M, Knip M, Hyoty H, Kulmala P, Hiltunen M, Vahasalo P, et al. Several different enterovirus serotypes can be associated with prediabetic autoimmune episodes and onset of overt IDDM. Childhood Diabetes in Finland (DiMe) Study Group. J Med Virol 1998;56(1):74-8.
203. Virtanen SM, Rasanen L, Aro A, Lindstrom J, Sippola H, Lounamaa R, et al. Infant feeding in Finnish children less than 7 yr of age with newly diagnosed IDDM. Childhood Diabetes in Finland Study Group. Diabetes Care 1991;14(5):415-7.
204. Virtanen SM, Rasanen L, Ylonen K, Aro A, Clayton D, Langholz B, et al. Early introduction of dairy products associated with increased risk of IDDM in Finnish children. The Childhood in Diabetes in Finland Study Group. Diabetes 1993;42(12):1786-90.
205. Virtanen SM, Saukkonen T, Savilahti E, Ylonen K, Rasanen L, Aro A, et al. Diet, cow's milk protein antibodies and the risk of IDDM in Finnish children. Childhood Diabetes in Finland Study Group. Diabetologia 1994;37(4):381-7.
206. Verge CF, Howard NJ, Irwig L, Simpson JM, Mackerras D, Silink M. Environmental factors in childhood IDDM. A population-based, case-control study. Diabetes Care 1994;17(12):1381-9.

207. Chang CH, Flavell RA. Class II transactivator regulates the expression of multiple genes involved in antigen presentation. J Exp Med 1995;181(2):765-767.
208. Chang CH, Guerder S, Hong SC, van Ewijk W, Flavell RA. Mice lacking the MHC class II transactivator (CIITA) show tissue-specific impairment of MHC class II expression. Immunity 1996;4(2):167-178.
209. Kern I, Steimle V, Siegrist CA, Mach B. The two novel MHC class II transactivators RFX5 and CIITA both control expression of HLA-DM genes. Int Immunol 1995;7(8):1295-1299.
210. LeibundGut-Landmann S, Waldburger JM, Krawczyk M, Otten LA, Suter T, Fontana A, et al. Specificity and expression of CIITA, the master regulator of MHC class II genes. Eur J Immunol 2004;34(6):1513-1525.
211. Steimle V, Otten LA, Zufferey M, Mach B. Complementation cloning of an MHC class II transactivator mutated in hereditary MHC class II deficiency (or bare lymphocyte syndrome). Cell 1993;75(1):135-46.
212. Heldt C, Listing J, Sozeri O, Blasing F, Frischbutter S, Muller B. Differential expression of HLA class II genes associated with disease susceptibility and progression in rheumatoid arthritis. Arthritis Rheum 2003;48(10):2779-87.
213. Etzel CJ, Chen WV, Shepard N, Jawaheer D, Cornelis F, Seldin MF, et al. Genome-wide meta-analysis for rheumatoid arthritis. Hum Genet 2006;119(6):634-41.
214. John S, Amos C, Shephard N, Chen W, Butterworth A, Etzel C, et al. Linkage analysis of rheumatoid arthritis in US and UK families reveals interactions between HLA-DRB1 and loci on chromosomes 6q and 16p. Arthritis and Rheumatism 2006;54(5):1482-1490.
215. Swanberg M, Lidman O, Padyukov L, Eriksson P, Akesson E, Jagodic M, et al. MHC2TA is associated with differential MHC molecule expression and susceptibility to rheumatoid arthritis, multiple sclerosis and myocardial infarction. Nat Genet 2005;37(5):486-494.
216. Todd JA, Walker NM, Cooper JD, Smyth DJ, Downes K, Plagnol V, et al. Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes. Nat Genet 2007;39(7):857-864.
217. Bates EE, Fournier N, Garcia E, Valladeau J, Durand I, Pin JJ, et al. APCs express DCIR, a novel C-type lectin surface receptor containing an immunoreceptor tyrosine-based inhibitory motif. J Immunol 1999;163(4):1973-83.
218. Geijtenbeek TB, van Vliet SJ, Engering A, t Hart BA, van Kooyk Y. Self- and nonself-recognition by C-type lectins on dendritic cells. Annu Rev Immunol 2004;22:33-54.
219. Hakonarson H, Grant SFA, Bradfield JP, Marchand L, Kim CE, Glessner JT, et al. A genome-wide association study identifies KIAA0350 as a type 1 diabetes gene. Nature 2007;448(7153):591-594.
220. Hafler DA, Compston A, Sawcer S, Lander ES, Daly MJ, De Jager PL, et al. Risk alleles for multiple sclerosis identified by a genomewide study. N Engl J Med 2007;357(9):851-62.
221. Zoledziewska M, Costa G, Pitzalis M, Cocco E, Melis C, Moi L, et al. Variation within the CLEC16A gene shows consistent disease association with both multiple sclerosis and type 1 diabetes in Sardinia. Genes Immun 2009;10(1):15-7.
222. Martinez A, Perdigones N, Cenit MC, Espino L, Varade J, Lamas JR, et al. Chromosomal region 16p13: further evidence of increased predisposition to immune diseases. Ann Rheum Dis 2010;69(1):309-11.

223. Barrett JC, Clayton DG, Concannon P, Akolkar B, Cooper JD, Erlich HA, et al. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. Nat Genet 2009;41(6):703-707.
224. Rubio JP, Stankovich J, Field J, Tubridy N, Marriott M, Chapman C, et al. Replication of KIAA0350, IL2RA, RPL5 and CD58 as multiple sclerosis susceptibility genes in Australians. Genes Immun 2008;9(7):624-30.
225. Nelson JL, Gillespie KM, Lambert NC, Stevens AM, Loubiere LS, Rutledge JC, et al. Maternal microchimerism in peripheral blood in type 1 diabetes and pancreatic islet beta cell microchimerism. Proc Natl Acad Sci USA 2007;104(5):1637-42.
226. Stevens AM, Tsao BP, Hahn BH, Guthrie K, Lambert NC, Porter AJ, et al. Maternal HLA class II compatibility in men with systemic lupus erythematosus. Arthritis Rheum 2005;52(9):2768-73.
227. Lambert NC, Erickson TD, Yan Z, Pang JM, Guthrie KA, Furst DE, et al. Quantification of maternal microchimerism by HLA-specific real-time polymerase chain reaction: studies of healthy women and women with scleroderma. Arthritis Rheum 2004;50(3):906-14.
228. Lo YMD, Lau TK, Chan LYS, Leung TN, Chang AMZ. Quantitative Analysis of the Bidirectional Fetomaternal Transfer of Nucleated Cells and Plasma DNA. Clin Chem 2000;46(9):1301-1309.
229. Lambert NC, Stevens AM, Tylee TS, Erickson TD, Furst DE, Nelson JL. From the simple detection of microchimerism in patients with autoimmune diseases to its implication in pathogenesis. Ann NY Acad Sci 2001;945:164-71.
230. Nelson JL. Microchimerism and human autoimmune diseases. Lupus 2002;11(10):651-654.
231. Nelson JJL. Microchimerism and HLA relationships of pregnancy: implications for autoimmune diseases. Curr Rheumatol Rep 2001;3(3):222-229.
232. Holzgreve W, Hahn S, Zhong XY, Lapaire O, Hosli I, Tercanli S, et al. Genetic communication between fetus and mother: short- and long-term consequences. Am J Obstet Gynecol 2007;196(4):372-381.
233. Artlett CM, Welsh KI, Black CM, Jimenez SA. Fetal-maternal HLA compatibility confers susceptibility to systemic sclerosis. Immunogenetics 1997;47(1):17-22.
234. Lunetta KL, Hayward LB, Segal J, Van Eerdewegh P. Screening large-scale association study data: exploiting interactions using random forests. Bmc Genetics 2004;5:13.
235. Rak JM, Maestroni L, Balandraud N, Guis S, Boudinet H, Guzian MC, et al. Transfer of the shared epitope through microchimerism in women with rheumatoid arthritis. Arthritis Rheum 2009;60(1):73-80.
236. Robertson KD, Wolffe AP. DNA methylation in health and disease. Nat Rev Genet 2000;1(1):11-9.
237. Marrosu MG, Sardu C, Cocco E, Costa G, Murru MR, Mancosu C, et al. Bias in parental transmission of the HLA-DR3 allele in Sardinian multiple sclerosis. Neurology 2004;63(6):1084-1086.
238. Ramagopalan SV, Herrera BM, Bell JT, Dyment DA, DeLuca GC, Lincoln MR, et al. Parental transmission of HLA-DRB1*15 in multiple sclerosis. Hum Genet 2008;122(6):661-663.
239. Vadheim CM, Rotter JI, Maclaren NK, Riley WJ, Anderson CE. Preferential transmission of diabetic alleles within the HLA gene complex. N Engl J Med 1986;315(21):1314-8.

240. Sasaki T, Nemoto M, Yamasaki K, Tajima N. Preferential transmission of maternal allele with DQA1*0301-DQB1*0302 haplotype to affected offspring in families with type 1 diabetes. J Hum Genet 1999;44(5):318-322.
241. Martin-Villa JM, Vicario JL, Martinez-Laso J, Serrano-Rios M, Lledo G, Damiano A, et al. Lack of preferential transmission of diabetic HLA alleles by healthy parents to offspring in Spanish diabetic families. J Clin Endocrinol Metab 1990;70(2):346-8.
242. Bain SC, Rowe BR, Barnett AH, Todd JA. Parental origin of diabetes-associated HLA types in sibling pairs with type I diabetes. Diabetes 1994;43(12):1462-8.
243. Undlien DE, Akselsen HE, Joner G, Dahl-Jorgensen K, Aagenaes O, Sovik O, et al. No difference in the parental origin of susceptibility HLA class II haplotypes among Norwegian patients with insulin-dependent diabetes mellitus. Am J Hum Genet 1995;57(6):1511-4.
244. Bianchi DDW, Zickwolf GGK, Weil GGJ, Sylvester SS, DeMaria MMA. Male fetal progenitor cells persist in maternal blood for as long as 27 years postpartum. Proc Natl Acad Sci USA 1996;93(2):705-708.
245. Hall JM, Lingenfelter P, Adams SL, Lasser D, Hansen JA, Bean MA. Detection of maternal cells in human umbilical cord blood using fluorescence in situ hybridization. Blood 1995;86(7):2829-32.
246. van Rood JJ, Claas F. Both self and non-inherited maternal HLA antigens influence the immune response. Immunol Today 2000;21(6):269-73.
247. van Rood JJ, Claas F. Noninherited maternal HLA antigens: a proposal to elucidate their role in the immune response. Hum Immunol 2000;61(12):1390-4.
248. Stevens AM. Microchimeric cells in systemic lupus erythematosus: targets or innocent bystanders? Lupus 2006;15(11):820-6.
249. Burlingham WJ, Grailer AP, Heisey DM, Claas FH, Norman D, Mohanakumar T, et al. The effect of tolerance to noninherited maternal HLA antigens on the survival of renal transplants from sibling donors. N Engl J Med 1998;339(23):1657-64.
250. Campbell DA, Jr., Lorber MI, Sweeton JC, Turcotte JG, Niederhuber JE, Beer AE. Breast feeding and maternal-donor renal allografts. Possibly the original donor-specific transfusion. Transplantation 1984;37(4):340-4.
251. Kois WE, Campbell DA, Lorber MI, Sweeton JC, Dafoe DC. Influence of Breast-Feeding on Subsequent Reactivity to a Related Renal-Allograft. J Surg Res 1984;37(2):89-93.
252. Andrassy J, Kusaka S, Jankowska-Gan E, Torrealba JR, Haynes LD, Marthaler BR, et al. Tolerance to noninherited maternal MHC antigens in mice. J Immunol 2003;171(10):5554-61.
253. Adams KM, Nelson JL. Microchimerism: an investigative frontier in autoimmunity and transplantation. JAMA-J Am Med Assoc 2004;291(9):1127-31.
254. Harney S, Newton J, Milicic A, Brown MA, Wordsworth BP. Non-inherited maternal HLA alleles are associated with rheumatoid arthritis. Rheumatology 2003;42(1):171-4.
255. Guthrie KA, Tishkevich NR, Nelson JL. Non-inherited maternal human leukocyte antigen alleles in susceptibility to familial rheumatoid arthritis. Ann Rheum Dis 2009;68(1):107-9.
256. Feitsma AL, Worthington J, van der Helm-van Mil AHM, Plant D, Thomson W, Ursum J, et al. Protective effect of noninherited maternal HLA-DR antigens on rheumatoid arthritis development. Proc Natl Acad Sci USA 2007;104(50):19966-19970.

Table 1. Descriptive summary of autoimmune diseases examined in this genetic study.

	RA	**SLE**	**MS**	**T1D**
Pop. Risk (%)	1	0.05	0.1	0.4
Sibling Risk	8	1.5	2	6
Familial clustering (λ_s)	8	30	20	15
Monozygotic Twin Concordance	15-30	24-69	25-30	35-40
Dizygotic Twin Concordance	2-4	2-9	3-5	5-6

CHAPTER TWO

The *CIITA* -168A/G Polymorphism and Risk for Rheumatoid Arthritis: A Meta-Analysis of 6,861 Patients and 9,270 Controls Reveals No Evidence for Association

Published: Bronson PG, Criswell LA, Barcellos LF. (2008) The MHC2TA -168A/G polymorphism and risk for rheumatoid arthritis: a meta-analysis of 6,861 patients and 9,270 controls reveals no evidence for association. Ann Rheum Dis 67(7): 933-6.

ABSTRACT

An association between major histocompatibility complex (MHC) genes, particularly those within the class II HLA region, and rheumatoid arthritis (RA) is well established, and accounts for an estimated 30% of the genetic component in RA. The MHC class II transactivator gene (*CIITA*) on chromosome 16p13 has recently emerged as one of the most important transcription factor regulating genes required for class II MHC-restricted antigen presentation. Previous studies of a promoter region polymorphism (-168A/G, rs3087456) in the *CIITA* gene and RA have yielded conflicting results. Our objective was to assess the association of the *CIITA* -168A/G polymorphism (rs3087456) and risk for RA by meta-analysis. Meta-analysis was performed for 6,861 RA patients and 9,270 controls from ten case-control studies. Odds ratios (ORs) and 95% confidence intervals (CIs) were calculated for each study. Summary ORs and 95% CIs were calculated for random effects models. No effect was observed for the G risk allele (OR = 1.02, 95% CI = 0.93 to 1.12, $P = 0.70$) or the GG risk genotype (OR = 1.14, 95% CI = 0.95 to 1.36, $P = 0.16$). Our results indicate that the *CIITA* -168A/G polymorphism (rs3087456) is not associated with RA yet underscore the importance of including shared epitope alleles, secondary phenotypes and more complete characterization of *CIITA* variation in future studies.

BACKGROUND

Rheumatoid arthritis (RA) is the most common systemic autoimmune disease with a worldwide prevalence approaching one percent (1, 2). It is a chronic inflammatory disease with the potential to cause substantial disability, primarily as a result of the erosive and deforming processes in joints, and is associated with increased mortality, particularly among individuals who develop extra-articular manifestations (3, 4). The prevalence of RA, in conjunction with the impact on the health status of affected individuals, results in tremendous associated costs to patients, their families and society (5, 6). While the etiology of RA is unknown, a significant genetic contribution to RA development is widely accepted and estimated to account for approximately 60 percent of the disease risk (7). A strong association between major histocompatibility complex (MHC) genes, particular those in the class II HLA region, and RA (8) has been well established, and is estimated to account for 30 to 50 percent of the genetic component in RA, at least in Northern Europeans (7). Thus, other genes (outside the MHC region) also contribute to disease risk.

The MHC class II transactivator (*CIITA*, previously called *MHC2TA*), located on chromosome 16p13, is a global regulator of expression of proteins involved in antigen presentation and processing including MHC class II molecules (9-12). Thus, *CIITA* is an attractive candidate for genetic studies of autoimmune diseases and other inflammatory conditions for which HLA associations have been well established. The CIITA protein functions as a chaperone for assembly of several transcription factors at MHC promoters (13). In addition, it may play a modulatory role in regulating the expression of other genes including some phosphatases, kinases and other genes involved in cell signaling. Due to the prominent role of the CIITA protein in immunoregulation and initial reports of association with RA, the majority of *CIITA* studies have focused on the rs3087456 (-168A/G) polymorphism located in the type III promoter, one of four promoters, all of which vary according to cell type (13, 14). The rs3087456 polymorphism has demonstrated biological relevance through functional studies in animals and humans (14) and has been shown to determine expression in antigen-presenting cells, including both B and activated T lymphocytes. To date, studies of *CIITA* variation and risk for RA have demonstrated conflicting results (Tables 1 and 2). The meta-analyses presented here utilize all available RA cases and controls genotyped for this putative causal *CIITA* polymorphism. The promoter polymorphism has been described as an A/G change in the literature but it is actually a G/A change because the G allele is ancestral (15). In this study, a meta-analysis was performed to test whether the G allele or the GG genotype are associated with susceptibility to RA.

METHODS

Identification of studies

A literature search of the PubMed database developed by the National Center for Biotechnology Information (NCBI) at the National Library of Medicine (NLM) was conducted using “rheumatoid arthritis”, “*CIITA*” and “*MHC2TA*” as search terms. Nine publications investigating the effect of the rs3084756 polymorphism and RA risk were identified (14, 16-22). Additional computerized searches of the Cochrane Library, Web of Science and Google Scholar databases did not identify any additional relevant publications. The search was not limited to

English language articles, but all publications were written in English. Seven publications were original research articles and two were letters to the editor (19, 21). Eight of the nine total publications presented results for a single study; one letter to the editor presented results for three studies (21). All publications prior to January 12, 2007 were identified for the meta-analysis.

Study selection criteria

The articles identified for this analysis were included if they met the following criteria: (1) All patients had to fulfill well-established classification criteria (1987 American College of Rheumatology) for RA (23); juvenile RA was not included (16); (2) Control individuals, if listed, had to be derived from a population within the same geographic area and ethnic background as patients; (3) Authors had to provide original *CIITA* rs3087456 genotype distributions; and (4) RA patient and control groups had to be unique for inclusion in the meta-analysis (it had to be evident that there was no overlap among studies); and (5) All languages, geographic areas and publication types were included. Two reviewers (P.G.B. and L.F.B.) applied inclusion/exclusion criteria. Ten studies of *CIITA* and RA met the inclusion criteria; the most recent study did not meet our inclusion criteria because original *CIITA* rs3087456 genotype distributions were not provided in the publication (24).

Definition of CIITA genotypes for meta-analysis

The minor G allele for rs3087456 was studied, either for G allele carriers under a dominant disease model (AA vs. GG+GA; 0 G alleles vs. 1 or 2 G alleles), or for G homozygotes under a recessive disease model (AA+AG vs. GG; 0 or 1 G alleles vs. 2 G alleles). The initial study of *CIITA* and RA demonstrated that carriers of the G allele were at increased risk for RA, and that the homozygous G genotype influenced expression of the CIITA protein and MHC class II molecules (14); therefore this study tested hypotheses specifically related to carrier status of the rs3087456G allele.

Methods for data extraction

One author (P.G.B.) extracted the following data from each study: authors; publication year; journal; publication type; place of study; study design; genotyping method; rs3087456 genotype frequencies; number of cases and controls; and a description of the sample, including recruitment method, age, race and ethnicity. Cases were also described by RA classification criteria, age at onset, and additional RA phenotypic data, if it was available. Controls were also described as population-based or matched controls and by matching characteristics. All ten studies had a case-control study design. Two of the studies presented genotype frequencies for a set of matched controls and a larger set of population-based controls (14, 22). Results were reported using the population-based controls because a larger sample size provided more power. Genotype frequencies were similar between the two sets of matched and population controls.

Methods of statistical analysis

Deviation from Hardy-Weinberg equilibrium (HWE) was tested for *CIITA* genotypes in control datasets from each individual study using the exact test implemented in Python for Population Genetics (PyPop) software version 0.6.0 (25). The meta-analysis included ten case-control studies with 6,861 RA patients and 9,270 controls. For each individual study, the odds ratios (OR) and 95% confidence intervals (95% CI) for the G risk allele and the GG risk genotype in RA patients against controls were calculated. Statistical analyses were performed in

R (http://www.r-project.org). ORs were calculated using unconditional maximum likelihood estimation. 95% CIs were calculated using normal approximation. Two-tailed *P*-values were calculated using the χ^2 test of independence. The summary ORs and 95% CIs for the G risk allele and the GG risk genotype in RA patients against controls were calculated for random effects models because it accounts for study-to-study heterogeneity in the association of interest, and thus does not assume that all populations used have the same underlying association (26). Between-study heterogeneity was assessed with the Cochran's Q test-statistic. Forest plots were generated in R to display the OR and 95% CI for each study (shown by a black square on a horizontal line, respectively) and the summary OR and 95% CI (shown by a dotted vertical line extending from a black diamond, respectively).

Publication bias was evaluated by constructing funnel plots and applying Egger's linear regression test (27). Funnel plot asymmetry was tested with the "metabias" command using the "linreg" method, which calculates a test statistic based on a weighted linear regression of the treatment effect on its standard error (28). The test statistic follows a t distribution with df equal to the number of studies minus two. A cumulative meta-analysis was conducted for G homozygotes under the fixed and random effects models to evaluate the cumulative summary OR over time. An influential analysis was conducted for G homozygotes under the fixed and random effects models to evaluate the influence of any one study on the summary OR. Studies were omitted one at a time. Omitting the hypothesis-generating study (14) yielded nine studies with 5,599 RA patients and 7,671 controls.

Three subgroup analyses were performed according to country of origin (Europe, Latin America and Asia), ethnicity (European and Asian), and SNP genotyping method. A logistic regression analysis of pooled genotype data from each study was also performed in SAS (PROC LOGISTIC) for G allele carriers and G homozygotes (SAS v. 9.1, Cary, NC). Resulting ORs and 95% CIs were adjusted for study location. Haploview (v. 3.32; www.broad.mit.edu/mpg/haploview) was used to identify haplotype tagging SNPs in the *CIITA* locus, using publicly available genotype data for 64 *CIITA* SNPs from the International Haplotype Website (www.hapmap.org) for 30 trio families sampled from Utah residents with ancestry from and Europe (CEU).

RESULTS

The ten studies used in the meta-analysis of *CIITA* rs3087456 are summarized in Table 1. There was no evidence for deviation from HWE in control individuals derived from any individual study (data not shown). Study results for the ten studies used in the meta-analysis are summarized in Table 2. Studies were weighted by size to contribute to the overall combined result. There was evidence of between-study heterogeneity for the G risk allele (χ^2 =17.7, df = 9, P = 0.04) but not for the GG risk genotype (χ^2 = 14.7, df = 9, P = 0.10). There was no evidence of association for the *CIITA* -168A/G polymorphism (rs3087456) with RA risk when comparing GG+AG against AA carriers (OR = 1.02, 95% CI = 0.93 to 1.12, P = 0.70) or when comparing GG against AG+AA carriers (OR = 1.14, 95% CI = 0.95 to 1.36, P = 0.16). The forest plots in Figure 1 show a schematic representation of the data that are presented in the bottom line of Table 2.

The subgroup analyses based upon country of origin, ethnicity and SNP genotyping method did not yield different results (data not shown). Results did not differ when the meta-analysis was limited to studies on populations of European descent or to studies of European populations (data not shown). Further, the logistic regression analysis of pooled genotype data, adjusted for study, also indicated that neither *CIITA* rs3087456 G allele carriers nor homozygotes were at an increased risk of RA (data not shown).

DISCUSSION

There is substantial evidence for genetic contribution of the MHC to RA. The MHC spans ~4.5 mega base pairs (Mb) on chromosome 6p21.3 and encodes >180 expressed genes; 40 percent are related to immune activation and response (29, 30). The MHC class II gene *HLA-DRB1* demonstrates the strongest association with RA, highlighting antigen presentation and subsequent T cell activation as a potential pathway in RA pathogenesis (8). All RA-associated *DRB1* alleles encode a shared epitope (SE) not present on non-RA associated alleles (31). Expression of MHC class II genes appears to be regulated almost exclusively by *CIITA* located on chromosome 16p13. Previous studies of a promoter region polymorphism (-168A/G, rs3087456) in *CIITA* and susceptibility to RA have demonstrated conflicting results. This study assessed the association of *CIITA* rs3087456 (-168A/G) and RA risk by meta-analysis of 6,861 patients and 9,270 controls, and is the largest investigation, to date, of *CIITA* variation and RA. Our findings indicate that *CIITA* rs3087456 G allele and GG genotype carrier status do not substantially increase risk for RA.

Information bias could have resulted from genotype misclassification. Most studies described laboratory quality control methods and specified the source of DNA as blood. Although genotyping methods varied between studies, there was no evidence of deviation from HWE in the controls, providing some level of support for a very low genotype misclassification rate. In addition, misdiagnosis bias is unlikely to be present in RA patients from different studies because the classification criteria were clearly defined in the inclusion criteria; all published studies met these criteria. Furthermore, while the potential for publication bias always exists in a meta-analysis due to the greater number of studies that are published with evidence for significant effects, our funnel plot included studies in the lower left section, indicating that small studies with negative results have been published (data not shown). In addition, the linear regression test of funnel plot asymmetry did not provide evidence for publication bias (data not shown).

Selection bias could also have been present in control individuals. While cases and controls from studies included in this meta-analysis were generally well-matched by ethnicity and geographic location, it is possible that population structure and resulting stratification could have influenced results (32), particularly in light of recent findings showing selected SNP ancestry informative markers (AIMs) can even distinguish genetic differences between closely related populations such as Northern and Southern Europeans (33, 34). Spurious associations may result from studies in populations with substructure if subgroups are not equally represented in both cases and controls. Studies of genetic associations in admixed populations, for example,

Argentineans, who are comprised of individuals with European, American Indian, and African ancestry, are particularly prone to stratification effects due to population structure (35). Matching as closely as possible should reduce or eliminate confounding due to population stratification; however, this is not always possible based on self-report, particularly in admixed populations where individuals may not know their precise ancestral background (36). The extent to which population substructure contributes to confounding in case-control studies remains controversial (38-40), and whether confounding was present, specifically, in any of the *CIITA* studies presented here is not known. Similar to other autoimmune-disease associated SNPs (37), ethnic differences in *CIITA* rs3087456 (-168 A/G) allele frequencies do exist. The use of a family-based study design (38) or the incorporation of AIMs in a case-control study design to identify structure and control for confounding (39) are needed to further evaluate the role of *CIITA* variation in RA.

Whole-genome microsatellite linkage screens in multiple affected RA families have been performed to identify non-MHC RA genes. Several studies have indicated linkage to microsatellites in chromosome 16p13 (40-42). A recent genome search meta-analysis (GSMA) of four linkage studies (41, 43-45) reported that in addition to the HLA locus ($P < 2x10^{-5}$), the highest linkage peak ($P = 0.004$) was on chromosome 16p13 (46). GSMA and pooled linkage analysis results obtained from different combinations of previous RA linkage studies (41-45, 47) have provided very similar results (47-49). Interestingly, conditioning for allele sharing at *DRB1* through two-locus joint linkage analysis (47), or stratifying by the presence of two copies of SE alleles (40) revealed stronger evidence for linkage to chromosome 16p (not specifically 16p13), suggesting the possibility of epistatic interaction between *HLA-DRB1* and 16p loci. Only two studies of *CIITA* variation and RA, to date, have accounted for the presence or absence of SE alleles (17, 20) and results so far are conflicting. Because this data was missing from the majority of studies included here, we were not able to incorporate SE carrier status into our analysis. Eyre et al. (17) did not find any effect of the *CIITA* -168A/G polymorphism (rs3087456) on RA susceptibility regardless of SE carrier status. Martinez et al. (20) did not find any effect of the *CIITA* -168A/G polymorphism (rs3087456) on RA susceptibility but reported two *CIITA* haplotypes (comprised of the -168A/G polymorphism (rs3087456) and rs4774) that were associated with RA, with a stronger effect revealed in analyses of SE-positive patients. Differential expression of class II genes has shown some evidence for association with both RA susceptibility and progressive disease in a recent study with direct implications for *CIITA* (50). A complex role in RA for the *CIITA* rs3087456 G allele and/or GG genotype carrier status, in conjunction with *DRB1* SE carrier status, remains plausible and can therefore not be excluded. Larger studies that incorporate both *DRB1* and *CIITA* genotype data are critically needed.

Clinical manifestations for RA are very heterogeneous; this may reflect differences in underlying disease mechanisms based partly on the influence of different genetic factors (51). Rheumatoid factor (RF) seropositivity, anti-cyclic citrullinated peptide (anti-CCP) seropositivity and erosive joint disease, for example, are common secondary phenotypes strongly associated with more severe disease outcomes (52-56) that have demonstrated evidence of specific genetic associations (57-59). In fact, among populations that exhibit an association between RA and the SE, the association is primarily with disease severity or related outcome rather than susceptibility (60-62). Only three published studies of *CIITA* variation and RA examined potential associations with some secondary phenotypes (17, 18, 22), including RF and anti-CCP positivity, age of

onset, presence of erosions, radiologic stage of disease and gender with negative results. Data was not shown in the publications for these phenotypes with the exception of two studies that described RF positivity. Thus, it was not possible to incorporate secondary phenotype data into this meta-analysis. Harrison et al. (18) reported slight trends when comparing the G risk allele in RF negative patients to controls (OR = 0.8, 95% CI = 0.6 to 1.0, P = 0.08) and 559 RF positive patients to 159 RF negative patients (OR = 1.3, 95% CI = 1.0 to 1.7, P = 0.08). Yazdani-Biuki et al. (22) reported no evidence for association between the *CIITA* -168A/G polymorphism (rs3087456) and RF positivity or anti-CCP positivity. *CIITA* rs3087456 G allele and/or GG genotype carrier status may influence specific RA phenotypes. Larger and more comprehensive studies of RA that incorporate both *CIITA* and clinical data are still required to answer this important question (63).

Although results from this meta-analysis of 6,861 RA patients and 9,270 controls do not support a prominent role for the *CIITA* rs3087456 (-168A/G) polymorphism and RA risk, additional functional variant(s) may exist within this locus. Three studies so far have investigated other *CIITA* polymorphisms (14, 17, 20). Eyre and colleagues (17) identified a total of five additional frequency validated SNPs from public databases mapping to the 5' region of the gene (rs7501204, rs6498114, rs6416647, rs7404672, rs6498116) and performed single SNP and extended haplotype analysis in 813 RA patients and 532 controls with negative findings. Swanberg and colleagues (14) evaluated haplotypes comprised of rs3087456 and two additional SNPs (rs4774 and rs2229320plus27bp in exon 11) in 1,288 RA cases and 709 controls; single-locus association results were not significant for these SNPs, and association with rs3087456 was not improved by haplotype analysis. These results are in contrast with findings from a recent haplotype analysis of *CIITA* rs3087456 and rs4774 SNPs in 350 RA patients and 509 controls from Spain (14, 20). The *CIITA* locus spans 47.6 kilo base pairs (kb) and contains 21 exons (http://genome.ucsc.edu). *CIITA* variant information from dbSNP build 126 shows that *CIITA* contains at least 107 validated SNPs. Our analyses of haplotype block structure utilizing publicly available data suggest that at least 21 tagging SNPs are needed to fully capture common *CIITA* variation; seven additional nonsynonymous coding region SNPs not previously evaluated may also be of interest (data not shown).

CONCLUSION

This study assessed association of the *CIITA* -168A/G promoter region polymorphism (rs3087456) and RA risk by performing a meta-analysis of 6,861 patients and 9,270 controls. Given that the large sample size of our study provided sufficient power to detect modest effects, our negative results indicate that the *CIITA* -168A/G polymorphism (rs3087456) is not causally associated with RA. Nevertheless, undiscovered functional variants may exist in the *CIITA*. In conclusion, a complex role for *CIITA* in RA, in conjunction with SE carrier status, remains plausible. Future studies that incorporate SE carrier status, secondary phenotypes and a more comprehensive screening of *CIITA* variation with haplotype tagging SNPs will help elucidate what role, if any, *CIITA* plays in RA.

ACKNOWLEDGEMENTS

This work was supported by the NIH/NIAID grant R01 AI065841.

REFERENCES

1. Gabriel SE. The epidemiology of rheumatoid arthritis. Rheum Dis Clin North Am 2001;27(2):269-81.
2. Silman AJ, Pearson JE. Epidemiology and genetics of rheumatoid arthritis. Arthritis Res 2002;4 Suppl 3:S265-72.
3. Gabriel SE, Crowson CS, Kremers HM, Doran MF, Turesson C, O'Fallon WM, et al. Survival in rheumatoid arthritis: a population-based analysis of trends over 40 years. Arthritis Rheum 2003;48(1):54-8.
4. Turesson C, O'Fallon WM, Crowson CS, Gabriel SE, Matteson EL. Occurrence of extraarticular disease manifestations is associated with excess mortality in a community-based cohort of patients with rheumatoid arthritis. J Rheumatol 2002;29(1):62-7.
5. Lajas C, Abasolo L, Bellajdel B, Hernandez-Garcia C, Carmona L, Vargas E, et al. Costs and predictors of costs in rheumatoid arthritis: a prevalence-based study. Arthritis Rheum 2003;49(1):64-70.
6. Yelin E, Wanke LA. An assessment of the annual and long-term direct costs of rheumatoid arthritis: the impact of poor function and functional decline. Arthritis Rheum 1999;42(6):1209-18.
7. MacGregor AJ, Snieder H, Rigby AS, Koskenvuo M, Kaprio J, Aho K, et al. Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins. Arthritis Rheum 2000;43(1):30-7.
8. Jawaheer D, Li W, Graham RR, Chen W, Damle A, Xiao X, et al. Dissecting the genetic complexity of the association between human leukocyte antigens and rheumatoid arthritis. Am J Hum Genet 2002;71(3):585-94.
9. Chang CH, Flavell RA. Class II transactivator regulates the expression of multiple genes involved in antigen presentation. J Exp Med 1995;181(2):765-767.
10. Chang CH, Guerder S, Hong SC, van Ewijk W, Flavell RA. Mice lacking the MHC class II transactivator (CIITA) show tissue-specific impairment of MHC class II expression. Immunity 1996;4(2):167-178.
11. Kern I, Steimle V, Siegrist CA, Mach B. The two novel MHC class II transactivators RFX5 and CIITA both control expression of HLA-DM genes. Int Immunol 1995;7(8):1295-1299.
12. Steimle V, Otten LA, Zufferey M, Mach B. Complementation cloning of an MHC class II transactivator mutated in hereditary MHC class II deficiency (or bare lymphocyte syndrome). Cell 1993;75(1):135-46.
13. LeibundGut-Landmann S, Waldburger JM, Krawczyk M, Otten LA, Suter T, Fontana A, et al. Specificity and expression of CIITA, the master regulator of MHC class II genes. Eur J Immunol 2004;34(6):1513-1525.
14. Swanberg M, Lidman O, Padyukov L, Eriksson P, Akesson E, Jagodic M, et al. MHC2TA is associated with differential MHC molecule expression and susceptibility to rheumatoid arthritis, multiple sclerosis and myocardial infarction. Nat Genet 2005;37(5):486-494.
15. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, et al. dbSNP: the NCBI database of genetic variation. Nucleic Acids Res 2001;29(1):308-11.
16. Akkad DA, Jagiello P, Szyld P, Goedde R, Wieczorek S, Gross WL, et al. Promoter polymorphism rs3087456 in the MHC class II transactivator gene is not associated with

susceptibility for selected autoimmune diseases in German patient groups. Int J Immunogenet 2006;33(1):59-61.
17. Eyre S, Bowes J, Spreckley K, Potter C, Ring S, Strachan D, et al. Investigation of the MHC2TA gene, associated with rheumatoid arthritis in a Swedish population, in a UK rheumatoid arthritis cohort. Arthritis Rheum 2006;54(11):3417-22.
18. Harrison P, Pointon JJ, Farrar C, Harin A, Wordsworth BP. MHC2TA promoter polymorphism (-168*G/A, rs3087456) is not associated with susceptibility to rheumatoid arthritis in British Caucasian rheumatoid arthritis patients. Rheumatology 2007;46(3):409-411.
19. Iikuni N, Ikari K, Momohara S, Tomatsu T, Hara M, Yamanaka H, et al. MHC2TA is associated with rheumatoid arthritis in Japanese patients. Ann Rheum Dis 2007;66:274-5.
20. Martinez A, Sanchez-Lopez M, Varade J, Mas A, Martin MC, Heras VDL, et al. Role of the MHC2TA gene in autoimmune diseases. Ann Rheum Dis 2007;66(3):325-329.
21. Orozco G, Robledo G, Reddy M, Garcia A, Pascual-Salcedo D, Balsa A, et al. Study of the role of a functional polymorphism of MHC2TA in rheumatoid arthritis in three ethnically different populations. Rheumatology 2006;45(11):1442-1444.
22. Yazdani-Biuki B, Brickmann K, Wohlfahrt K, Mueller T, Marz W, Renner W, et al. The MHC2TA -168A>G gene polymorphism is not associated with rheumatoid arthritis in Austrian patients. Arthritis Research & Therapy 2006;8(4):R97.
23. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988;31(3):315-324.
24. Newman WG, Zhang Q, Liu XD, Walker E, Ternan H, Owen J, et al. Rheumatoid arthritis association with the FCRL3-169C polymorphism is restricted to PTPN22 1858T-homozygous individuals in a Canadian population. Arthritis and Rheumatism 2006;54(12):3820-3827.
25. Lancaster A, Nelson MP, Meyer D, Thomson G, Single RM. PyPop: a software framework for population genomics: analyzing large-scale multi-locus genotype data. Pac Symp Biocomput 2003:514-25.
26. Rothman JR, Greenland, S. Modern Epidemiology. 2 ed. Philadelphia: Lippincott Raven Publishers; 1998.
27. Egger M, Smith GD, Sterne JA. Uses and abuses of meta-analysis. Clin Med 2001;1(6):478-84.
28. Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997;315(7109):629-634.
29. Complete sequence and gene map of a human major histocompatibility complex. The MHC sequencing consortium. Nature 1999;401(6756):921-3.
30. Stewart CA, Horton R, Allcock RJ, Ashurst JL, Atrazhev AM, Coggill P, et al. Complete MHC Haplotype Sequencing for Common Disease Gene Mapping. Genome Res 2004;14(6):1176-87.
31. Gregersen PK, Silver J, Winchester RJ. The shared epitope hypothesis: an approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis. Arthritis Rheum 1987;30:1205-1213.
32. Freedman ML, Reich D, Penney KL, McDonald GJ, Mignault AA, Patterson N, et al. Assessing the impact of population stratification on genetic association studies. Nat Genet 2004;36(4):388-93.
33. Campbell CD, Ogburn EL, Lunetta KL, Lyon HN, Freedman ML, Groop LC, et al.

Demonstrating stratification in a European American population. Nat Genet 2005;37(8):868-72.
34. Seldin MF, Shigeta R, Villoslada P, Selmi C, Tuomilehto J, Silva G, et al. European population substructure: clustering of northern and southern populations. PLoS Genet 2006;2(9):e143.
35. Seldin MF, Tian C, Shigeta R, Scherbarth HR, Silva G, Belmont JW, et al. Argentine population genetic structure: Large variance in Amerindian contribution. Am J Phys Anthropol 2006.
36. Ziv E, Burchard EG. Human population structure and genetic association studies. Pharmacogenomics 2003;4(4):431-41.
37. Mori M, Yamada R, Kobayashi K, Kawaida R, Yamamoto K. Ethnic differences in allele frequency of autoimmune-disease-associated SNPs. J Hum Genet 2005;50(5):264-6.
38. Spielman RS, Ewens WJ. The TDT and other family-based tests for linkage disequilibrium and association. Am J Hum Genet 1996;59:983-989.
39. Choudhry S, Coyle NE, Tang H, Salari K, Lind D, Clark SL, et al. Population stratification confounds genetic association studies among Latinos. Hum Genet 2006;118(5):652-64.
40. Eyre S, Barton A, Shephard N, Hinks A, Brintnell W, MacKay K, et al. Investigation of susceptibility loci identified in the UK rheumatoid arthritis whole-genome scan in a further series of 217 UK affected sibling pairs. Arthritis Rheum 2004;50(3):729-35.
41. Cornelis F, Faure S, Martinez M, Prud'homme J-F, Fritz P, Dib C, et al. New susceptibility locus for rheumatoid arthritis suggested by a genome-wide linkage study. Proc Natl Acad Sci USA 1998;95:10746-50.
42. Jawaheer D, Seldin MF, Amos CI, Chen WV, Shigeta R, Etzel C, et al. Screening the genome for rheumatoid arthritis susceptibility genes: a replication study and combined analysis of 512 multicase families. Arthritis Rheum 2003;48(4):906-16.
43. Jawaheer D, Seldin MF, Amos CI, Chen WV, Monteiro J, Kern M, et al. A genome-wide screen in multiplex rheumatoid arthritis families suggests genetic overlap with other autoimmune diseases. Am J Hum Genet 2001;68:927-936.
44. MacKay K, Eyre S, Myerscough A, Milicic A, Barton A, Laval S, et al. Whole-genome linkage analysis of rheumatoid arthritis susceptibility loci in 252 affected sibling pairs in the United Kingdom. Arthritis Rheum 2002;46(3):632-9.
45. Shiozawa S, Hayashi S, Tsukamoto Y, Goko H, Kawasaki H, Wada T, et al. Identification of the gene loci that predispose to rheumatoid arthritis. Int Immunol 1998;10(12):1891-1895.
46. Fisher SA, Lanchbury JS, Lewis CM. Meta-analysis of four rheumatoid arthritis genome-wide linkage studies - confirmation of a susceptibility locus on chromosome 16. Arthritis and Rheumatism 2003;48(5):1200-1206.
47. John S, Amos C, Shephard N, Chen W, Butterworth A, Etzel C, et al. Linkage analysis of rheumatoid arthritis in US and UK families reveals interactions between HLA-DRB1 and loci on chromosomes 6q and 16p. Arthritis and Rheumatism 2006;54(5):1482-1490.
48. Choi SJ, Rho YH, Ji JD, Song GG, Lee YH. Genome scan meta-analysis of rheumatoid arthritis. Rheumatology 2006;45(2):166-170.
49. Etzel CJ, Chen WV, Shepard N, Jawaheer D, Cornelis F, Seldin MF, et al. Genome-wide meta-analysis for rheumatoid arthritis. Hum Genet 2006;119(6):634-41.
50. Heldt C, Listing J, Sozeri O, Blasing F, Frischbutter S, Muller B. Differential expression of HLA class II genes associated with disease susceptibility and progression in rheumatoid

arthritis. Arthritis Rheum 2003;48(10):2779-87.
51. Jawaheer D, Lum RF, Amos CI, Gregersen PK, Criswell LA. Clustering of disease features within 512 multicase rheumatoid arthritis families. Arthritis Rheum 2004;50(3):736-41.
52. Combe B, Dougados M, Goupille P, Cantagrel A, Eliaou JF, Sibilia J, et al. Prognostic factors for radiographic damage in early rheumatoid arthritis: a multiparameter prospective study. Arthritis Rheum 2001;44(8):1736-43.
53. Suarez-Almazor ME, Soskolne CL, Saunders LD, Russell AS. Outcome in rheumatoid arthritis. A 1985 inception cohort study. J Rheumatol 1994;21(8):1438-46.
54. van Schaardenburg D, Hazes JMW, de Boer A, Zwinderman AH, Meijers KA, Breedveld FC. Outcome of rheumatoid arthritis in relation to age and rheumatoid factor at diagnosis. J Rheumatol 1993;20:45-52.
55. Young A, Bielawska C, Corbett M, Roitt I. A prospective study of early onset rheumatoid arthritis over fifteen years: prognostic features and outcome. Clin Rheumatol 1987;6 Suppl 2:12-9.
56. Van Gaalen FA, Van Aken J, Huizinga TW, Schreuder GM, Breedveld FC, Zanelli E, et al. Association between HLA class II genes and autoantibodies to cyclic citrullinated peptides (CCPs) influences the severity of rheumatoid arthritis. Arthritis Rheum 2004;50(7):2113-21.
57. Gorman JD, Lum RF, Chen JJ, Suarez-Almazor ME, Thomson G, Criswell LA. Meta-analysis of the shared epitope and severity of rheumatoid arthritis. Arthritis Rheum 2001;44(9 Supplement):S296.
58. Gorman JD, Lum RF, Chen JJ, Suarez-Almazor ME, Thomson G, Criswell LA. Impact of shared epitope genotype and ethnicity on erosive disease: a meta-analysis of 3,240 rheumatoid arthritis patients. Arthritis Rheum 2004;50(2):400-12.
59. Huizinga TWJ, Amos CI, van der Helm-van Mil AHM, Chen W, Van Gaalen FA, Jawaheer D, et al. Refining the complex rheumatoid arthritis phenotype based on specificity of the HLA-DRB1 shared epitope for antibodies to citrullinated proteins. Arthritis Rheum 2005;52(11):3433-8.
60. Gorman JD, Criswell LA. The shared epitope and severity of rheumatoid arthritis. Rheum Dis Clin North Am 2002;28:59-78.
61. MacGregor A, Ollier W, Thomson W, Jawaheer D, Silman A. HLA-DRB1*0401/0404 genotype and rheumatoid arthritis: increased association in men, young age at onset, and disease severity. J Rheumatol 1995;22(6):1032-6.
62. Ollier W, Thomson W. Population genetics of rheumatoid arthritis. Rheum Dis Clin North Am 1992;18:741-759.
63. Burwick RM, Ramsay PP, Haines JL, Hauser SL, Oksenberg JR, Pericak-Vance MA, et al. APOE epsilon variation in multiple sclerosis susceptibility and disease severity: Some answers. Neurology 2006;66(9):1373-1383.

Table 1. Genotype and allele frequencies in RA patients and controls for ten studies included in a meta-analysis of the *CIITA* -168A/G polymorphism (rs3087456) and RA risk.

Reference	Cases/ Controls	Genotypes, % (n)			Alleles, %		OR (95% CI), *P*-value
		AA	AG	GG	A	G	
Akkad et al.(16)	319	55.2 (176)	38.9 (124)	5.9 (19)	74.6	25.4	Neutral
Germany	463	53.8 (249)	39.5 (183)	6.7 (31)	73.5	26.5	
Eyre et al.(17)	1401	54.2 (760)	39.8 (557)	6.0 (84)	74.1	25.9	*Neutral
UK	2476	56.2 (1391)	37.2 (922)	6.6 (163)	74.8	25.2	
Harrison et al.(18)	733	55.0 (404)	37.4 (274)	7.5 (55)	73.8	26.2	Neutral
UK	613	54.8 (336)	39.5 (242)	5.7 (35)	74.6	25.4	
Iikuni et al.(19)	1121	2.8 (31)	22.6 (253)	74.7 (837)	14.0	84.0	Susceptible GG genotype: OR=1.47, 95% CI=1.16-1.87, *P*=0.001
Japan	450	3.3 (15)	30.0 (135)	66.7 (300)	18.3	81.7	
Martinez et al.(20)	350	52.9 (185)	40.3 (141)	6.9 (24)	73.0	27.0	Neutral. **Susceptible G allele: OR=1.81, 95% CI=1.10-2.97, *P*=0.01
Spain	519	57.0 (296)	37.0 (192)	6.0 (31)	75.5	24.5	
Orozco et al.(21)	287	39.7 (114)	40.8 (117)	19.5 (56)	60.1	39.9	Neutral
Argentina	287	38 (109)	48.0 (139)	13.6 (39)	62.2	37.8	
Orozco et al. (21)	748	59.4 (444)	35.0 (262)	5.6 (42)	76.9	23.1	Protective G allele: OR=0.76,

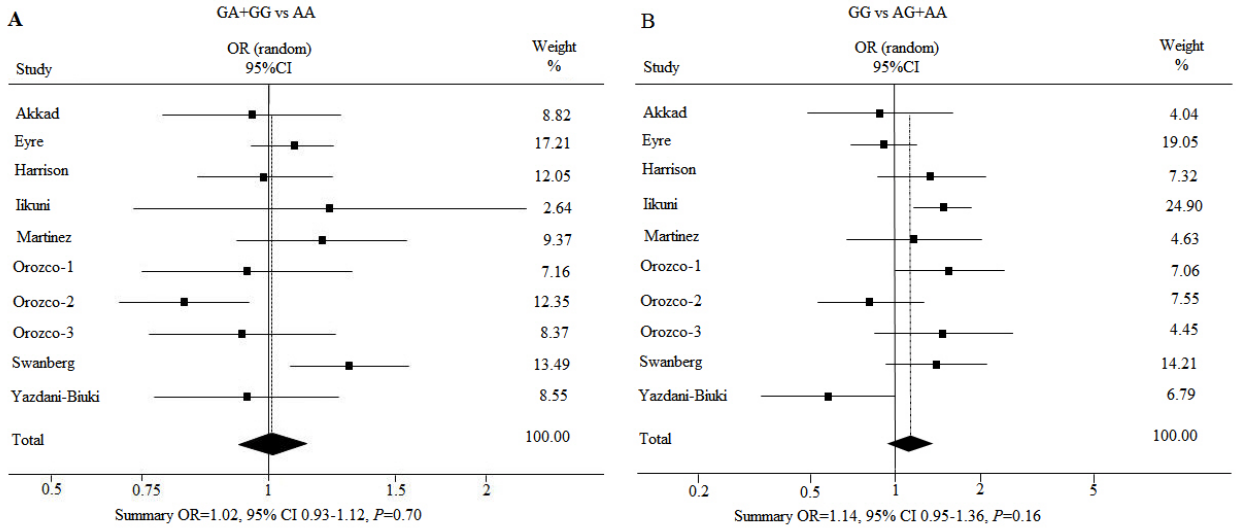
Spain	676	52.7 (356)	40.5 (274)	6.8 (46)	72.9	27.1	
Orozco et al. (21)	278	57.6 (160)	33.8 (94)	8.6 (24)	74.5	25.5	Neutral
Sweden	478	55.4 (265)	38.5 (184)	6.1 (29)	74.7	25.3	
Swanberg et al.(14)	1262	57.7 (728)	35.7 (451)	6.6 (83)	75.6	24.4	Susceptible G allele: OR=1.19, 95% CI=1.02-1.38, *P*=0.02
Sweden	1599	61.9 (989)	33.0 (528)	5.1 (82)	78.4	21.6	
Yazdani-Biuki et al.(22)	362	50.8 (184)	42.8 (155)	6.4 (23)	72.2	27.8	Protective GG genotype: OR=0.58, 95% CI=0.34-0.99, *P*=0.04
Austria	1709	52.3 (894)	39.1 (669)	8.5 (146)	71.9	28.1	
Meta-analysis	6861	46.4 (3186)	35.4 (2428)	18.2 (1247)	64.1	35.9	
All studies	9270	52.9 (4900)	37.4 (3468)	9.7 (902)	71.6	28.4	

[*]The minor allele frequency for rs3087456 was reported incorrectly as the major allele frequency' (A. Barton, pers. comm.).
[**]In a haplotype with rs4774 in patients with one or more shared epitope alleles.

Table 2. Odds ratios (ORs) and 95% confidence intervals (95% CI) for the *CIITA* -168 G risk allele and the G risk genotype in RA patients and controls for ten studies included in a meta-analysis.

Reference	GG+AG vs AA		GG vs AG+AA	
	OR (95% CI)	***P***	**OR (95% CI)**	***P***
Akkad et al.(16) Germany	0.95 (0.71 to 1.26)	0.70	0.88 (0.49 to 1.59)	0.68
Eyre et al.(17) UK	1.08 (0.95 to 1.23)	0.24	0.91 (0.69 to 1.19)	0.47
Harrison et al.(18) UK	0.99 (0.80 to 1.23)	0.91	1.34 (0.86 to 2.08)	0.19
Iikuni et al.(19) Japan	1.21 (0.65 to 2.27)	0.55	1.47 (1.16 to 1.87)	**0.001**
Martinez et al.(20) Spain	1.18 (0.90 to 1.56)	0.22	1.16 (0.67 to 2.01)	0.60
Orozco et al. (21) Argentina	0.93 (0.66 to 1.30)	0.67	1.54 (0.99 to 2.41)	0.06
Orozco et al. (21) Spain	0.76 (0.62 to 0.94)	**0.011**	0.81 (0.53 to 1.25)	0.35
Orozco et al. (21) Sweden	0.92 (0.68 to 1.24)	0.57	1.46 (0.83 to 2.57)	0.18
Swanberg et al.(14) Sweden	1.19 (1.02 to 1.38)	**0.024**	1.30 (0.95 to 1.78)	0.10
Yazdani-Biuki et al. (22) Austria	1.06 (0.85 to 1.33)	0.61	0.73 (0.46 to 1.15)	0.17
Meta-analysis All studies	1.02 (0.93 to 1.12)	0.70	1.14 (0.95 to 1.36)	0.16

Figure 1. Summary odds ratio (OR) and 95% confidence interval (95% CI) for a) the G risk allele and b) the GG risk genotype in RA patients and controls.



CHAPTER THREE

A Candidate Gene Study of *CIITA* Does Not Provide Evidence of Association with Risk for Rheumatoid Arthritis

Submitted: Bronson PG, Ramsay PP, Seldin MF, Gregersen PK, Criswell LA, Barcellos LF. A candidate gene study of CIITA does not provide evidence of association with risk for rheumatoid arthritis. Genes Immun.

ABSTRACT

The major histocompatibility complex (MHC) class II transactivator gene (*CIITA*) encodes an important transcription factor regulating genes required for human leukocyte antigen (HLA) class II MHC-restricted antigen presentation. Major histocompatibility complex (MHC) genes, particularly HLA class II, are strongly associated with risk of developing rheumatoid arthritis (RA). Given the strong biological relationship between *CIITA* and HLA class II genes, a comprehensive investigation of *CIITA* variation in RA gene was conducted. This study tested 31 *CIITA* SNPs in 2,542 RA cases and 3,690 controls (N = 6,232). All individuals were of European ancestry, as determined by ancestry informative genetic markers. No evidence for association between *CIITA* variation and RA was observed after a correction for multiple testing was applied. This is the largest study to fully characterize common genetic variation in *CIITA*, including an assessment of haplotypes. Results do not provide evidence that common variation in *CIITA* plays a role in susceptibility to RA.

BACKGROUND

Rheumatoid arthritis (RA) is the most common systemic autoimmune disease with a prevalence of 1% (1). This chronic inflammatory disease can cause substantial disability from the erosive and deforming processes in joints (2). RA has a strong genetic component, as demonstrated by twin studies; however the etiology is unknown (3). Major histocompatibility complex (MHC) genes, particularly the class II *HLA-DRB1* locus, as well as variants within other MHC regions, are strongly associated with risk of developing RA.

The MHC class II transactivator gene (*CIITA*, also called *MHC2TA*) encodes the CIITA protein, a transcription factor required for the expression of HLA class II molecules (4-7). *CIITA* spans 48 kb on chromosome 16p13 and has four alternate first exons in a 12kb promoter region (I-IV) (8). Mutations in *CIITA* cause a rare and severe immunodeficiency characterized by HLA class II deficiency (bare lymphocyte syndrome) (9). In addition, *CIITA* is located on 16p13, a region that has been implicated in RA linkage studies (10). Thus, *CIITA* is an attractive candidate for genetic studies of autoimmune diseases for which HLA associations have been well established. A comprehensive haplotype-based investigation of *CIITA* as a candidate RA gene was conducted. The study sample consisted of 682 RA cases and 752 controls collected by the North American RA Consortium (NARAC) [RA1], 1860 RA cases and 2938 controls collected by the Wellcome Trust Case Control Consortium (WTCCC) RA Group in the U.K. [RA2] (total $N = 6232$) (Table 1).

METHODS

Patients

RA cases met the American College of Rheumatology classification criteria for RA (11). RA1 controls were frequency matched by age and gender to the cases. RA2 controls were frequency matched by geographical region and gender to the 1958 Birth cohort (which included all births in England, Wales and Scotland, during one week in 1958) so as to be nationally representative. Based on the available genetic ancestry data for all individuals, and to apply the most stringent criteria possible for genetic analysis of *CIITA*, only RA1 subjects with ≥90% N. European ancestry and RA2 subjects with European ancestry were analyzed.
Previous GWA studies provided genotypes for 5 *CIITA* SNPs in RA1 (Illumina HumanHap550 Genotyping BeadChip (San Diego, CA, USA)) and 19 *CIITA* SNPs in RA2 (Affymetrix GeneChips Mapping 500K Array Set (Santa Clara, CA, USA)) as previously described (12-14).

Statistical analysis

European ancestry was estimated in RA1 using a Bayesian clustering algorithm (Structure v. 2.0) and data for 112 European and 246 Northern European ancestry informative markers.(15, 16) For RA2, European ancestry was estimated by principal components analysis (12). Two intronic SNPs in RA2 (rs7404615, rs8062961) were excluded from analysis due to low minor allele frequency (MAF) (<0.01). Deviation from Hardy-Weinberg equilibrium (HWE) was examined in controls separately for each cohort using the exact test (PLINK v. 1.05 (Boston, MA, USA)) (17, 18). There was no evidence for deviation from HWE in the RA1 or RA2 controls ($P<0.01$).

Sufficient power for this study was confirmed with PGA v.2.0 (Bethesda, MD, USA) (MAF 0.1-0.5, two-sided $\alpha=2.9x10^{-3}$).(19) Haplotype blocks were estimated in RA1 and RA2 controls and CEU separately (Haploview v.4.1 (Cambridge, MA, USA)) (20). Percent of *CIITA* variation captured was based on $r^2 \geq 0.8$ in CEU using two- and three-marker haplotypes (HAPLOVIEW).

Allelic association was tested by creating *2x2* contingency tables and estimating ORs with Fisher's exact test (PLINK). Haplotypes were estimated with the expectation-maximization (EM) algorithm (HAPLOVIEW). Maximum likelihood estimates of haplotype probabilities were computed with the EM algorithm and score statistics were used for global haplotype association tests, assuming a dominant genetic model (HaploStats v.1.4.4 (Rochester, MN, USA), R v.2.10.1 (Vienna, AT)) (21). Haplotypes with inferred frequencies <5% were excluded. A significance threshold of $P=2.9x10^{-3}$ was set using a Bonferroni correction for the number of *CIITA* haplotype blocks (four) and SNPs that were not located in haplotype blocks (13), based on CEU. Empirical *P*-values based on 10, 000 simulations were reported for all allelic and haplotype tests. Allelic and haplotype empirical *P*-values were estimated in PLINK (max(T) permutation procedure) and HaploStats, respectively, by permuting the ordering of the disease status, counting the number of times the permuted test was greater than the observed test, and dividing by the total number of simulations (10,000) (17, 21). Because there is no evidence of an association of age or gender with the polymorphisms of interest we decided not to adjust for either.

In order to conduct a combined analysis of RA1 + RA2, missing genotypes were imputed for 16 SNPs in RA1, four SNPs in RA2 and 11 SNPs in the combined RA sample. A hidden Markov Model based algorithm was used to infer missing genotypes from known haplotypes (IMPUTE (v.0.5.0 (Oxford, UK)) (22). The robustness of the imputation accuracy rate for this standard imputation method has been demonstrated (23). Known haplotypes were obtained from publicly available genotype data for CEU, using observed linkage disequilibrium patterns ($r^2 \geq 0.8$) in two 500 kb regions adjacent to each side of *CIITA* (22). Association tests of imputed genotypes accounted for the uncertainty of imputed genotypes in missing data likelihood score tests, using the frequentist proper option and a dominant genetic model in SNPTEST (v.1.1.5 (Oxford, UK)) (22). Imputed genotypes with <90% probability were omitted. There was no evidence for deviation from HWE in the controls. After one SNP with low MAF that was imputed in RA1+RA2 (rs12925158) was omitted from further analyses, 31 SNPs in RA1+RA2 were tested for allelic association.

RESULTS

We conducted allelic tests of association for 5 SNPs and global haplotype tests (one haplotype block encompassing two SNPs) in 682 anti-cyclic citrullinated peptide positive (anti-CCP positive) RA cases and 752 controls (N = 1,434 [RA1]) (Figure 1). All results were negative after correcting for multiple testing (Figure 2, Table 2). Next, we conducted allelic tests of 17 SNPs and global haplotype tests (two haplotype blocks encompassing 11 SNPs) in the second RA dataset comprised of 1860 RA cases and 2938 controls (N = 4798 [RA2]). No evidence for association was present (Figure 2, Table 2). Furthermore, allelic tests of 31 imputed SNPs within *CIITA* derived for the combined RA sample (2542 cases and 3690 controls, total N = 6232 [RA1+RA2]) revealed no evidence for disease association (Figure 2, Table 2).

DISCUSSION

Association between the -168A/G variant in the type III *CIITA* promoter region (rs3087456) and RA was previously reported (24). However, a recent meta-analysis of 10 studies including more than 15,000 individuals revealed no evidence for association between the -168A/G variant and RA (25). Negative findings from the meta-analysis have been further supported by additional reports (26, 27). The current study did not examine the -168A/G variant, as the data were not available from either GWA study, nor could genotypes be imputed using CEU samples from HapMap (see Figure 1).

Modest association between RA and a haplotype containing the -168A/G variant and the +1614G/C missense mutation (rs4774) has been reported in two independent Spanish populations (28, 29). Martinez et al. (2007) report a global haplotype test result of P=0.04, and the odds ratio (OR) and 95% confidence interval (95% CI) for the -168G/+1614C haplotype vs. all other haplotypes was 1.6 (1.05 to 2.44) P=0.02). Martinez et al. (2010) did not report a global haplotype test result, and the OR and 95% CI) for the -168G/+1614C haplotype vs. all other haplotypes was 1.93 (1.10 to 3.45) (P=0.02). Our study was well-powered to detect a modest effect size, with 80% power to detect an allelic OR as low as 1.22. Similar to previous studies, we did not observe any evidence for association between the +1614G/C variant and RA (24, 28, 29). Though genotype data and imputed genotypes were available for the +1614G/C variant in RA2 and RA1, respectively, we didn't examine the -168G/+1614C haplotype because data for the -168A/G variant were not available. Nevertheless, the associations reported by Martinez et al. do not reach significance criteria following correction for multiple testing. Furthermore, in contrast to the current study, Martinez et al. did not estimate European ancestry to protect against spurious association due to population stratification effects.

Although rare variants in *CIITA* were not directly investigated here, for the first time all common genetic variation within *CIITA* was interrogated for a role in RA susceptibility. The 31 SNPs in the combined RA sample captured 94% of common variation based on Caucasian HapMap population [CEU] data (see Figure 2). The combined sample tagged all but two of the common HapMap variants (rs6498122 (intronic variant) and rs8046121 (missense mutation)). The RA1 dataset (n = 5 SNPs) captured 27% and the RA2 dataset (n = 17 SNPs) captured 76% of common variation. The data used in this study were taken from two genome-wide association (GWA) studies that did not identify *CIITA* as a risk locus for RA based on stringent significance criteria. A focused candidate gene study that captures a much larger portion of genetic variation when compared to initial GWA studies is a useful and complementary strategy.

CONCLUSION

In conclusion, this is the first genetic study of RA to fully characterize common genetic variation in *CIITA* including assessment of haplotypes. Results do not provide evidence that common variation in *CIITA* plays a role in susceptibility to RA.

ACKNOWLEDGEMENTS

We would like to thank Farren Briggs, Benjamin Goldstein, Alan Hubbard, Ira Tager and Gary Artim. This work was supported by an Abbott Graduate Student Achievement Award (ACR REF), grants F31 AI075609, R01 AI065841 and R01 AI059829 (NIH/NIAID), and grants

RO1 AR44422, NO1 AR22263, R01 AR050267, K24 AR02175 (NIH/NIAMS). The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the NIH, NIAID or NIAMS. This study makes use of data generated by the WTCCC; a full list of the investigators who contributed to the generation of the data is available from www.wtccc.org.uk, and funding for the project was provided by the Wellcome Trust under award 076113. These studies were performed in part in the General Clinical Research Center, Moffitt Hospital, University of California, San Francisco, with funds provided by the National Center for Research Resources, 5 M01 RR-00079, U.S. Public Health Service.

REFERENCES

1. Gabriel SE. The epidemiology of rheumatoid arthritis. Rheum Dis Clin North Am 2001;27(2):269-81.
2. Gabriel SE, Crowson CS, Kremers HM, Doran MF, Turesson C, O'Fallon WM, et al. Survival in rheumatoid arthritis: a population-based analysis of trends over 40 years. Arthritis Rheum 2003;48(1):54-8.
3. MacGregor AJ, Snieder H, Rigby AS, Koskenvuo M, Kaprio J, Aho K, et al. Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins. Arthritis Rheum 2000;43(1):30-7.
4. Chang CH, Flavell RA. Class II transactivator regulates the expression of multiple genes involved in antigen presentation. J Exp Med 1995;181(2):765-767.
5. Chang CH, Guerder S, Hong SC, van Ewijk W, Flavell RA. Mice lacking the MHC class II transactivator (CIITA) show tissue-specific impairment of MHC class II expression. Immunity 1996;4(2):167-178.
6. Kern I, Steimle V, Siegrist CA, Mach B. The two novel MHC class II transactivators RFX5 and CIITA both control expression of HLA-DM genes. Int Immunol 1995;7(8):1295-1299.
7. LeibundGut-Landmann S, Waldburger JM, Krawczyk M, Otten LA, Suter T, Fontana A, et al. Specificity and expression of CIITA, the master regulator of MHC class II genes. Eur J Immunol 2004;34(6):1513-1525.
8. Muhlethaler-Mottet A, Otten LA, Steimle V, Mach B. Expression of MHC class II molecules in different cellular and functional compartments is controlled by differential usage of multiple promoters of the transactivator CIITA. Embo J 1997;16(10):2851-60.
9. Steimle V, Otten LA, Zufferey M, Mach B. Complementation cloning of an MHC class II transactivator mutated in hereditary MHC class II deficiency (or bare lymphocyte syndrome). Cell 1993;75(1):135-46.
10. Etzel CJ, Chen WV, Shepard N, Jawaheer D, Cornelis F, Seldin MF, et al. Genome-wide meta-analysis for rheumatoid arthritis. Hum Genet 2006;119(6):634-41.
11. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988;31(3):315-324.
12. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007;447(7145):661-678.
13. Plenge RM, Seielstad M, Padyukov L, Lee AT, Remmers EF, Ding B, et al. TRAF1-C5 as a Risk Locus for Rheumatoid Arthritis -- A Genomewide Study. N Engl J Med 2007;357(12):1199-1209.
14. Hom G, Graham RR, Modrek B, Taylor KE, Ortmann W, Garnier S, et al. Association of Systemic Lupus Erythematosus with C8orf13-BLK and ITGAM-ITGAX. N Engl J Med 2008;358(9):900-909.
15. Falush D, Stephens M, Pritchard JK. Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. Genetics 2003;164(4):1567-1587.
16. Tian C, Plenge RM, Ransom M, Lee A, Villoslada P, Selmi C, et al. Analysis and Application of European Genetic Substructure Using 300 K SNP Information. PLoS Genetics 2008;4(1):29-39.

17. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81(3):559-75.
18. Wigginton JE, Cutler DJ, Abecasis GR. A note on exact tests of Hardy-Weinberg equilibrium. Am J Hum Genet 2005;76(5):887-93.
19. Menashe I, Rosenberg PS, Chen BE. PGA: power calculator for case-control genetic association analyses. BMC Genet 2008;9:36.
20. Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, et al. The structure of haplotype blocks in the human genome. Science 2002;296(5576):2225-9.
21. Schaid DJ, Rowland CM, Tines DE, Jacobson RM, Poland GA. Score tests for association between traits and haplotypes when linkage phase is ambiguous. Am J Hum Genet 2002;70(2):425-34.
22. Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. Nat Genet 2007;39(7):906-913.
23. Pei Y-F, Li J, Zhang L, Papasian CJ, Deng H-W. Analyses and Comparison of Accuracy of Different Genotype Imputation Methods. PLoS ONE 2008;3(10):e3551.
24. Swanberg M, Lidman O, Padyukov L, Eriksson P, Akesson E, Jagodic M, et al. MHC2TA is associated with differential MHC molecule expression and susceptibility to rheumatoid arthritis, multiple sclerosis and myocardial infarction. Nat Genet 2005;37(5):486-494.
25. Bronson PG, Criswell LA, Barcellos LF. The MHC2TA -168A/G polymorphism and risk for rheumatoid arthritis: a meta-analysis of 6861 patients and 9270 controls reveals no evidence for association. Ann Rheum Dis 2008;67(7):933-6.
26. Plant D, Barton A, Thomson W, Ke X, Eyre S, Hinks A, et al. A re-evaluation of three putative functional single nucleotide polymorphisms in rheumatoid arthritis. Ann Rheum Dis 2009;68(8):1373-1375.
27. Dieguez-Gonzalez R, Akar S, Calaza M, Gonzalez-Alvaro I, Fernandez-Gutierrez B, Lamas JR, et al. Lack of Association with Rheumatoid Arthritis of Selected Polymorphisms in 4 Candidate Genes: CFH, CD209, Eotaxin-3, and MHC2TA. Journal of Rheumatology 2009;36(8):1590-1595.
28. Martinez A, Sanchez-Lopez M, Varade J, Mas A, Martin MC, Heras VDL, et al. Role of the MHC2TA gene in autoimmune diseases. Ann Rheum Dis 2007;66(3):325-329.
29. Martinez A, Perdigones N, Cenit MC, Espino L, Varade J, Lamas JR, et al. Chromosomal region 16p13: further evidence of increased predisposition to immune diseases. Ann Rheum Dis 2010;69(1):309-11.

Table 1. RA study cohorts utilized for *CIITA* analyses.

	RA1	**Controls**	**RA2**	**Controls**
N	682	752	1860	2938
Site	N.A.	N.A.	U.K.	U.K.
Mean age, years	56.2	48.5	-	-
Age range, years	21-87	30-82	-	21-72
Female, *N (%)*	503 (73.7)	525 (69.8)	1390 (74.7)	1492 (50.8)
Mean age-at-onset, years	45.7		-	
Rheumatoid factor positive, *N (%)*	580 (85)		1310 (83.9)	
Shared epitope[a] (no. of copies), *N (%)*				
0	15 (2.3)	401 (53.3)	286 (20.7)	
1	362 (56.5)	301 (40)	680 (49.2)	
2	264 (41.2)	50 (6.6)	416 (30.1)	
Erosions, *N (%)*	211 (66.6)		-	
Anti-CCP Positive, *N (%)*	681 (100)		884 (79.8)	

[1]*HLA-DRB1*0101, *0102, *0104, *0401, *0404, *0405, *0408, *0413, *0416, *1001* alleles.

Table 2 Minor allele frequencies (MAF), odds ratios (OR), 95% confidence intervals (95% CI) and *P*-values from allelic tests and *P*-values from global haplotype tests in healthy controls and RA cases.

				MAF		Allelic test		Haplotype test	
Marker	Location	Function	Sample	Cases	Controls	OR (95% CI)	*P*	Block	*P*
rs12928665	A10971474G	intron	RA1	0.227	0.227	1.00 (0.84-1.19)	0.97	1	0.05
			RA1+RA2	0.241	0.246	0.98 (0.88-1.08)	0.66		
rs12932187	C10971880G	intron	RA1+RA2	0.020	0.020	0.94 (0.76-1.15)	0.53		
rs4781011	C10975311A	intron	RA1	0.230	0.238	0.96 (0.80-1.14)	0.62	1	0.05
			RA1+RA2	0.254	0.258	0.99 (0.89-1.09)	0.78		
rs11074934	C10979440T	intron	RA2	0.285	0.285	1.00 (0.91-1.09)	0.99	1	0.86
			RA1+RA2	0.278	0.280	1.01 (0.91-1.12)	0.87		
rs8043545	G10982345C	intron	RA2	0.273	0.273	1.00 (0.91-1.09)	0.99	1	0.86
			RA1+RA2	0.266	0.267	1.05 (0.86-1.28)	0.63		
rs8063850	A10991621T	intron	RA2	0.271	0.273	0.99 (0.91-1.09)	0.91	1	0.86
			RA1+RA2	0.173	0.175	1.00 (0.89-1.12)	0.96		
rs6498119	T10991878C	intron	RA2	0.062	0.066	0.93 (0.78-1.10)	0.39	1	0.86
			RA1+RA2	0.054	0.059	1.03 (0.44-2.38)	0.97		
rs4781015	A10991952G	intron	RA1+RA2	0.172	0.174	1.21 (0.91-1.61)	0.19		
rs8048002	T10991988C	intron	RA2	0.062	0.066	0.94 (0.79-1.11)	0.44	1	0.86
			RA1+RA2	0.051	0.058	1.04 (0.43-2.49)	0.95		
rs7189406	A10993488G	intron	RA2	0.073	0.069	1.06 (0.91-1.25)	0.47		
			RA1+RA2	0.010	0.011	1.08 (0.87-1.34)	0.51		
rs6498124	G10995850T	intron	RA1+RA2	0.460	0.431	1.08 (0.88-1.33)	0.43		
rs4781016	C10996399A	intron	RA1	0.287	0.287	1.00 (0.85-1.18)	0.99		
			RA2	0.268	0.275	0.97 (0.88-1.06)	0.51		
			RA1+RA2	0.224	0.230	1.09 (0.86-1.37)	0.49		
rs4774	G11000848C	missense	RA2	0.285	0.287	0.99 (0.90-1.08)	0.81		
			RA1+RA2	0.256	0.260	1.07 (0.87-1.33)	0.52		
rs4781018	C11002133G	intron	RA1+RA2	0.286	0.290	1.09 (0.90-1.33)	0.36		
rs4780334	A11002626G	intron	RA1+RA2	0.286	0.289	1.00 (0.90-1.12)	0.97		
rs4781019	A11004150G	intron	RA1	0.433	0.439	0.98 (0.84-1.13)	0.74		
			RA1+RA2	0.429	0.427	1.07 (0.92-1.25)	0.39		
rs4780335	C11004328G	intron	RA2	0.434	0.433	1.00 (0.92-1.09)	0.97		
			RA1+RA2	0.428	0.432	1.04 (0.93-1.16)	0.51		
rs6498126	C11004363G	intron	RA2	0.205	0.200	1.03 (0.93-1.14)	0.60		
			RA1+RA2	0.114	0.121	1.02 (0.89-1.15)	0.81		
rs7204799	G11004549C	intron	RA2	0.081	0.081	1.00 (0.86-1.16)	0.99	2	0.97
			RA1+RA2	0.060	0.067	0.62 (0.29-1.33)	0.21		
rs12598246	G11004732A	intron	RA1	0.296	0.277	1.10 (0.94-1.30)	0.25		
			RA1+RA2	0.224	0.227	1.04 (0.92-1.17)	0.58		
rs11074938	A11006543G	intron	RA1+RA2	0.374	0.377	0.94 (0.80-1.10)	0.44		
rs4781020	A11008262G	intron	RA1+RA2	0.284	0.280	1.00 (0.83-1.20)	0.96		
rs6498130	G11010150T	intron	RA1+RA2	0.418	0.426	1.01 (0.90-1.14)	0.84		

rs6498132	C11010644T	intron	RA2	0.104	0.105	1.00 (0.87-1.14)	0.97	2	0.97
			RA1+RA2	0.091	0.098	0.94 (0.83-1.07)	0.36		
rs11074939	G11011709A	intron	RA2	0.280	0.282	0.99 (0.90-1.08)	0.85	2	0.97
			RA1+RA2	0.250	0.249	0.99 (0.81-1.22)	0.91		
rs9788916	G11011728A	intron	RA2	0.105	0.108	0.97 (0.85-1.11)	0.63	2	0.97
			RA1+RA2	0.091	0.098	1.30 (0.78-2.15)	0.31		
rs7404786	C11012550G	intron	RA2	0.438	0.439	1.00 (0.92-1.08)	0.95	2	0.97
			RA1+RA2	0.420	0.427	1.01 (0.90-1.14)	0.85		
rs8056269	C11012567G	intron	RA2	0.435	0.437	0.99 (0.92-1.08)	0.90	2	0.97
			RA1+RA2	0.420	0.427	1.01 (0.90-1.14)	0.83		
rs2229322	C11016045T	syncod[1]	RA2	0.110	0.117	0.94 (0.82-1.07)	0.33		
			RA1+RA2	0.049	0.051	0.98 (0.84-1.15)	0.79		
rs4781024	A11017058G	intron	RA1+RA2	0.409	0.399	1.00 (0.86-1.17)	0.97		
rs1139564	C11018622T	utr-3[2]	RA1+RA2	0.101	0.111	0.89 (0.79-1.00)	0.05		

[1]Synonymous coding
[2]Untranslated region

Figure 1. Schematic of our analysis strategy in stages (**a**) 1, (**b**) 2 and (**c**) 3.

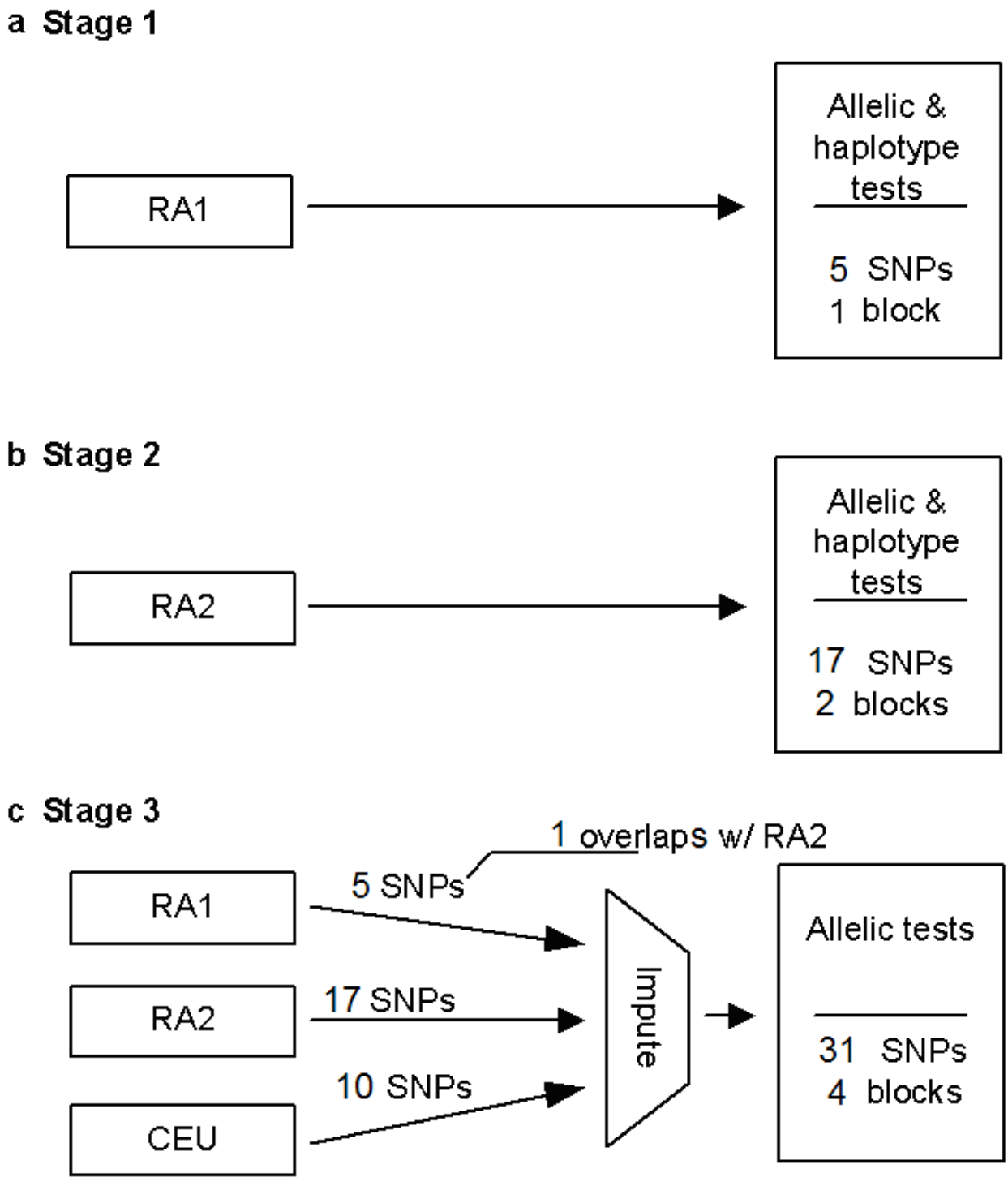
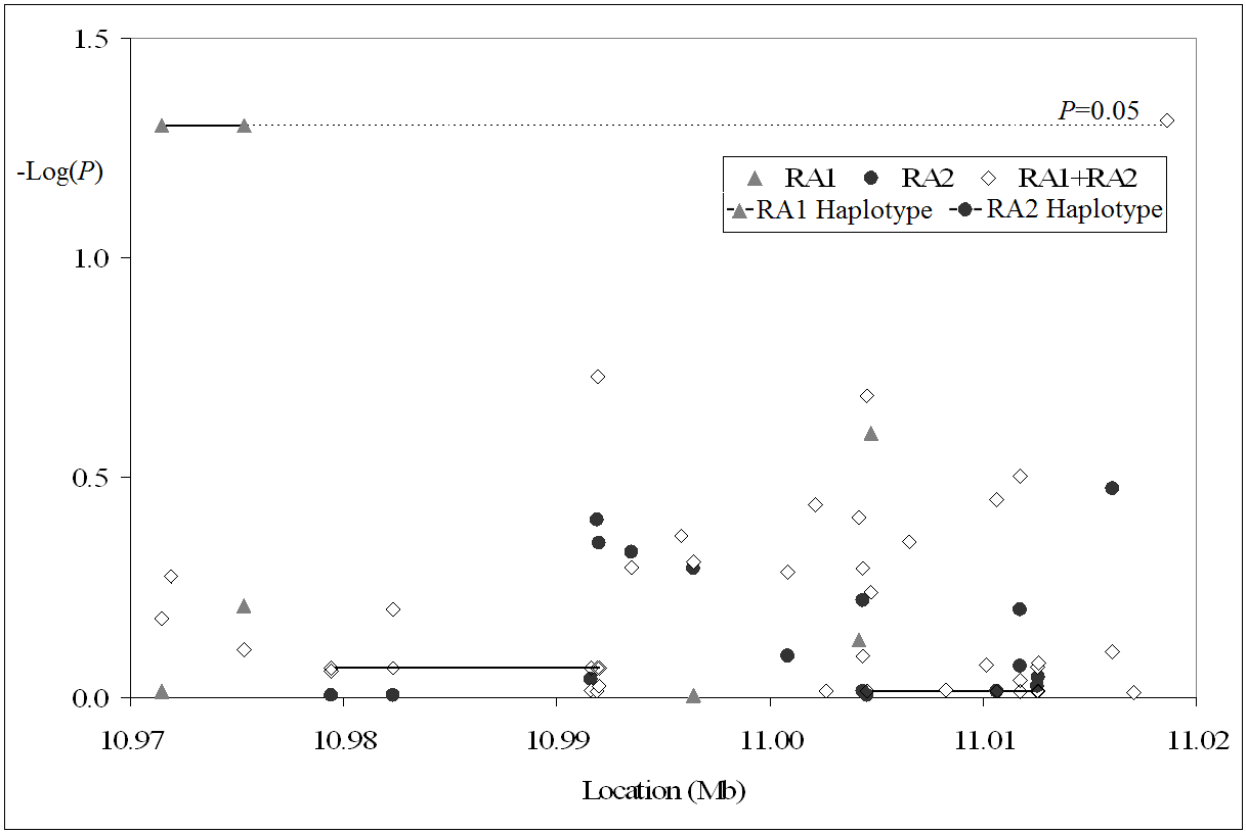


Figure 2. *P*-values from allelic and haplotype tests of *CIITA* SNPs in RA.



CHAPTER FOUR

CIITA Does Not Provide Evidence of Association with Risk for Systemic Lupus Erythematosus

Submitted: Bronson PG, May S, Ramsay PP, Beckman KB, Noble JA, Lane JA, Seldin MF, Kelly JA, Harley JB, Moser KL, Gaffney PM, Behrens TW, Criswell LA, Barcellos LF. CIITA Does Not Provide Evidence of Association with Risk for Systemic Lupus Erythematosus. Arthritis Rheum.

ABSTRACT

The major histocompatibility complex (MHC) class II transactivator gene (*CIITA*) is an important transcription factor regulating gene required for human leukocyte antigen (HLA) class II MHC-restricted antigen presentation. Association with HLA class II variation, particularly *DRB1*1501* and **0301*, has been well-established for systemic lupus erythematosus (SLE). Thus, we investigated *CIITA* variation, *HLA-DRB1*1501*, *DRB1*0301* and secondary phenotypes in SLE. We tested 18 *CIITA* SNPs in 637 SLE trio families and 826 independent, unrelated SLE cases (n = 2,737). Family-based association tests were conducted to compare transmitted vs. nontransmitted alleles. Case-control association tests were conducted in the combined trios and unrelated cases, utilizing nontransmitted parental alleles as controls. Analyses were stratified by *DRB1*1501*, *DRB1*0301*, LN, arthritis, serositis, neurological involvement, Sm and Ro autoantibody production. Case-only analyses of *DRB1*1501*, *DRB1*0301* and clinical phenotypes were conducted. No evidence for association was observed between *CIITA* and SLE in family-based and case-control analyses after correcting for multiple testing. Cases exhibited modest evidence for association between the rs11074938**G* variant and the presence of either *DRB1*1501* or **0301* (OR = 1.27, 95% CI = 1.11 to 1.47, $P = 2 \times 10^{-3}$), and the rs45617532**C* variant and neurological involvement (OR = 4.07, 95% CI = 1.83 to 9.06, $P = 1.2 \times 10^{-3}$). This is the first study of SLE to fully characterize common genetic variation in *CIITA*, including secondary phenotypes and HLA risk alleles. Results do not provide evidence that common variation in *CIITA* plays a role in susceptibility to SLE.

BACKGROUND

Systemic lupus erythematosus (SLE) is the prototypic systemic autoimmune disease and is characterized by autoantibody production and altered immune complex formation and clearance (1). SLE has a strong genetic component, as demonstrated by twin and other family studies; however the etiology is unknown (2). Major histocompatibility complex (MHC) genes on chromosome 6p21, particularly the class II human leukocyte antigen (HLA)-*DRB1*1501* and **0301* alleles, are strongly associated with increased risk of developing SLE (3-5). However, MHC genes only account for a portion of the genetic risk. Several non-MHC genes have recently been associated with risk for SLE, including *PTPN22*, *STAT4* and *TNFAIP3* (6-9).

The MHC class II transactivator gene (*CIITA*, also called *MHC2TA*) encodes the CIITA protein, a transcription factor essential for the expression of HLA class II molecules and involved in the expression of HLA class I molecules (10-18). *CIITA* spans 53 kb on chromosome 16p13, with four alternate first exons in a 12kb promoter region (I-IV) (19). The gene is adjacent to a recently identified MS and T1D risk locus (C-type lectin domain family 16, member A gene, *CLEC16A*) (10-13). Mutations in *CIITA* cause a rare and severe immunodeficiency characterized by HLA class II deficiency (bare lymphocyte syndrome) (20). Thus, *CIITA* is an attractive candidate for genetic studies of autoimmune diseases for which HLA associations have been well established.

The purpose of this study was to conduct a comprehensive investigation of common variation in *CIITA* and susceptibility to SLE, accounting for lupus nephritis (LN) and other secondary phenotypes and the presence of *DRB1*1501* and **0301* risk alleles.

METHODS

Patients

The cohort consisted of 637 trio families with one SLE-affected child and two parents and 826 unrelated, independent SLE cases (n = 2,737) of European ancestry. All SLE cases met the American College of Rheumatology classification criteria for SLE (21). SLE cases were enrolled through the University of California, San Francisco (UCSF) Lupus Genetics Project collection (n = 827), the Lupus Genetics Studies and Lupus Family Registry and Repository at the Oklahoma Medical Research Foundation (OMRF) (n = 389), and the University of Minnesota collection (n = 247) as previously described (Table 1) (5). Data were collected by questionnaire and chart review. *DRB1* genotypes with four-digit DRB1 resolution, obtained through PCR-SSO methodology, were available for all cases. Data for 20 *CIITA* SNPs were derived from Sequenom iPLEX genotyping (San Diego, CA, USA) at the Children's Hospital Oakland Research Institute.

European ancestry was estimated based on genotypes available for 112 European ancestry informative markers as previously described (5, 22, 23). To apply the most stringent criteria possible for genetic analysis of *CIITA*, only SLE patients with ≥80% European ancestry were analyzed.

Statistical analysis

Stringent quality control criteria were applied to all samples and genotype data. One rare missense mutation SNP (rs4781022) was excluded from analysis because it was monomorphic. No additional SNPs were omitted due to low minor allele frequency (MAF) (<0.01). Mendelian inconsistencies for one, two and three SNPs were identified in 19, 7 and four SLE trio families, respectively, and these specific genotypes were omitted from analysis (PLINK v1.05) (24). One promoter polymorphism (rs9302456) was dropped because of excessive Mendelian error rate in the trio families (>0.01). Deviation from Hardy-Weinberg equilibrium (HWE) was examined in the SLE patients and unaffected parents separately using the exact test (PLINK v1.05) (24, 25). There were no SNPs with evidence for deviation from HWE (P<0.01).

The transmission disequilibrium test (TDT) was used to examine differences between transmitted and nontransmitted *CIITA* alleles in affected offspring (PLINK) (24, 26). A case-control analysis was also performed with a log likelihood ratio test of the cases vs. affected family-based controls (AFBACs, or nontransmitted parental alleles), implemented in the UNPHASED program v3.0.13 (27). The TDT and log likelihood ratio tests were stratified by the presence and absence of *DRB1*1501*, *DRB1*0301*, *DRB1*1501* or **0301* and secondary phenotypes (LN, arthritis, serositis, history of neurological involvement, Sm autoantibody production, Ro autoantibody production). A case-only test was conducted by using logistic regression (PLINK) to test for association between *CIITA* variation and the presence of HLA risk alleles and secondary phenotypes, assuming an additive genetic model (24).

All reported P-values are empirically based on ≥10,000 permutations and are two-tailed. Empirical P-values were estimated by permuting tests, counting the number of times the permuted test was greater than the observed test, and dividing by the total number of simulations (10,000). PLINK (max(T) permutation procedure) was used to permute the transmitted and untransmitted status of SNPs in each family in the family-based analyses (24). UNPHASED was used to permute disease status in the case-control analyses.(27). PLINK (max(T) permutation procedure) was used to permute the ordering of the outcome status (24).

A significance threshold of $P = 4.16 \times 10^{-3}$ was set using a Bonferroni correction for the number of independent single-locus association tests in the overall sample (4 haplotype blocks and 8 SNPs that did not fall into haplotype blocks). Power was estimated (PGA v2.0) assuming a two-sided type I error of $\alpha = 4.16 \times 10^{-3}$, to account for the number of statistical tests (28). The current study had 80% power to detect an OR ranging from 1.3 to 1.55 under varying MAF (0.1-0.5).

Haplotypes were estimated in the cases and unaffected parents separately with the expectation-maximization algorithm and haplotype blocks were determined with the confidence bound algorithm (HAPLOVIEW v4.1) (**Figure 1**) (29). Percent of *CIITA* variation captured was based on $r^2 \geq 0.8$ in HapMap European individuals of Northern and Western European origin (CEU) using two- and three- marker haplotypes (HAPLOVIEW).

RESULTS

We conducted allelic tests of association for 18 SNPs (**Table 2**) stratified by the presence and absence of *DRB1*1501* and/or **0301* and secondary phenotypes. No associations were observed after correcting for multiple testing in the overall family-based and case-control analyses (**Figure 2A**). In the case-only analyses, modest evidence for association was observed between the rs11074938**G* variant and the presence of either *DRB1*1501* or **0301* (odds ratio [OR] = 1.27, 95% confidence interval [95% CI] = 1.11-1.47, $P = 2x10^{-3}$), and the rs45617532**C* variant and history of neurological involvement (OR = 4.07, 95% CI = 1.83-9.06, $P = 1.2x10^{-3}$) (**Figure 2B**).

DISCUSSION

One of the most serious clinical outcomes of SLE is lupus nephritis (LN), which develops in up to half of SLE patients and, despite treatment, progressed to end stage renal disease in one-fifth of patients with SLE (30, 31). LN is a marker of severe disease, a strong predictor of adverse outcomes and a leading cause of damage associated with SLE (32-36). The Fcγ receptor IIIA gene (FCGR3A) V/F158 polymorphism has been shown to increase LN risk in SLE patients by 20%, yet it is not associated with SLE risk (37). Due to its common frequency, a tenth of LN cases may be attributable to the F158 variant (37).Therefore, it appears worthwhile to evaluate LN, and other secondary phenotypes, in genetic studies of SLE.(38)

Although rare variants in *CIITA* were not directly investigated here, for the first time all common genetic variation within *CIITA* was interrogated for SLE. Approximately sixty percent of the common genetic variation (MAF ≥ 0.01) in *CIITA* and its promoter region was captured in the current study, based on $r^2 \geq 0.8$ in CEU. An additional 11 *CIITA* HapMap variants were captured (**Table 2**). Though this study failed to capture 15 common *CIITA* HapMap variants, all of them were intronic, and perhaps less likely to play role in SLE susceptibility, with the exception of one missense mutation and one promoter polymorphism (**Table 3**). In addition, this study did not capture the rare rs7197779 HapMap missense mutation (MAF 0.009).

The *CIITA* −168A/G promoter polymorphism (rs3087456) has been reported to be associated with SLE risk in a Japanese population (39). However, no evidence for association was observed in studies of Swedish and Spanish populations (40, 41). The -168A/G polymorphism was not investigated in this study. Sanchez et al. also evaluated the *CIITA* +1614C/G missense mutation (rs4774), recently reported to be associated with multiple sclerosis (MS) in the presence of the *DRB1*1501* MS risk allele; similar to the current study, Sanchez et al. did not observe evidence for association between the +1614C/G variant and SLE risk (41, 42).

Modest evidence for association was observed between the untransalated 3' rs11074938**G* variant and the presence of either *DRB1*1501* or **0301*, and the intronic rs45617532**C* variant and history of neurological involvement in the cases only. However, in light of the large number of tests conducted in the current study, these results should be interpreted with caution. It may be worthwhile to examine these variants in future genetic studies of SLE phenotypes.

CONCLUSION

In conclusion, this is the first SLE genetic study of *CIITA* to fully characterize common genetic variation in *CIITA* and assess secondary phenotypes and HLA risk alleles. Results do not provide evidence that common variation in *CIITA* plays a role in susceptibility to SLE.

ACKNOWLEDGEMENTS

We thank Farren Briggs, Benjamin Goldstein, Alan Hubbard and Gary Artim. Support was provided by the Kirkland Scholar Award (LAC, JBH) from the Mary Kirkland Center for Lupus Research, the Abbott/REF Graduate Student Achievement Award (PGB) from the American College of Rheumatology, the Alliance for Lupus Research (JBH, KLM, LAC), the U.S. Department of Veterans Affairs (JBH), Lupus Foundation Minnesota (KLM) and NIH grants R01-AR44804 (LAC), R01-AR052300 (LAC), K24-AR02175 (LAC), P01 AR049084 (JBH), R01 AR42460 (JBH), R01 AR043274-14 (KLM) and N01 AR62277 (JBH) from NIAMS, R37 AI24717 (JBH), P01 AI083194 (JBH), 5R01 A1063274-06 (PMG) and F31-AI075609 (PGB) from NIAID, and P20-RR020143 (JBH) from NCRR. This study was performed in part in the General Clinical Research Center, Moffitt Hospital, UCSF, with funds provided by NCRR, 5-M01-RR-00079, U.S. Public Health Service.

REFERENCES

1. Kelley W, Harris E, Jr., Ruddy S, Sledge C. Textbook of Rheumatology. Philadelphia: Harcourt Brace Jovanovich, Inc.; 1993.
2. Deapen D, Escalante A, Weinrib L, Horwitz D, Bachman B, Roy-Burman P, et al. A revised estimate of twin concordance in systemic lupus erythematosus. Arthritis Rheum 1992;35:311-317.
3. Graham RR, Ortmann WA, Langefeld CD, Jawaheer D, Selby SA, Rodine PR, et al. Visualizing human leukocyte antigen class II risk haplotypes in human systemic lupus erythematosus. Am J Hum Genet 2002;71(3):543-53.
4. Graham RR, Ortmann W, Rodine P, Espe K, Langefeld C, Lange E, et al. Specific combinations of HLA-DR2 and DR3 class II haplotypes contribute graded risk for disease susceptibility and autoantibodies in human SLE. Eur J Hum Genet 2007;15(8):823-30.
5. Barcellos LF, May SL, Ramsay PP, Quach HL, Lane JA, Nititham J, et al. High-density SNP screening of the major histocompatibility complex in systemic lupus erythematosus demonstrates strong evidence for independent susceptibility regions. PLoS Genet 2009;5(10):e1000696.
6. Kyogoku C, Langefeld CD, Ortmann WA, Lee A, Selby S, Carlton VE, et al. Genetic association of the R620W polymorphism of protein tyrosine phosphatase PTPN22 with human SLE. Am J Hum Genet 2004;75(3):504-7.
7. Remmers EF, Plenge RM, Lee AT, Graham RR, Hom G, Behrens TW, et al. STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus. N Engl J Med 2007;357(10):977-86.
8. Musone SL, Taylor KE, Lu TT, Nititham J, Ferreira RC, Ortmann W, et al. Multiple polymorphisms in the TNFAIP3 region are independently associated with systemic lupus erythematosus. Nat Genet 2008;40(9):1062-4.
9. Graham RR, Cotsapas C, Davies L, Hackett R, Lessard CJ, Leon JM, et al. Genetic variants near TNFAIP3 on 6q23 are associated with systemic lupus erythematosus. Nat Genet 2008;40(9):1059-61.
10. Rubio JP, Stankovich J, Field J, Tubridy N, Marriott M, Chapman C, et al. Replication of KIAA0350, IL2RA, RPL5 and CD58 as multiple sclerosis susceptibility genes in Australians. Genes Immun 2008;9(7):624-30.
11. Hafler DA, Compston A, Sawcer S, Lander ES, Daly MJ, De Jager PL, et al. Risk alleles for multiple sclerosis identified by a genomewide study. N Engl J Med 2007;357(9):851-62.
12. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007;447(7145):661-678.
13. Hakonarson H, Grant SFA, Bradfield JP, Marchand L, Kim CE, Glessner JT, et al. A genome-wide association study identifies KIAA0350 as a type 1 diabetes gene. Nature 2007;448(7153):591-594.
14. Todd JA, Walker NM, Cooper JD, Smyth DJ, Downes K, Plagnol V, et al. Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes. Nat Genet 2007;39(7):857-864.
15. Chang CH, Flavell RA. Class II transactivator regulates the expression of multiple genes involved in antigen presentation. J Exp Med 1995;181(2):765-767.

16. Chang CH, Guerder S, Hong SC, van Ewijk W, Flavell RA. Mice lacking the MHC class II transactivator (CIITA) show tissue-specific impairment of MHC class II expression. Immunity 1996;4(2):167-178.
17. Kern I, Steimle V, Siegrist CA, Mach B. The two novel MHC class II transactivators RFX5 and CIITA both control expression of HLA-DM genes. Int Immunol 1995;7(8):1295-1299.
18. LeibundGut-Landmann S, Waldburger JM, Krawczyk M, Otten LA, Suter T, Fontana A, et al. Specificity and expression of CIITA, the master regulator of MHC class II genes. Eur J Immunol 2004;34(6):1513-1525.
19. Muhlethaler-Mottet A, Otten LA, Steimle V, Mach B. Expression of MHC class II molecules in different cellular and functional compartments is controlled by differential usage of multiple promoters of the transactivator CIITA. Embo J 1997;16(10):2851-60.
20. Steimle V, Otten LA, Zufferey M, Mach B. Complementation cloning of an MHC class II transactivator mutated in hereditary MHC class II deficiency (or bare lymphocyte syndrome). Cell 1993;75(1):135-46.
21. Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum 1997;40(9):1725.
22. Tian C, Plenge RM, Ransom M, Lee A, Villoslada P, Selmi C, et al. Analysis and Application of European Genetic Substructure Using 300 K SNP Information. PLoS Genetics 2008;4(1):29-39.
23. Kosoy R, Nassir R, Tian C, White PA, Butler LM, Silva G, et al. Ancestry informative marker sets for determining continental origin and admixture proportions in common populations in America. Hum Mutat 2009;30(1):69-78.
24. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81(3):559-75.
25. Wigginton JE, Cutler DJ, Abecasis GR. A note on exact tests of Hardy-Weinberg equilibrium. Am J Hum Genet 2005;76(5):887-93.
26. Spielman RS, Ewens WJ. The TDT and other family-based tests for linkage disequilibrium and association. Am J Hum Genet 1996;59:983-989.
27. Dudbridge F. Likelihood-based association analysis for nuclear families and unrelated subjects with missing genotype data. Hum Hered 2008;66(2):87-98.
28. Menashe I, Rosenberg PS, Chen BE. PGA: power calculator for case-control genetic association analyses. BMC Genet 2008;9:36.
29. Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, et al. The structure of haplotype blocks in the human genome. Science 2002;296(5576):2225-9.
30. Huong DL, Papo T, Beaufils H, Wechsler B, Bletry O, Baumelou A, et al. Renal involvement in systemic lupus erythematosus. A study of 180 patients from a single center. Medicine (Baltimore) 1999;78(3):148-66.
31. Ponticelli C, Moroni G. Flares in lupus nephritis: incidence, impact on renal survival and management. Lupus 1998;7(9):635-8.
32. Golbus J, McCune WJ. Lupus nephritis: classification, prognosis, immunopathogenesis, and treatment. In; 1994. p. 213-242.
33. Ginzler EM, Diamond HS, Weiner M, Schlesinger M, Fries JF, Wasner C, et al. A multicenter study of outcome in systemic lupus erythematosus. I. Entry variables as predictors of prognosis. Arthritis Rheum 1982;25:601-611.

34. Reveille JD, Bartolucci A, Alarcon GS. Prognosis in systemic lupus erythematosus. Negative impact of increasing age at onset, black race, and thrombocytopenia, as well as causes of death. Arthritis Rheum 1990;33(1):37-48.
35. Gladman DD, Urowitz MB. Systemic lupus erythematosus: clinical manifestations. In: Klippel JH, Dieppe PA, editors. Rheumatology. London: Mosby-Year Book Europe Limited; 1994. p. 6.2.1-6.2.20.
36. Boumpas DT, Austin HA, 3rd., Fessler BJ, Balow JE, Klippel JH, Lockshin MD. Systemic lupus erythematosus: emerging concepts. Part 1: Renal, neuropsychiatric, cardiovascular, pulmonary, and hematologic disease. Ann Intern Med 1995;122(12):940-950.
37. Karassa FB, Trikalinos TA, Ioannidis JPA. The Fc[ggr]RIIIA-F158 allele is a risk factor for the development of lupus nephritis: A meta-analysis. Kidney Int 2003;63(4):1475-1482.
38. Morel L. Genetics of human lupus nephritis. Semin Nephrol 2007;27(1):2-11.
39. Okamoto H, Kaneko H, Terai C, Kamatani N. Protective effect of A at position ‚Äì168 in the type III promoter of the MHCIITA gene in systemic lupus erythematosus. Annals of the Rheumatic Diseases 2007;66(9):1263-1264.
40. Linga-Reddy MVP, Gunnarsson I, Svenungsson E, Sturfelt G, Jonsen A, Truedsson L, et al. A polymorphic variant in the MHC2TA gene is not associated with systemic lupus erythematosus. Tissue Antigens 2007;70(5):412-414.
41. Sanchez E, Sabio JM, Jimenez-Alonso J, Callejas JL, Camps M, de Ramon E, et al. Study of two polymorphisms of the MHC2TA gene with systemic lupus erythematosus. Rheumatology 2008;47(1):102-103.
42. Bronson PG, Caillier S, Ramsay PP, McCauley JL, Zuvich RL, De Jager PL, et al. CIITA variation in the presence of HLA-DRB1*1501 increases risk for multiple sclerosis. Hum Mol Genet 2010 0:ddq101v2.

Table 1. Characteristics of SLE cases analyzed in this study (n = 1,463).

		SLE
Site for trio families (n = 637), n (%)		
	UCSF	264 (41.4)
	OMRF	148 (23.2)
	UMN	225 (35.3)
Site for independent cases (n = 826), n (%)		
	UCSF	563 (68.2)
	OMRF	241 (29.2)
	UMN	22 (2.7)
Age, mean (standard deviation)		52.8 (13.4)
Age range		15-94
Female, n (%)		1,352 (92.4)
Age at SLE diagnosis, mean (standard deviation)		33.1 (13.1)
Double strand DNA autoantibody production, n (%)		604 (51.2)[1]
Lupus nephritis, n (%)		205 (26.1)[2]
Arthritis, n (%)		850 (71.0)[3]
Serositis, n (%)		398 (33.2)[4]
History of neurological involvement, n (%)		115 (9.6)[5]
Sm autoantibody production, n (%)		125 (11.7)[6]
Ro autoantibody production, n (%)		265 (26.4)[7]

[1]Data available for 1,179 cases.
[2]Data available for 785 cases.
[3]Data available for 1,197 cases.
[4]Data available for 1,199 cases.
[5]Data available for 1,198 cases.
[6]Data available for 1,068 cases.
[7]Data available for 1,003 cases.

Figure 1. R^2 plot illustrating the LD structure in SLE patients (n = 1,463) for the 18 *CIITA* SNP variants analyzed in this study, where darker gray indicates higher r^2 between pairs of SNPs.

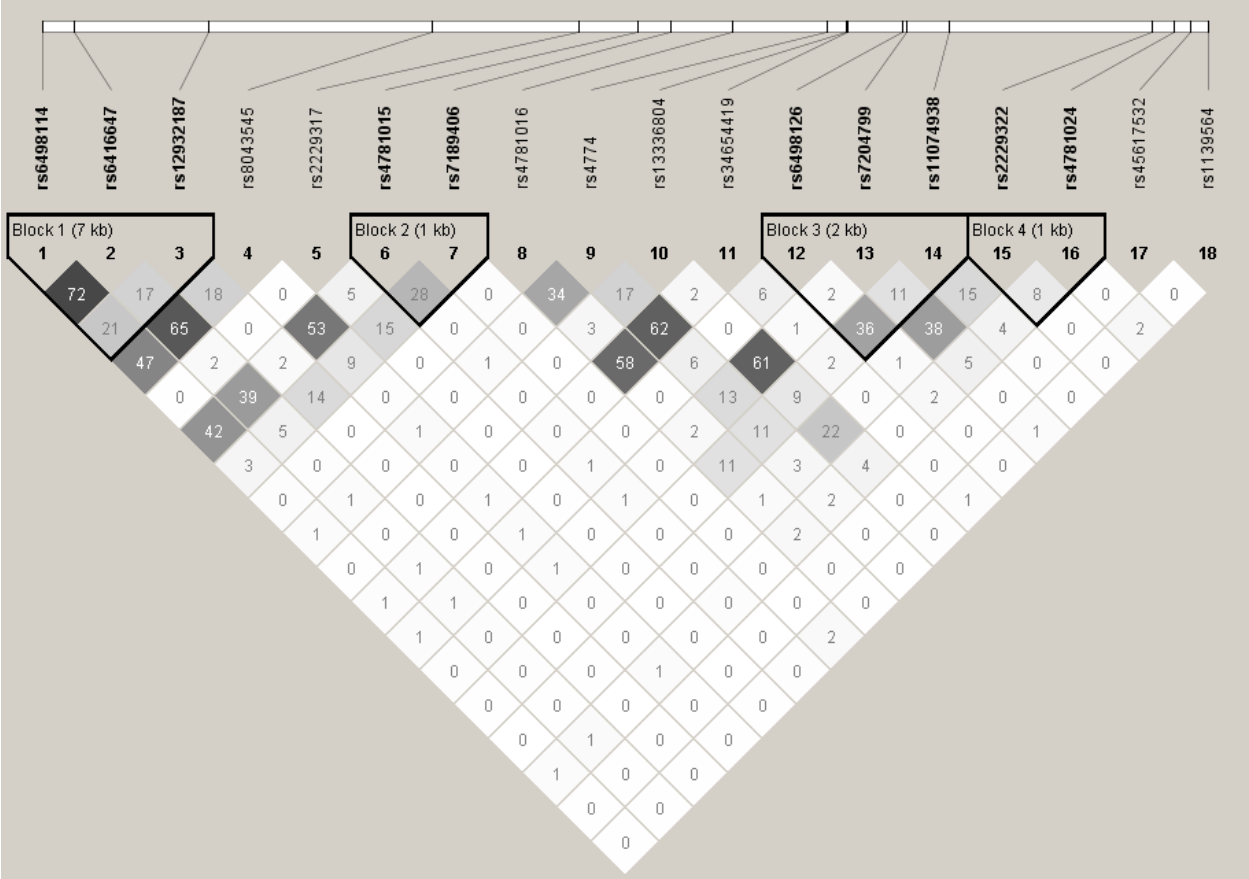
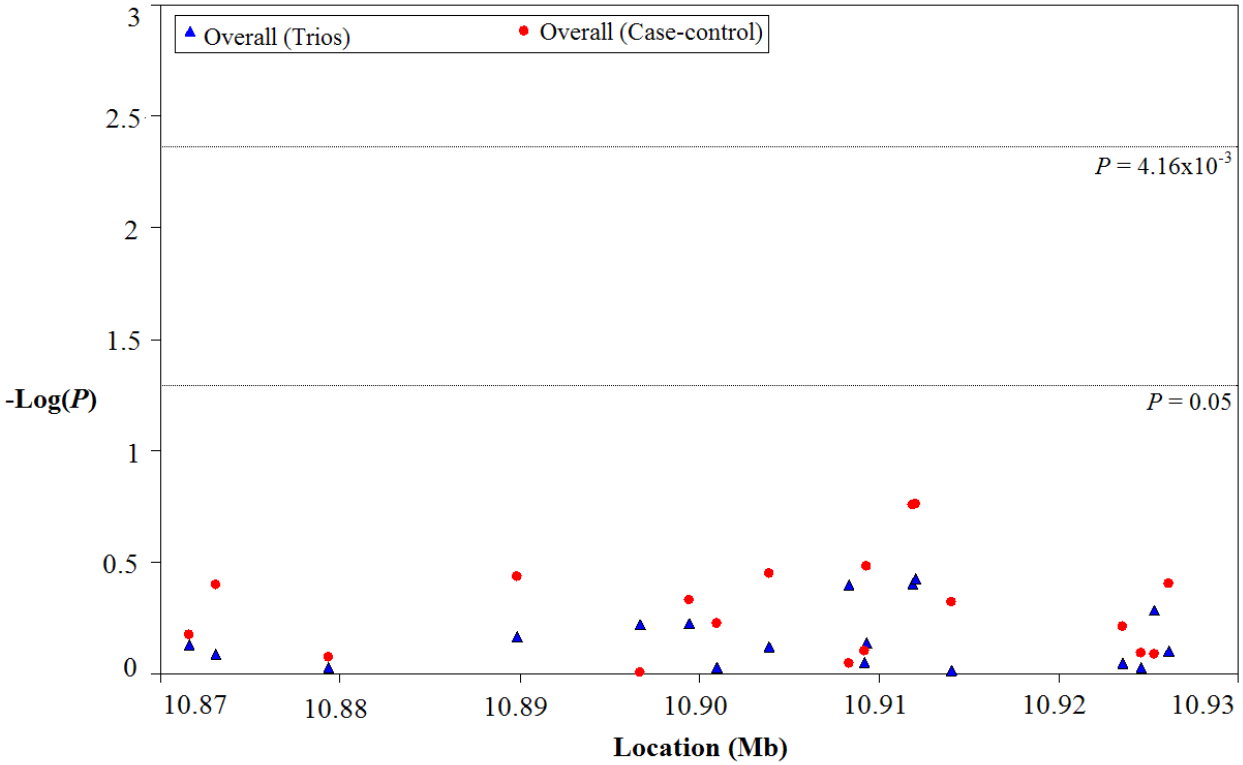


Figure 2. *P*-values from (**a**) overall and (**b**) case-only tests of *CIITA* SNPs in SLE.

a.



b.

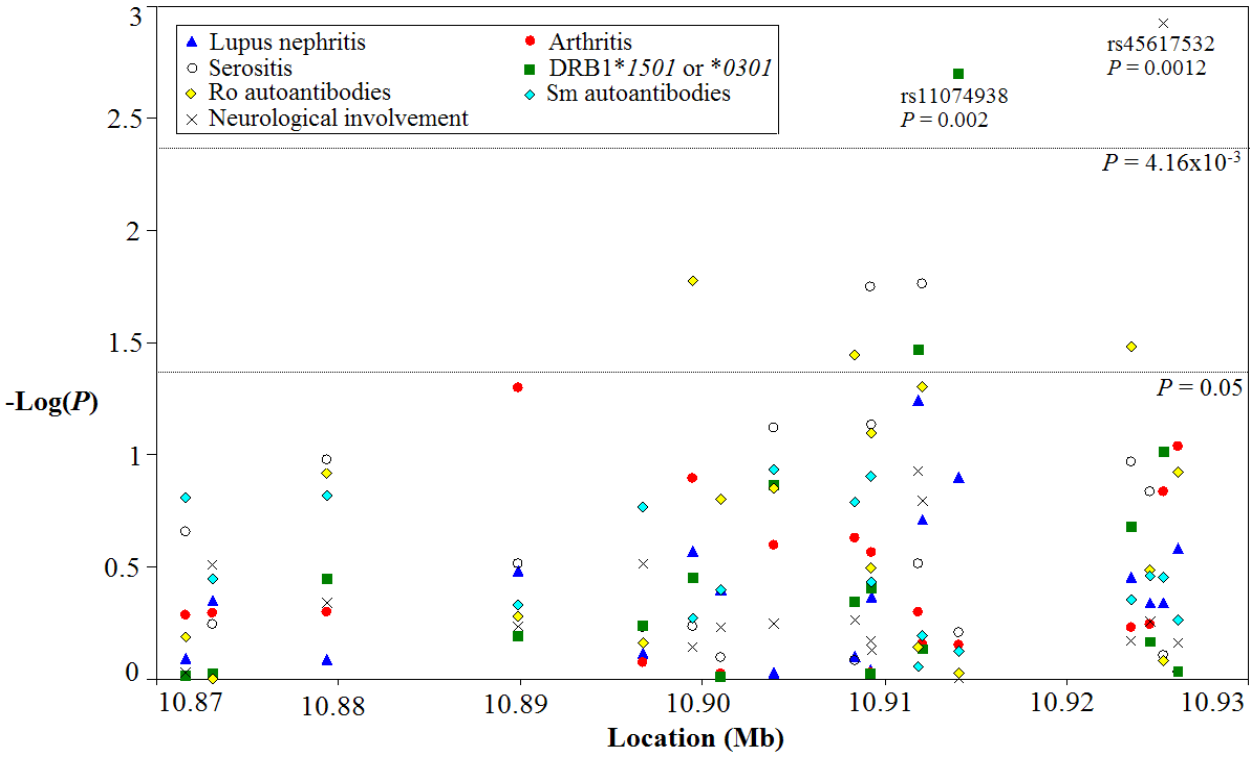


Table 2. Description of 18 *CIITA* SNP variants included in association tests of 637 SLE trio families and 826 unrelated SLE cases.

Marker	Position	Alleles	MAF SLE	Type
rs6498114	10964118	G/T	0.25	Promoter region
rs6416647	10965597	C/T	0.29	Promoter region
rs12932187	10971880	G/C	0.07	Intronic
rs8043545	10982345	C/G	0.26	Intronic
rs2229317	10989219	G/C	0.01	Missense mutation
rs4781015	10991952	A/G	0.20	Intronic
rs7189406	10993488	G/A	0.08	Intronic
rs4781016	10996399	A/C	0.28	Intronic
rs4774	11000848	C/G	0.31	Missense mutation
rs13336804	11001694	C/T	0.09	Missense mutation
rs34654419	11001770	A/C	0.22	Synonymous coding
rs6498126	11004363	G/C	0.20	Intronic
rs7204799	11004549	C/G	0.07	Intronic
rs11074938	11006543	A/G	0.39	Intronic
rs2229322	11016045	T/C	0.10	Synonymous coding
rs4781024	11017058	A/G	0.42	Intronic
rs45617532	11017815	C/T	0.01	Untranslated region
rs1139564	11018622	T/C	0.18	Untranslated region

Table 3. This study captured 11 additional common *CIITA* HapMap SNP variants and did not capture the 15 remaining common *CIITA* HapMap SNP variants, based on HapMap CEU ($r^2 \geq 0.8$).

Marker	Position	Captured	Type
rs12596540	10970476	Yes	Promoter region
rs12928665	10971474	Yes	Intronic
rs4781011	10975311	Yes	Intronic
rs11074934	10979440	Yes	Intronic
rs11647384	10997289	Yes	Intronic
rs4780333	10998482	Yes	Intronic
rs2229321	11001914	Yes	Synonymous coding
rs4780334	11002626	Yes	Intronic
rs7196089	11003101	Yes	Intronic
rs6498132	11010644	Yes	Intronic
rs9788916	11011728	Yes	Intronic
rs9302456	10968472	No	Promoter region
rs6498119	10991878	No	Intronic
rs8048002	10991988	No	Intronic
rs6498122	10994182	No	Intronic
rs8046121	10995933	No	Missense mutation
rs4781019	11004150	No	Intronic
rs4780335	11004328	No	Intronic
rs12598246	11004732	No	Intronic
rs4781020	11008262	No	Intronic
rs6498130	11010150	No	Intronic
rs6498131	11010626	No	Intronic
rs11074939	11011709	No	Intronic
rs7404786	11012550	No	Intronic
rs8056269	11012567	No	Intronic
rs2229322	11016045	No	Synonymous coding

CHAPTER FIVE

CIITA Variation in the Presence of *HLA-DRB1*1501* Increases Risk for Multiple Sclerosis

*Published: Bronson PG, Caillier S, Ramsay PP, McCauley JL, Zuvich RL, De Jager PL, Rioux JD, Ivinson AJ, Compston A, Hafler DA, Sawcer SJ, Pericak-Vance MA, Haines JL, The International Multiple Sclerosis Genetics Consortium, Hauser SL, Oksenberg JR, Barcellos LF. (2010) CIITA variation in the presence of HLA-DRB1*1501 increases risk for multiple sclerosis. Hum Molec Genet 19(11): 2331-40.*

ABSTRACT

The MHC class II transactivator gene (*CIITA*) is an important transcription factor regulating gene required for HLA class II MHC-restricted antigen presentation. Association with HLA class II variation, particularly *HLA-DRB1*1501*, has been well-established for multiple sclerosis (MS). In addition, the -168A/G *CIITA* promoter variant (rs3087456) has been reported to be associated with MS. Thus, a multi-stage investigation of variation within *CIITA, DRB1*1501* and MS was undertaken in 6,108 individuals. In stage 1, 24 SNPs within *CIITA* were genotyped in 1,320 cases and 1,363 controls (n = 2,683). Rs4774 (missense +1614G/C; G500A) was associated with MS ($P = 4.9 \times 10^{-3}$), particularly in *DRB1*1501*+ individuals ($P = 1 \times 10^{-4}$). No association was observed for the -168A/G promoter variant. In stage 2, rs4774 was genotyped in 973 extended families; rs4774**C* was also associated with increased risk for MS in *DRB1*1501*+ families ($P = 2.3 \times 10^{-2}$). In a third analysis, rs4774 was tested in cases and controls (stage 1) combined with one case per family (stage 2) for increased power. Rs4774**C* was associated with MS ($P = 1 \times 10^{-3}$), particularly in *DRB1*1501*+ cases and controls ($P = 1 \times 10^{-4}$). Results obtained from logistic regression analysis showed evidence for interaction between rs4774**C* and *DRB1*1501* associated with risk for MS (ratio of ORs=1.72, 95% CI 1.28 to 2.32, $P = 3 \times 10^{-4}$). Furthermore, rs4774**C* was associated with *DRB1*1501*+ MS when conditioned on the presence (OR = 1.67, 95% CI = 1.19 to 2.37, $P = 1.9 \times 10^{-3}$) and absence (OR = 1.49, 95% CI = 1.15 to 1.95, $P = 2.3 \times 10^{-3}$) of *CLEC16A* rs6498169**G*, a putative MS risk allele adjacent to *CIITA*. Our results provide strong evidence supporting a role for *CIITA* variation in MS risk, which appears to depend on the presence of *DRB1*1501*.

BACKGROUND

Multiple sclerosis (MS) is a chronic inflammatory disorder of the central nervous system characterized by demyelination, astrogliosis, varying degrees of axonal pathology and a relapsing or progressive course (1). A strong but complex genetic component in MS pathogenesis is indicated by both an increased relative risk in non-twin siblings compared with the general population, and by an increased concordance rate in monozygotic compared with dizygotic twins (25% vs. 5%) (2-4). The strongest and most consistent evidence for a susceptibility gene in MS is within the major histocompatibility complex (MHC) on chromosome 6p21.3. Associations with the human leukocyte antigen (*HLA)-DR15* haplotype (*DRB1*1501-DQA1*0102-DQB1*0602*) have been repeatedly demonstrated in multiple populations, primarily in those of Northern European descent (5, 6). Haplotype analysis of *HLA* class II genes in admixed African Americans has demonstrated *HLA-DRB1*15* is the primary susceptibility allele for MS (7). However, a complex pattern of allelic heterogeneity at the *DRB1* locus in MS highlights the intricate nature of this genetic association (8). Transgenic animal models of autoimmune demyelination confirm the critical role of *DRB1* gene products in initiating and maintaining a damaging anti-myelin immune response and suggest *DQB1*0602* is associated with anti-myelin autoimmunity (9, 10).

The identification of all non-MHC genetic risk factors in MS, while progressing steadily over the past couple of years, is far from complete. Several whole genome linkage screens in MS previously identified a large number of regions of modest contributions with little overlap, underscoring a complex polygenic pattern of inheritance contributing to disease susceptibility (11-14). Recent genome-wide association (GWA) and replication studies have identified several genetic risk loci for MS, including *IL2RA*, *IL7RA*, *CLEC16A*, *CD58, TNFRSFA1* and *IRF8* (15-20), though each contributes very modestly to the overall genetic risk for disease. Therefore, a substantial component of the genetic susceptibility to MS remains unknown. While GWA studies are attractive for many reasons, including that in principle they are 'hypothesis free', it is clear that experiments using current technology will be limited in their ability to identify the entire genetic contribution for most complex diseases, including MS (21). Candidate gene studies have historically failed to identify susceptibility loci with conclusive evidence. However, revisiting candidate gene studies with well-powered datasets and strong hypotheses based on prior research remains an important strategy for disease gene identification.

The MHC class II transactivator gene (*CIITA,* also called *MHC2TA*) encodes the CIITA protein, a transcription factor essential for the expression of HLA class II molecules and involved in the expression of HLA class I molecules (16, 18, 22-28). *CIITA* spans 48 kb on chromosome 16p13 and has four alternate first exons in a 12kb promoter region (I-IV) (29). The gene is adjacent to a recently identified MS risk locus (C-type lectin domain family 16, member A gene, *CLEC16A,* previously called *KIAA0350*) (16, 18). The CIITA protein contains a highly conserved and relatively rare domain also encoded by the neuronal apoptosis inhibitory protein gene (*NAIP*, also called *BIRC1*), associated with spinal muscular atrophy, and the nucleotide-binding oligomerization domain protein 2 gene (*NOD2*), associated with Crohn's disease (30-33). Mutations in *CIITA* cause a rare and severe immunodeficiency characterized by HLA class II deficiency (bare lymphocyte syndrome) (34). Thus, *CIITA* is an attractive candidate for genetic studies of autoimmune diseases and other inflammatory conditions for which HLA associations

have been well established. In this multi-stage study, we investigated whether genetic variation in *CIITA* is associated with MS risk and assessed effect modification by *DRB1*1501*.

METHODS

Subjects

MS cases and controls (first stage) were collected by the Brigham and Women's Hospital in Boston (BWH), School of Clinical Medicine in Cambridge (CSU), University of California in San Francisco (UCSF) and Wellcome Trust Case-Control Consortium (Table 1). The second stage included 973 extended MS families of self-identified European ancestry collected by BWH, CSU and UCSF (total n = 978 offspring MS cases, 32 parental MS cases, 2,917 individuals) (Table 1). 518 trios from a GWA study by the International Multiple Sclerosis Genetics Consortium (IMSGC) were part of this family sample (16). There was no overlap between MS cases or other individuals in the two datasets. All MS cases in this study met the revised McDonald MS criteria (35). A total of 5,600 genotyped individuals were studied. Based on the available genetic ancestry data for all cases and controls, and to apply the most stringent criteria possible for genetic analysis of *CIITA*, only MS cases and controls with ≥90% European ancestry were analyzed (stage 1). These steps reduced the potential impact of population stratification on our investigation.

Genotyping

Initially, 29 *CIITA* single nucleotide polymorphisms (SNPs) and the *CLEC16A* rs6498169 MS risk variant were genotyped in the case-control sample with a custom Illumina iSelect 48K chip (San Diego, CA). Four monomorphic missense variants (rs34648899, rs35451230, rs4781022, rs8046121) and one missense variant (rs7197779) with a low minor allele frequency (MAF) (<0.001) were omitted from further analyses. Hardy-Weinberg equilibrium (HWE) for all marker genotypes was examined in cases and controls separately with the exact test (PLINK v1.07) (36, 37). There was no evidence of deviation from HWE in the cases or controls ($P<1x10^{-4}$). Ultimately, 24 *CIITA* variants were tested in the first stage (Table 2). Rs4774 was also genotyped in an independent extended family sample (n = 455 families) using an Applied Biosystems TaqMan assay (Foster City, CA). There were no Mendelian errors (PedCheck v1.1) and the exact test showed no deviation from HWE in pedigree founders (PEDSTATS v0.6.8) (38, 39). In addition, genotyping data for rs4774 and *CLEC16A* rs6498169 were available for 518 trio families (*DRB1*1501*+ n = 223) from a GWA study that used the Affymetrix GeneChip Human Mapping 500K array (16). *DRB1* or rs3135388 genotype data were available for all cases and controls in this study. The *DRB1* locus was characterized as previously described (8, 16, 40).

Statistical analyses

European ancestry was estimated from 1,000 markers included on the Illumina iSelect 48K chip using a Bayesian clustering algorithm based on two populations under the admixture model using a burn-in length of 10,000 for 10,000 repetitions (STRUCTURE v2.3.1) (41). European ancestry estimates were tested for association with MS in the combined cases and controls (all individuals from stage 1) and subgroups stratified by *DRB1*1501* status with logistic regression (R v2.9; http://www.r-project.org). European ancestry estimates were statistically indistinguishable when compared between stage 1 MS cases (mean = 0.99, standard

deviation [SD] = 0.018) and controls (mean = 0.99, SD = 0.017) (asymptotic $P = 0.20$), even in the *DRB1*1501*+ ($P = 0.71$) and *DRB1*1501*- subsets ($P = 0.27$).

Association was tested in the case-control sample with a Cochran-Armitage test for trend (degrees of freedom [*df*]=1) (PLINK) (42, 43). Trend tests for *CIITA* SNPs were stratified by the presence or absence of *DRB1*1501* in the cases and controls. To address multiple testing concerns for *CIITA*, we determined significance in stage 1 ($P \leq 4.9 \times 10^{-3}$) by controlling the False Discovery Rate (FDR) using the Benjamini-Hochberg method at a level of 10% (44). Allelic ORs and 95% CIs are also reported. *DRB1*1501* was also tested with the Cochran-Armitage test. Haplotypes were estimated in the controls with the expectation-maximization algorithm and haplotype blocks were determined with the confidence bound algorithm (HAPLOVIEW v4.1) (Figure 1) (45). We computed maximum likelihood estimates of haplotype probabilities for the MS cases and controls with the expectation-maximization algorithm and conducted global haplotype tests of 6 haplotype blocks encompassing 16 SNPs using score statistics under the additive genetic model (HAPLOSTATS v1.4.3, R) (46). Haplotypes with inferred frequencies <5% were excluded. Analyses were stratified by *DRB1*1501* status. ORs and 95% CIs are reported.

Association between rs4774 and MS was tested in extended families with the pedigree disequilibrium test (PDT v6.2.4), an extension of the transmission disequilibrium test (TDT) (47, 48). We chose the PDT to take advantage of the data available for families with more than one affected offspring and families with one or more unaffected siblings. 892 families (674 affected offspring trios and 489 discordant sibling pairs [DSPs]) were informative for the PDT, and a one-tailed asymptotic *P*-value was calculated. 497 *DRB1*1501*+ families (358 affected offspring trios and 288 DSPs) were informative for the PDT. *DRB1*1501* was also tested with the PDT and the informative sample consisted of 848 families with 630 affected offspring trios and 488 DSPs (*data not shown*).

A combined analysis (stage 3) of cases and controls (from stage 1) and one proband per family (from stage 2) was conducted using a 1-*df* Cochran-Armitage test for trend (PLINK). Rs4774 allele frequencies between the cases from stages 1 and 2, as well as between the controls from stage 1 and the affected family-based controls (AFBACs, or nontransmitted parental alleles) from stage 2, were calculated with AFBAC v1.13 (49). Frequencies were tested with chi-square tests of heterogeneity ($df = 1$) in R and asymptotic *P*-values were calculated. Rs4774 frequencies in controls from stage 1 (MAF 0.268) and the AFBACs from stage 2 (MAF 0.311) were different ($P = 2.5 \times 10^{-4}$). Therefore, we chose not to utilize the AFBACs as controls in the combined analysis (stage 3). The AFBACs were utilized only in family-based analyses robust against effects of population stratification.

We also tested rs4774 for association in the combined datasets from stages 1 and 2 with a TDT that incorporates the sibling TDT and parental phenotypes and tests unrelated cases and controls as sibships with the Cochran-Mantel-Haenszel test for stratified tables (PLINK) (42, 47, 50-52). We considered conducting a conditional logistic regression analysis of one proband per family and three "pseudo-controls" based on nontransmitted parental alleles, along with the cases and controls (53-55). However, this method would have substantially reduced power because a

third of the families (n = 301) did not have rs4774 genotypes available for both of the proband's parents.

Interaction between the presence of rs4774*C and *DRB1*1501* (*gene x gene*) in the combined cases and controls (stage 3) was tested using logistic regression (R); we report these results as the ratio of ORs (ROR_{GxG}) and include 95% CIs. A case-only interaction test was also conducted by using logistic regression (R) to test for association between the rs4774 genotype and the presence of *DRB1*1501* (56, 57). The conditional haplotype method (58) (HAPLOVIEW, R) was used to test for association between rs4774 and MS, conditional on the presence or absence of the *CLEC16A* rs6498169*G MS risk allele; asymptotic P-values were calculated. The presence of rs4774*C was tested for association with MS in a logistic regression (R), adjusted for presence/absence of *DRB1*1501* and the presence of the *CLEC16A* rs6498169*G MS risk allele.

All reported P-values are empirically based on ≥10,000 permutations and are two-tailed, unless otherwise noted in the methods. Power was estimated (PGA v2.0) assuming a two-sided type I error of $\alpha=4.9\text{x}10^{-3}$, to account for number of statistical tests (59). Stage one of the current study had 80% power to detect an OR ranging from 1.22 to 1.38, under varying MAF (0.1-0.5).

A meta-analysis of the -168A/G variant and MS was performed for 3,322 cases and 4,260 controls in stage 3 plus four additional case-control studies (60-63). ORs and 95% CIs were calculated to test whether the G allele (*AA* compared against *GA+GG* carriers [dominant model]) or the *GG* genotype (*AA+AG* compared against *GG* carriers [recessive model]) increased risk for MS. A meta-analysis of the rs4774 variant and MS was also performed for 2,669 cases and 3,773 controls in stage 3 plus three additional case-control studies (60, 62, 63). ORs and 95% CIs were calculated to test whether the C allele (*GG* compared against *GC+CC* carriers) or the *CC* genotype (*GG+CG* compared against *CC* carriers) increased risk for MS. Results did not differ under the recessive models (*data not shown*). Between-study heterogeneity was assessed with a X^2 test of heterogeneity and publication bias was evaluated with a funnel plot. We calculated summary ORs and 95% CIs using a random effects model and asymptotic P-values, as previously described (64, 65).

RESULTS

In stage 1, 24 *CIITA* SNPs were genotyped and analyzed in a European dataset of 1,320 MS cases and 1,363 independent healthy controls (total n = 2,683) (Tables 1 and 2). In the second stage, we genotyped and analyzed rs4774 in an independent dataset of 973 MS families of self-identified European ancestry (total n = 978 offspring MS cases, 2,917 individuals) (see Table 1) (16). *DRB1*1501* was strongly associated with MS in case-control (odds ratio [OR]=2.71, 95% confidence interval [95% CI] = 2.36 to 3.11, $P = 1\text{x}10^{-6}$) and family ($P < 1\text{x}10^{-6}$) datasets, as expected (*data not shown;* see Methods).

Of the 24 *CIITA* SNPs tested, evidence for association with MS was observed only for rs4774 ($X^2 = 8.14$, $P = 4.9\text{x}10^{-3}$) after application of a conservative correction for number of statistical tests (Methods, Table 3). Furthermore, rs4774 and MS were more strongly associated in individuals carrying *DRB1*1501* ($X^2 = 18.8$, $P = 1\text{x}10^{-4}$) (Table 3). In both comparisons

(overall and stratified by presence of *DRB1*1501*), the minor rs4774**C* allele frequency was increased in MS cases vs. controls (overall: OR = 1.19, 95% CI = 1.06 to 1.34, *DRB1*1501*+: OR = 1.6, 95% CI = 1.3 to 1.97) (Table 4). This result exceeded our threshold for statistical significance. No evidence for association was observed for the previously reported -168A/G *CIITA* promoter polymorphism (rs3087456). *CIITA* SNP haplotypes were assigned based on block structure (Methods) and also compared between MS cases and controls. A total of six blocks were observed (Figure 1). Results did not indicate the presence of any other *CIITA* SNP associations stronger than rs4774. Interestingly, rs4774 did not fall within surrounding SNP blocks, and LD with neighboring SNP variant rs3087456 was not present ($r^2 = 0.01$). No evidence for any global haplotype associations between *CIITA* and MS were observed (Table 5). The family-based analysis of rs4774**C* in 539 *DRB1*1501*+ MS extended families also showed evidence of association with increased MS risk, albeit weak ($P = 2.3 \times 10^{-2}$) (Table 6). MS families were stratified based on *DRB1*1501* status in the proband.

Rs4774 frequencies did not differ between MS cases from the case-control (MAF 0.304) and family-based (MAF 0.306) datasets utilized in this study ($P = 0.83$). Therefore, a pooled analysis, whereby both MS case groups were combined (from stages 1 and 2) and compared to controls (from stage 1), was performed for increased statistical power (stage 3). One MS case per family was selected. A total of 3,656 individuals (n = 2,293 cases and 1,363 controls) were included in the analysis. Rs4774**C* was associated with increased MS risk ($X^2 = 10.9$, $P = 1 \times 10^{-3}$), and this association was stronger in *DRB1*1501*+ cases and controls ($X^2 = 18.9$, $P = 1 \times 10^{-4}$) (Table 7). We also conducted a combined analysis that maintained the robustness of the family-based component against potential confounding due to population stratification and controlled for potential differences between study populations in stages 1 and 2. Specifically, we used the sibling TDT to analyze the extended families and the Cochran-Mantel-Haenszel test for stratified tables to incorporate the independent cases and controls. Results supported those obtained from the pooled analyses. Rs4774**C* was associated with increased risk of MS in the overall sample ($X^2 = 5.4$ [1,249 expected, 1,298 observed], $P = 3.2 \times 10^{-2}$), and this association was stronger in *DRB1*1501*+ individuals ($X^2 = 17.4$ [687.8 expected, 746 observed], $P = 3 \times 10^{-4}$). No association was observed in *DRB1*1501*- individuals ($X^2 = 1.1$ [568.2 expected, 552 observed], $P = 0.33$).

Results obtained from a logistic regression analysis of MS cases and controls demonstrated evidence of interaction between *DRB1*1501* and rs4774**C* (ratio of ORs [ROR_{GxG}] = 1.72, 95% CI = 1.28 to 2.32, $P = 3 \times 10^{-4}$) (Table 7). Furthermore, the rs4774 genotype was associated with the presence of *DRB1*1501* in a case-only analysis (OR = 1.19, 95% CI = 1.05 to 1.36, $P = 6.8 \times 10^{-3}$), adding further evidence for interaction. Also, the presence of both *DRB1*1501* and rs4774**C*, when compared to the presence of *DRB1*1501* alone, was associated with MS (OR = 1.79, 95% CI = 1.39 to 2.30, $P = 2.6 \times 10^{-6}$) (Table 7).

To determine whether the *CIITA* rs4774 association observed in the current study was independent from the previously reported MS association with nearby *CLEC16A* rs6498169 (16, 18), conditional haplotype analysis using the larger stage 3 dataset and genotypes for both loci was performed. Here, rs4774**C* was associated with *DRB1*1501*+ MS when conditioned on both presence (OR = 1.67, 95% CI = 1.19-2.37, $P = 1.9 \times 10^{-3}$) and absence (OR = 1.49, 95% CI = 1.15-1.95, $P = 2.3 \times 10^{-3}$) of the *CLEC16A* rs6498169**G* MS risk allele (Table 8 for haplotype frequencies). *CLEC16A* rs6498169**G* was associated with MS when conditioned on the presence

(OR = 1.32, 95% CI = 1.09 to 1.61, $P = 4.9x10^{-3}$) and absence (OR = 1.19, 95% CI = 1.05-1.35, $P = 6.6x10^{-3}$) of rs4774**C*, and this association trended toward significance in the *DRB1*1501* stratified subsets (*DRB1*1501*+/rs4774**C*+, OR = 1.38, 95% CI = 0.96 to 2.01, $P = 0.08$; *DRB1*1501*+/rs4774**C*-, OR = 1.23, 95% CI = 0.90 to 1.54, $P = 0.06$; *DRB1*1501*-/rs4774*C+, OR = 1.33, 95% CI = 1.03 to 1.72, $P = 2.6x10^{-2}$; *DRB1*1501*-/rs4774**C*-, OR = 1.16, 95% CI = 0.98 to 1.37, $P = 0.07$). These results indicate that the association between *CIITA* and MS is independent of *CLEC16A*. Furthermore, in a logistic regression, rs4774**C* demonstrated association with MS (OR = 1.24, 95% CI 1.07 to 1.44, $P = 6x10^{-3}$) even after adjusting for the presence of *DRB1*1501* (OR = 3.33, 95% CI 2.85 to 3.89, $P < 2x10^{-6}$) and *CLEC16A* rs6498169**G* (OR = 1.24, 95% CI 1.06 to 1.44, $P = 4.9x10^{-3}$) (*data not shown*).

DISCUSSION

Due to the strong association between *HLA-DRB1*1501* and MS, and the influence of *CIITA* on the expression of HLA class II genes, the *CIITA* locus has long been considered a strong MS candidate gene. Almost a decade ago, Rasmussen et al. screened for variants in 111 MS cases and 105 controls from the U.K., and through sequencing, identified the -168A/G variant in the type III *CIITA* promoter region (rs3087456), as well as five variants in the 3' untranslated region (UTR) (66). Association between MS and *CIITA* in the overall or *DR15*-stratified sample was not detected; however, an association between the -168A/G variant and primary progressive MS ($P<0.04$) was reported. Shortly thereafter, Patarroyo et al. identified four additional *CIITA* SNPs, including the rs4774 (+1614G/C) missense mutation in exon 11 (discovered by Steimle et al.), through bidirectional sequencing of lymphocyte cDNAs from 50 healthy individuals of Northern European ancestry (34, 67). A more recent study reported association between the -168A/G variant and increased susceptibility to both MS and rheumatoid arthritis (RA), as well as lower expression of *CIITA* after stimulation of leukocytes with interferon γ (60). In addition, ex vivo stimulation with interferon-γ of peripheral blood cells from RA cases with the -168G/G genotype exhibited decreased expression of *CIITA* and the HLA class II alleles DQA1 and DRA, compared to RA patients with the -168A/G or -168A/A genotype, and the difference was greater with increased stimulation (60). However, a meta-analysis of ten studies revealed no evidence for association between the -168A/G variant and RA (64).

Thus far, studies of the *CIITA* -168A/G variant and MS have yielded conflicting results (60-63). No evidence for association between the -168A/G variant and MS was observed in the current study, despite achieving 80% power to detect an allelic OR of 1.25 (MAF 0.27) under a two-sided $\alpha = 4.9x10^{-3}$, or even an allelic OR of 1.18 under relaxed significance criteria ($\alpha=0.05$). As part of the current investigation, we also performed a meta-analysis of allelic association between the -168A/G polymorphism and MS in 3,322 cases and 4,260 controls, which were obtained from stage 3 of this study plus four additional published case-control studies (60-63). Between-study heterogeneity and publication bias were not present (*data not shown*). There was no evidence for association (summary OR = 1.06, 95% CI = 0.91 to 1.24, $P = 0.47$). Our results collectively and definitively exclude any major effect of the -168A/G variant on risk for MS (*data not shown*).

This study is the first to report evidence for interaction between the rs4774 (+1614G/C) missense mutation and *DRB1*1501* associated with MS. This variant, located in exon 12, causes an amino acid substitution from glycine to alanine. Based on sequence homology and physical properties of amino acids, this amino acid substitution is predicted to have a tolerable effect on protein function (68). However, the exact functional consequences that result are not known. We have replicated this finding in an extended family-based sample; though the effect was in the same direction, the significance of the observed association was modest. A combined analysis of MS cases, controls and one MS case from each family also supports a role for rs4774 in MS susceptibility, particularly in the presence of *DRB1*1501*. Three previous studies have examined association between rs4774 and MS and reported negative findings (60, 62, 63). We performed a meta-analysis of rs4774 and MS in 2,669 cases and 3,773 controls, obtained from stage 3 of this study plus three additional published case-control studies (60, 62, 63). While some evidence for between-study heterogeneity was detected ($P<0.02$), publication bias was not present (*data not shown*). The meta-analysis revealed no evidence for allelic association when all MS cases and controls were considered (summary OR = 1, 95% CI = 0.81 to 1.23, $P = 0.99$). Unfortunately, *DRB1*1501* data were not available for published studies to perform stratified analyses. The association seen in the current study between rs4774 and MS appears to depend on the presence of *DRB1*1501*.

Rs4774**C* has been reported to be over-represented in MS cases with active replication of human herpes virus 6A (HHV-6A), compared to MS cases and controls without active replication of HHV-6A (69, 70), although this association needs to be replicated in an independent sample. Actively replicating HHV-6A was defined as at least one positive serum sample for HHV-6A among the 5 serum samples collected in a two-year period. The importance of this finding is not yet known. Larger studies of MS and *CIITA* that include environmental exposure data are warranted.

Variation within the MHC class II transactivator gene in animal studies (rat *Ciita*, 10q11) has been reported to affect the quantity of MHC class II expression in the brain and on immune cells, as well as risk and severity of experimental allergic encephalomyelitis (EAE) (71). The rat strain most susceptible to neurodegeneration and central nervous system inflammation from experimental nerve injury (dark Agouti [DA] strain) was bred with the rat strain most resistant to experimental nerve injury (the Piebald Virol Glaxo (PVG) strain). Compared to DA rats carrying the DA *Ciita* locus, DA rats with the PGV *Ciita* locus exhibited decreased MHC class II expression in the brain upon stimulation with IFN-γ, decreased MHC class II expression on B cells and dendritic cells, and reduced risk and severity of EAE. In MS, the relationship between variation at the *CIITA* locus and gene expression for both *CIITA* and MHC class II loci, as well as the resulting biological implications for the immune response and MS pathogenesis, are poorly understood. Large and comprehensive studies, particularly ones that can also fully explore clinical MS phenotypes, are needed.

We carefully considered the potential impact of population stratification on the current study (72). In stage 1, European ancestry was estimated in MS cases and controls using genetic markers and only individuals with ≥90% European ancestry were included in further analyses of *CIITA* variation. In the second stage, the family-based analysis used to replicate our initial finding was not subject to population stratification. Finally, one MS case per family was selected

and combined with other cases for a larger case-control analysis. Because ancestry informative marker information was not available for all familial MS cases in this final stage, it is possible that ancestral differences in frequencies could have contributed to a spurious association in stage 3. However, the majority of families utilized here were subjected to rigorous testing for population outliers, as previously described (16). Despite these efforts it is possible that within European population stratification may still be present. A conservative correction for multiple testing was also employed to help guide interpretation of testing; however, further replication studies will be required for confirmation.

GWA studies have not identified *CIITA* as a susceptibility locus for MS (16, 73, 74). Further, results for *CIITA* analysis in the current study would not meet criteria for genome-wide significance. Because the recently confirmed *CLEC16A* MS locus is adjacent to *CIITA* on chromosome 16, we examined LD patterns between 14 *CLEC16A* SNPs and 24 *CIITA* SNPs in the controls from stage 1, and 274 *CLEC16A* SNPs and 40 *CIITA* SNPs in HapMap samples of northern and western European origin (CEU, release 24) (75). There was no evidence of LD between *CIITA* and *CLEC16A* in the controls ($r^2 \leq 0.10$). The rs1139564 *CIITA* 3' UTR variant exhibited weak LD with the intronic rs8055876 *CLEC16A* variant in CEU ($r^2 = 0.46$). The rs8055876 *CLEC16A* variant was neither genotyped nor tagged in our controls, but the rs1139564 *CIITA* 3' UTR variant was not in LD with rs4774 in either our controls ($r^2 = 0.001$) or CEU ($r^2 = 0.001$). Also, rs8055876 and the rs6498169 *CLEC16A* MS risk variant were not in LD in CEU ($r^2 = 0.09$). Thus, based on patterns of LD derived from two independent samples and results from our comprehensive analyses, including logistic regression modeling, it does not appear that association observed between *CIITA* and MS, specifically the effect on disease risk conferred by rs4774, is due to *CLEC16A*.

Approximately eighty percent of the common genetic variation in *CIITA* was captured in the current study, based on $r^2 > 0.8$ in CEU. An additional 19 *CIITA* HapMap variants were captured (Table 9). Though this study failed to capture 8 common *CIITA* variants, all of these were intronic, and perhaps less likely to play a role in MS susceptibility (Table 10). In addition, rare variants in *CIITA* were not directly investigated in the current study, and must be considered in future studies.

CONCLUSION

In conclusion, this is the first large study of *CIITA* in MS to fully characterize common genetic variation in *CIITA*, including the assessment of haplotypes and *gene x gene* interaction with *DRB1*1501*. Our results confirm that the previously reported -168A/G promoter variant (rs3087456) is not associated with MS, and provide strong evidence for association between MS and the *CIITA* non-synonymous coding variant (rs4774; missense G/C; G500A) in the presence of *DRB1*1501*. Given the functional relevance of *CIITA*, and the relationship between *CIITA* and the class II *DRB1* locus, our results will help further the understanding of biological mechanisms contributing to MS pathogenesis.

ACKNOWLEDGEMENTS

We would like to thank the study participants and the IMSGC (https://www.imsgc.org/ for list of all members). We would also like to thank Farren B.S. Briggs, Benjamin A. Goldstein and Alan Hubbard for helpful discussions. This work was supported by the National MS Society (RG4201-A-1, RG4201, RG2901), Medical Research Council (G0700061), Cambridge NIHR Biomedical Research Centre, and grants R01 NS049477 (NIH/NINDS), R01 AI059829, F31 AI075609 and U19 AI067152 (NIH/NIAID). The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the NIH or NIAID.

REFERENCES

1. Hauser SL, Goodin DS. Multiple sclerosis and other demyelinating diseases. In: Kasper DL, Braunwald E, Fauci AD, Hauser SL, Longo DL, Jameson JL, editors. Harrison's Principles of Internal Medicine. 16 ed. NY: McGraw Hill; 2005.
2. Willer CJ, Dyment DA, Risch NJ, Sadovnick AD, Ebers GC. Twin concordance and sibling recurrence rates in multiple sclerosis. Proc Natl Acad Sci USA 2003;100(22):12877-12882.
3. Nielsen NM, Westergaard T, Rostgaard K, Frisch M, Hjalgrim H, Wohlfahrt J, et al. Familial risk of multiple sclerosis: a nationwide cohort study. Am J Epidemiol 2005;162(8):774-8.
4. Hemminki K, Li X, Johansson SE, Sundquist K, Sundquist J. Re: "Familial risk of multiple sclerosis: a nationwide cohort study". Am J Epidemiol 2006;163(9):873-4.
5. Barcellos LF, Oksenberg JR, Begovich AB, Martin ER, Schmidt S, Vittinghoff E, et al. HLA-DR2 dose effect on susceptibility to multiple sclerosis and influence on disease course. Am J Hum Genet 2003;72(3):710-6.
6. Lincoln MR, Montpetit A, Cader MZ, Saarela J, Dyment DA, Tiislar M, et al. A predominant role for the HLA class II region in the association of the MHC region with multiple sclerosis. Nat Genet 2005;37(10):1108-12.
7. Oksenberg JR, Barcellos LF, Cree BAC, Baranzini SE, Bugawan TL, Khan O, et al. Mapping multiple sclerosis susceptibility to the HLA-DR locus in African Americans. Am J Hum Genet 2004;74(1):160-167.
8. Barcellos LF, Sawcer S, Ramsay PP, Baranzini SE, Thomson G, Briggs F, et al. Heterogeneity at the HLA-DRB1 locus and risk for multiple sclerosis. Hum Mol Genet 2006;15(18):2813-2824.
9. Fugger L. Human autoimmunity genes in mice. Curr Opin Immunol 2000;12(6):698-703.
10. Kaushansky N, Altmann DM, Ascough S, David CS, Lassmann H, Ben-Nun A. HLA-DQB1*0602 determines disease susceptibility in a new "humanized" multiple sclerosis model in HLA-DR15 (DRB1*1501;DQB1*0602) transgenic mice. J Immunol 2009;183(5):3531-41.
11. Haines JL, Ter-Minassian M, Bazyk A, Gusella JF, Kim DJ, Terwedow H, et al. A complete genomic screen for multiple sclerosis underscores a role for the major histocompatability complex. The Multiple Sclerosis Genetics Group. Nat Genet 1996;13(4):469-71.
12. Sawcer S, Jones HB, Feakes R, Gray J, Smaldon N, Chataway J, et al. A genome screen in multiple sclerosis reveals susceptibility loci on chromosome 6p21 and 17q22. Nat Genet 1996;13(4):464-8.
13. Ebers GC, Kukay K, Bulman DE, Sadovnick AD, Rice G, Anderson C, et al. A full genome search in multiple sclerosis. Nat Genet 1996;13(4):472-6.
14. Sawcer S, Ban M, Maranian M, Yeo TW, Compston A, Kirby A, et al. A high-density screen for linkage in multiple sclerosis. Am J Hum Genet 2005;77(3):454-67.
15. Perera D, Stankovich J, Butzkueven H, Taylor BV, Foote SJ, Kilpatrick TJ, et al. Fine mapping of multiple sclerosis susceptibility genes provides evidence of allelic heterogeneity at the IL2RA locus. J Neuroimmunol 2009;211(1-2):105-9.
16. Hafler DA, Compston A, Sawcer S, Lander ES, Daly MJ, De Jager PL, et al. Risk alleles for multiple sclerosis identified by a genomewide study. N Engl J Med 2007;357(9):851-62.
17. Refining genetic associations in multiple sclerosis. Lancet Neurology 2008;7(7):567-569.

18. Rubio JP, Stankovich J, Field J, Tubridy N, Marriott M, Chapman C, et al. Replication of KIAA0350, IL2RA, RPL5 and CD58 as multiple sclerosis susceptibility genes in Australians. Genes Immun 2008;9(7):624-30.
19. De Jager PL, Baecher-Allan C, Maier LM, Arthur AT, Ottoboni L, Barcellos L, et al. The role of the CD58 locus in multiple sclerosis. Proc Natl Acad Sci USA 2009;106(13):5264-5269.
20. De Jager PL, Jia X, Wang J, de Bakker PIW, Ottoboni L, Aggarwal NT, et al. Meta-analysis of genome scans and replication identify CD6, IRF8 and TNFRSF1A as new multiple sclerosis susceptibility loci. Nat Genet 2009;41(7):776-782.
21. Altshuler D, Daly MJ, Lander ES. Genetic mapping in human disease. Science 2008;322(5903):881-8.
22. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007;447(7145):661-678.
23. Hakonarson H, Grant SFA, Bradfield JP, Marchand L, Kim CE, Glessner JT, et al. A genome-wide association study identifies KIAA0350 as a type 1 diabetes gene. Nature 2007;448(7153):591-594.
24. Todd JA, Walker NM, Cooper JD, Smyth DJ, Downes K, Plagnol V, et al. Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes. Nat Genet 2007;39(7):857-864.
25. Chang CH, Flavell RA. Class II transactivator regulates the expression of multiple genes involved in antigen presentation. J Exp Med 1995;181(2):765-767.
26. Chang CH, Guerder S, Hong SC, van Ewijk W, Flavell RA. Mice lacking the MHC class II transactivator (CIITA) show tissue-specific impairment of MHC class II expression. Immunity 1996;4(2):167-178.
27. Kern I, Steimle V, Siegrist CA, Mach B. The two novel MHC class II transactivators RFX5 and CIITA both control expression of HLA-DM genes. Int Immunol 1995;7(8):1295-1299.
28. LeibundGut-Landmann S, Waldburger JM, Krawczyk M, Otten LA, Suter T, Fontana A, et al. Specificity and expression of CIITA, the master regulator of MHC class II genes. Eur J Immunol 2004;34(6):1513-1525.
29. Muhlethaler-Mottet A, Otten LA, Steimle V, Mach B. Expression of MHC class II molecules in different cellular and functional compartments is controlled by differential usage of multiple promoters of the transactivator CIITA. Embo J 1997;16(10):2851-60.
30. Hughes A. Evolutionary relationships of vertebrate NACHT domain-containing proteins. Immunogenetics 2006;58(10):785-791.
31. Roy N, Mahadevan MS, McLean M, Shutler G, Yaraghi Z, Farahani R, et al. The gene for neuronal apoptosis inhibitory protein is partially deleted in individuals with spinal muscular atrophy. Cell 1995;80(1):167-78.
32. Ogura Y, Bonen DK, Inohara N, Nicolae DL, Chen FF, Ramos R, et al. A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. Nature 2001;411(6837):603-6.
33. Hugot JP, Chamaillard M, Zouali H, Lesage S, Cezard JP, Belaiche J, et al. Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. Nature 2001;411(6837):599-603.
34. Steimle V, Otten LA, Zufferey M, Mach B. Complementation cloning of an MHC class II transactivator mutated in hereditary MHC class II deficiency (or bare lymphocyte syndrome). Cell 1993;75(1):135-46.

35. Polman CH, Reingold SC, Edan G, Filippi M, Hartung HP, Kappos L, et al. Diagnostic criteria for multiple sclerosis: 2005 revisions to the "McDonald Criteria". Ann Neurol 2005;58(6):840-6.
36. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81(3):559-75.
37. Wigginton JE, Cutler DJ, Abecasis GR. A note on exact tests of Hardy-Weinberg equilibrium. Am J Hum Genet 2005;76(5):887-93.
38. O'Connell JR, Weeks DE. PedCheck: a program for identification of genotype incompatibilities in linkage analysis. Am J Hum Genet 1998;63(1):259-66.
39. Wigginton JE, Abecasis GR. PEDSTATS: descriptive statistics, graphics and quality assessment for gene mapping data. Bioinformatics 2005;21(16):3445-7.
40. Caillier SJ, Briggs F, Cree BA, Baranzini SE, Fernandez-Vina M, Ramsay PP, et al. Uncoupling the roles of HLA-DRB1 and HLA-DRB5 genes in multiple sclerosis. J Immunol 2008;181(8):5473-80.
41. Falush D, Stephens M, Pritchard JK. Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. Genetics 2003;164(4):1567-1587.
42. Cochran WG. Some methods for strengthening the comon X2 tests. Biometrics 1954;10(4):417-451.
43. Armitage P. Tests for linear trends in proportions and frequencies. Biometrics 1955;11(3):375-386.
44. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B-Methodological 1995;57(1):289-300.
45. Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, et al. The structure of haplotype blocks in the human genome. Science 2002;296(5576):2225-9.
46. Schaid DJ, Rowland CM, Tines DE, Jacobson RM, Poland GA. Score tests for association between traits and haplotypes when linkage phase is ambiguous. Am J Hum Genet 2002;70(2):425-34.
47. Spielman RS, McGinnis RE, Ewens WJ. Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). Am J Hum Genet 1993;52:506-516.
48. Martin ER, Monks SA, Warren LL, Kaplan NL. A test for linkage and association in general pedigrees: the pedigree disequilibrium test. Am J Hum Genet 2000;67(1):146-54.
49. Thomson G. Mapping disease genes: family-based association studies. Am J Hum Genet 1995;57(2):487-98.
50. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. J Natl Cancer Inst 1959;22(4):719-48.
51. Spielman RS, Ewens WJ. A sibship test for linkage in the presence of association: the sib transmission/disequilibrium test. Am J Hum Genet 1998;62(2):450-8.
52. Purcell S, Sham P, Daly MJ. Parental phenotypes in family-based association analysis. Am J Hum Genet 2005;76(2):249-59.
53. Self SG, Longton G, Kopecky KJ, Liang KY. On estimating HLA/disease association with application to a study of aplastic anemia. Biometrics 1991;47(1):53-61.
54. Schaid DJ. General score tests for associations of genetic markers with disease using cases and their parents. Genet Epidemiol 1996;13(5):423-49.

55. Cordell HJ, Clayton DG. A unified stepwise regression procedure for evaluating the relative effects of polymorphisms within a gene using case/control or family data: application to HLA in type 1 diabetes. Am J Hum Genet 2002;70(1):124-41.
56. Yang Q, Khoury MJ, Sun F, Flanders WD. Case-Only Design to Measure Gene-Gene Interaction. Epidemiology 1999;10(2):167-170.
57. Clarke GM, Pettersson FH, Morris AP. A comparison of case-only designs for detecting gene x gene interaction in rheumatoid arthritis using genome-wide case-control data in Genetic Analysis Workshop 16. BMC Proc 2009;3 Suppl 7:S73.
58. Thomson G, Valdes AM. Conditional genotype analysis: detecting secondary disease loci in linkage disequilibrium with a primary disease locus. BMC Proc 2007;1 Suppl 1:S163.
59. Menashe I, Rosenberg PS, Chen BE. PGA: power calculator for case-control genetic association analyses. BMC Genet 2008;9:36.
60. Swanberg M, Lidman O, Padyukov L, Eriksson P, Akesson E, Jagodic M, et al. MHC2TA is associated with differential MHC molecule expression and susceptibility to rheumatoid arthritis, multiple sclerosis and myocardial infarction. Nat Genet 2005;37(5):486-494.
61. Akkad DA, Jagiello P, Szyld P, Goedde R, Wieczorek S, Gross WL, et al. Promoter polymorphism rs3087456 in the MHC class II transactivator gene is not associated with susceptibility for selected autoimmune diseases in German patient groups. Int J Immunogenet 2006;33(1):59-61.
62. Martinez A, Sanchez-Lopez M, Varade J, Mas A, Martin MC, Heras VDL, et al. Role of the MHC2TA gene in autoimmune diseases. Ann Rheum Dis 2007;66(3):325-329.
63. O'Doherty C, Hawkins S, Rooney M, Vandenbroeck K. The MHC2TA -168A/G and +1614G/C polymorphisms and risk for multiple sclerosis or chronic inflammatory arthropathies. Tissue Antigens 2007;70(3):247-251.
64. Bronson PG, Criswell LA, Barcellos LF. The MHC2TA -168A/G polymorphism and risk for rheumatoid arthritis: a meta-analysis of 6861 patients and 9270 controls reveals no evidence for association. Ann Rheum Dis 2008;67(7):933-6.
65. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7(3):177-88.
66. Rasmussen HB, Kelly MA, Clausen J. Genetic susceptibility to multiple sclerosis: detection of polymorphic nucleotides and an intron in the 3' untranslated region of the major histocompatibility complex class II transactivator gene. Hum Immunol 2001;62(4):371-377.
67. Patarroyo JC, Stuve O, Piskurich JF, Hauser SL, Oksenberg JR, Zamvil SS. Single nucleotide polymorphisms in MHC2TA, the gene encoding the MHC class II transactivator (CIITA). Genes Immun 2002;3(1):34-7.
68. Ng PC, Henikoff S. Predicting deleterious amino acid substitutions. Genome Res 2001;11(5):863-74.
69. Martinez A, Alvarez-Lafuente R, Mas A, Bartolome M, Garcia-Montojo M, de Las Heras V, et al. Environment-gene interaction in multiple sclerosis: human herpesvirus 6 and MHC2TA. Hum Immunol 2007;68(8):685-9.
70. Alvarez-Lafuente R, Martinez A, Garcia-Montojo M, Mas A, De Las Heras V, Dominguez-Mozo MI, et al. MHC2TA rs4774C and HHV-6A active replication in multiple sclerosis patients. Eur J Neurol 2009;17(1):129-135.
71. Harnesk K, Swanberg M, Ockinger J, Diez M, Lidman O, Wallstrom E, et al. Vra4 Congenic rats with allelic differences in the class II transactivator gene display altered

susceptibility to experimental autoimmune encephalomyelitis. J Immunol 2008;180(5):3289-3296.
72. Price AL, Butler J, Patterson N, Capelli C, Pascali VL, Scarnicci F, et al. Discerning the ancestry of European Americans in genetic association studies. PLoS Genet 2008;4(1):e236.
73. ANZgene. Genome-wide association study identifies new multiple sclerosis susceptibility loci on chromosomes 12 and 20. Nat Genet 2009;41(7):824-8.
74. Baranzini SE, Wang J, Gibson RA, Galwey N, Naegelin Y, Barkhof F, et al. Genome-wide association analysis of susceptibility and clinical phenotype in multiple sclerosis. Hum Mol Genet 2009;18(4):767-78.
75. A haplotype map of the human genome. Nature 2005;437(7063):1299-320.
76. de Bakker PI, McVean G, Sabeti PC, Miretti MM, Green T, Marchini J, et al. A high-resolution HLA and SNP haplotype map for disease association studies in the extended human MHC. Nat Genet 2006;38(10):1166-72.

Table 1. Characteristics of MS cases and controls analyzed in this study (n = 3,656).

Clinical characteristic	MS % (n = 1,320)	Controls % (n = 1,363)	MS family probands % (n = 973)
Site			
Brigham and Women's Hospital, Boston	22.5 (294)	0.6 (8)	5.4 (53)
School of Clinical Medicine, Cambridge	50.2 (664)		46.6 (453)
University of California, San Francisco	27.3 (362)	44 (599)	48.0 (467)
Wellcome Trust Case Control Consortium		55.4 (756)	
Sex			
Males	37.7 (502)	35 (473)	48.4 (471)
Females	62.3 (818)	65 (890)	51.6 (502)
Age, years			
Range	25-91	25-86	21-92
Mean ± S.D.	52.0 ± 10.2	49.9 ± 12.3	50.7 ± 10.8
Age-at-onset, years			
Range	4-64		11-60
Mean ± S.D.	33.3 ± 9.9		30.3 ± 8.7
Disease course			
Relapsing-remitting	52.5 (693)		66.9 (651)
Secondary-progressive	23.2 (306)		20.3 (198)
Primary-progressive	12.0 (159)		4.7 (46)
Clinically isolated syndrome	3.4 (45)		3.3 (32)
Progressive-relapsing	1.5 (20)		1.2 (12)
Unknown	7.4 (97)		3.5 (34)
Number of *HLA-DRB1*1501* alleles[1]			
0	47.3 (625)	74.7 (1,018)	44.6 (434)
1	45.4 (599)	23.3 (318)	46.5 (452)
2	7.3 (96)	2.0 (27)	8.9 (87)

[1] The rs3135388 (A/G) SNP was genotyped to characterize *DRB1*1501* status, due to very strong correlation between the presence of rs3135388**A* and *DRB1*1501* as previously described (76)

Table 2. **Description of 24 *CIITA* SNP variants included in Cochran-Armitage trend and global haplotype association tests.**

SNP	Marker	Type	Position & Alleles	MAF	MS	HWE *P*	Controls	HWE *P*	Block
1	rs4436808	Promoter	C10870067T	0.01	0/26/1293	1	0/27/1336	1	
2	rs6498114	Promoter	T10871619G	0.25	75/478/767	0.94	96/499/768	0.22	
3	rs9302456	Promoter	C10875973T	0.34	164/574/581	0.24	160/601/599	0.63	
4	rs4781010	Promoter	A10877959T	0.01	0/24/1296	1	0/27/1336	1	1
5	rs3087456	Promoter	T10878403C	0.26	87/493/740	0.72	108/519/736	0.22	
6	rs12928665	Intron	T10878975C	0.25	85/470/764	0.26	98/505/759	0.29	1
7	rs12932187	Intron	G10879381C	0.06	5/156/1159	1	10/159/1194	0.08	1
8	rs12925158	Intron	C10881806T	0.02	0/52/1268	1	0/56/1307	1	1
9	rs8043545	Intron	G10889846C	0.27	96/510/714	0.72	119/526/718	0.11	
10	rs6498119	Intron	T10899379C	0.07	7/170/1143	0.83	7/153/1203	0.34	1
11	rs4781015	Intron	C10899453T	0.21	52/417/851	0.93	66/464/833	0.87	
12	rs7189406	Intron	T10900989C	0.07	6/173/1140	1	9/170/1184	0.29	2
13	rs6498124	Intron	A10903351C	0.43	241/687/391	0.05	251/639/473	0.18	2
14	rs4781016	Intron	G10903900T	0.28	98/563/658	0.14	98/540/725	0.89	3
15	rs4774	Missense[1]	G10908349C	0.29	125/552/643	0.70	108/515/740	0.17	3
16	rs13330686	Missense	G10909192A	0.08	11/216/1092	0.87	10/183/1169	0.32	
17	rs13336804	Missense	A10909195G	0.08	11/217/1092	0.87	10/183/1170	0.32	4
18	rs4781019	Intron	T10911651C	0.45	271/636/381	0.87	266/630/432	0.20	4
19	rs6498126	Intron	C10911864G	0.20	52/430/838	0.8	60/445/858	0.80	
20	rs4781020	Intron	C10915763T	0.30	127/538/653	0.29	128/566/668	0.61	5
21	rs6498132	Intron	G10918145A	0.11	15/282/1023	0.43	19/259/1085	0.40	5
22	rs2229322	SynCoding[2]	C10923546T	0.11	14/261/1045	0.78	15/255/1093	1	6
23	rs4781024	Intron	C10924559T	0.41	225/628/464	0.61	234/654/469	0.82	6
24	rs1139564	3' UTR[3]	G10926123A	0.20	68/402/843	0.03	72/388/900	$7.6x10^{-4}$	6

[1]Missense, missense mutation.
[2]Synonymous coding mutation.
[3]UTR, untranslated region.

Table 3. *P*-values[1] from the Cochran-Armitage test of association in 1,320 MS cases and 1,363 controls stratified by the presence of the *HLA-DRB1*1501*[2] risk allele (n = 2,683).

SNP	Marker	MS MAF	Controls MAF	Overall	*DRB1*1501+*	*DRB1*1501-*
1	rs4436808	0.01	0.01	1.00	0.24	0.37
2	rs6498114	0.24	0.25	0.19	0.64	0.22
3	rs9302456	0.34	0.34	0.81	0.42	0.38
4	rs4781010	0.01	0.01	0.79	0.20	0.37
5	rs3087456	0.25	0.27	0.16	0.87	0.07
6	rs12928665	0.24	0.26	0.22	0.90	0.10
7	rs12932187	0.06	0.07	0.70	0.38	0.59
8	rs12925158	0.02	0.02	0.84	0.71	0.63
9	rs8043545	0.27	0.28	0.24	1.00	0.13
10	rs6498119	0.07	0.06	0.22	0.06	0.78
11	rs4781015	0.20	0.22	0.06	0.51	0.07
12	rs7189406	0.07	0.07	0.86	0.91	0.43
13	rs6498124	0.44	0.42	0.07	**$1.8\text{x}10^{-3}$***	0.79
14	rs4781016	0.29	0.27	0.15	$8.5\text{x}10^{-3}$	0.82
15	rs4774	0.30	0.27	**$4.9\text{x}10^{-3}$***	**$1\text{x}10^{-4}$***	0.81
16	rs13330686	0.09	0.07	0.04	0.05	0.39
17	rs13336804	0.09	0.07	0.03	0.05	0.37
18	rs4781019	0.46	0.44	0.15	0.09	0.54
19	rs6498126	0.20	0.21	0.66	0.20	0.40
20	rs4781020	0.30	0.30	0.92	0.65	0.53
21	rs6498132	0.12	0.11	0.30	**$1.3\text{x}10^{-3}$***	0.12
22	rs2229322	0.11	0.10	0.57	0.25	0.81
23	rs4781024	0.41	0.41	0.76	0.07	0.15
24	rs1139564	0.20	0.20	0.42	0.40	0.86

[1]*P*-values based on 10, 000 permutations.
[2]Characterized by rs3135388 genotyping.
* Statistical significance.

Table 4. ORs and 95% CIs from an allelic test of association in 1,320 MS cases and 1,363 controls stratified by the presence of the *HLA-DRB1*1501*[1] risk allele (n = 2,683).

SNP	Marker	Overall	*DRB1*1501+*	*DRB1*1501-*
1	rs4436808	1.00 (0.58-1.71)	2.00 (0.67-6.00)	0.71 (0.34-1.49)
2	rs6498114	0.92 (0.81-1.04)	0.95 (0.77-1.18)	0.90 (0.76-1.06)
3	rs9302456	1.01 (0.91-1.14)	0.92 (0.76-1.12)	1.07 (0.92-1.24)
4	rs4781010	0.92 (0.53-1.59)	2.33 (0.67-8.13)	0.68 (0.32-1.42)
5	rs3087456	0.92 (0.81-1.04)	1.02 (0.83-1.26)	0.86 (0.73-1.01)
6	rs12928665	0.93 (0.82-1.05)	1.02 (0.82-1.25)	0.87 (0.74-1.03)
7	rs12932187	0.96 (0.77-1.19)	0.83 (0.57-1.21)	1.08 (0.82-1.43)
8	rs12925158	0.96 (0.65-1.40)	0.84 (0.42-1.67)	1.14 (0.71-1.83)
9	rs8043545	0.93 (0.83-1.05)	1.00 (0.81-1.22)	0.88 (0.75-1.03)
10	rs6498119	1.15 (0.93-1.43)	1.48 (0.99-2.21)	1.04 (0.79-1.38)
11	rs4781015	0.88 (0.77-1.00)	0.93 (0.74-1.16)	0.85 (0.71-1.01)
12	rs7189406	1.02 (0.83-1.26)	0.98 (0.67-1.43)	1.12 (0.86-1.46)
13	rs6498124	1.11 (0.99-1.23)	1.36 (1.13-1.63)	0.98 (0.85-1.13)
14	rs4781016	1.09 (0.97-1.23)	1.32 (1.07-1.62)	0.98 (0.84-1.15)
15	rs4774	1.19 (1.06-1.34)	1.60 (1.30-1.97)	0.98 (0.84-1.15)
16	rs13330686	1.23 (1.01-1.50)	1.40 (1.00-1.98)	1.12 (0.87-1.46)
17	rs13336804	1.24 (1.02-1.50)	1.40 (1.00-1.98)	1.14 (0.88-1.47)
18	rs4781019	1.08 (0.97-1.21)	1.18 (0.98-1.42)	1.05 (0.91-1.21)
19	rs6498126	0.97 (0.85-1.11)	0.86 (0.68-1.08)	1.08 (0.91-1.28)
20	rs4781020	0.99 (0.88-1.11)	0.95 (0.78-1.16)	1.05 (0.90-1.23)
21	rs6498132	1.10 (0.93-1.30)	1.60 (1.19-2.17)	0.83 (0.66-1.04)
22	rs2229322	1.05 (0.89-1.25)	1.20 (0.89-1.62)	0.97 (0.77-1.22)
23	rs4781024	0.98 (0.88-1.09)	0.84 (0.69-1.01)	1.11 (0.96-1.28)
24	rs1139564	1.06 (0.93-1.21)	1.11 (0.88-1.39)	1.02 (0.85-1.21)

[1]Chacterized by rs3135388 genotyping.

Table 5. Global *P*-values[1] from global haplotype association tests in MS cases (n = 1,320) and controls (n = 1,363) and also stratified by the presence of the *HLA-DRB1*1501*[2] risk allele (total n = 2,683).

Haplotype Block	Global *P*			Number of observed haplotypes[3]	Frequency of most common haplotype in haplotype block
	Overall	*DRB1*1501+*	*DRB1*1501-*		
1	0.73	0.66	0.22	4	0.37
2	0.10	0.41	0.05	3	0.79
3	0.31	$7x10^{-3}$	0.96	3	0.57
4	$3.6x10^{-2}$	0.04	0.26	2	0.92
5	0.61	0.10	0.44	3	0.50
6	0.81	$5x10^{-3}$	0.32	4	0.41

[1]*P*-values based on 10, 000 permutations, under the additive genetic model.
[2]Characterized by rs3135388 genotyping.
[3]Excluding haplotypes with inferred frequencies <5%.

Table 6. Results for the pedigree disequilibrium test (PDT) of rs4774 in 973 extended MS families stratified by the presence of the *HLA-DRB1*1501* risk allele (n = 2,917).

Results	Overall (n = 973)	*DRB1*1501+* (n = 539)	*DRB1*1501-* (n = 434)
Number of informative families	892	497	395
Number of informative DSPs[1]	489	288	201
Siblings affected : not affected, % (n) : % (n)			
rs4774**G*	68.8 (530) : 72.1 (705)	68.5 (304) : 73.4 (423)	69.3 (226) : 70.1 (282)
rs4774**C*	31.2 (240) : 27.9 (273)	31.5 (140) : 26.6 (153)	30.7 (100) : 29.9 (120)
Number of informative trios	674	357	317
Alleles transmitted : not transmitted, % (n) : % (n)			
rs4774**G*	70.3 (948) : 69.0 (930)	69.3 (495) : 70.5 (504)	71.5 (453) : 67.2 (426)
rs4774**C*	29.7 (400) : 31.0 (418)	30.7 (219) : 29.5 (211)	28.5 (181) : 32.8 (208)
***P*-value**[2]	0.25	**2.3×10^{-2}***	0.10

[1]The informative discordant sibling pairs (DSPs) contained 385, 222 and 163 affected offspring in the overall, *DRB1*1501*+ and *DRB1*1501*- samples, respectively.
[2]One-sided, asymptotic *P*-values.
* Statistical significance.

Table 7. Frequencies for the rs4774 variant and *HLA-DRB1*1501*[1] in MS cases (n = 1,320), controls (n = 1,363) and one MS case per family (n = 973).[2]

Number of rs4774*_C_ alleles	**Presence of *DRB1*1501***	**MS % (*n*)**	**Controls % (*n*)**
0	-	23.7 (543)	39.0 (531)
1	-	18.8 (430)	29.7 (405)
2	-	3.8 (86)	6.0 (82)
0	+	24.9 (570)	15.3 (209)
1	+	23.2 (532)	8.1 (110)
2	+	5.8 (132)	1.9 (26)

[1]*DRB1* or rs3135388 genotype data were available for all cases and controls in this study (8, 16, 40).
[2]Total n = 3,656 individuals (2,293 affected)

Table 8. Rs4774-rs6498169 haplotype frequencies in the combined *HLA-DRB1-1501+*[1] MS cases (n = 918)[2] and controls (n = 345) (total n = 1,263).

Presence of rs4774*C[3]	**MS % (*n*)**	**Controls % (*n*)**	**OR (95% CI)**	***P***
C*G*	14.9(274)	8.8(61)	1.67 (1.19-2.37)	1.9×10^{-3}
G*G*	24.4(448)	24.2(167)		
C*A*	17.9(329)	14.7(101)	1.49 (1.15-1.95)	2.3×10^{-3}
G*A*	42.8(786)	52.3(361)		

[1]*DRB1* or rs3135388 genotype data were available for all cases and controls in this study (8, 16, 40).

[2]Genotyping data for the rs6498169 *CLEC16A* variant was available for all 1,320 cases from the case-control sample (stage 1) (*DRB1*1501+* n = 695) and 518 cases from the family sample (stage 2) (*DRB1*1501+* n = 223).

[3]Association testing was performed for rs4774 (*C/G* in bold) conditioned on the presence or absence of the *CLEC16A* rs6498169**G* risk allele.

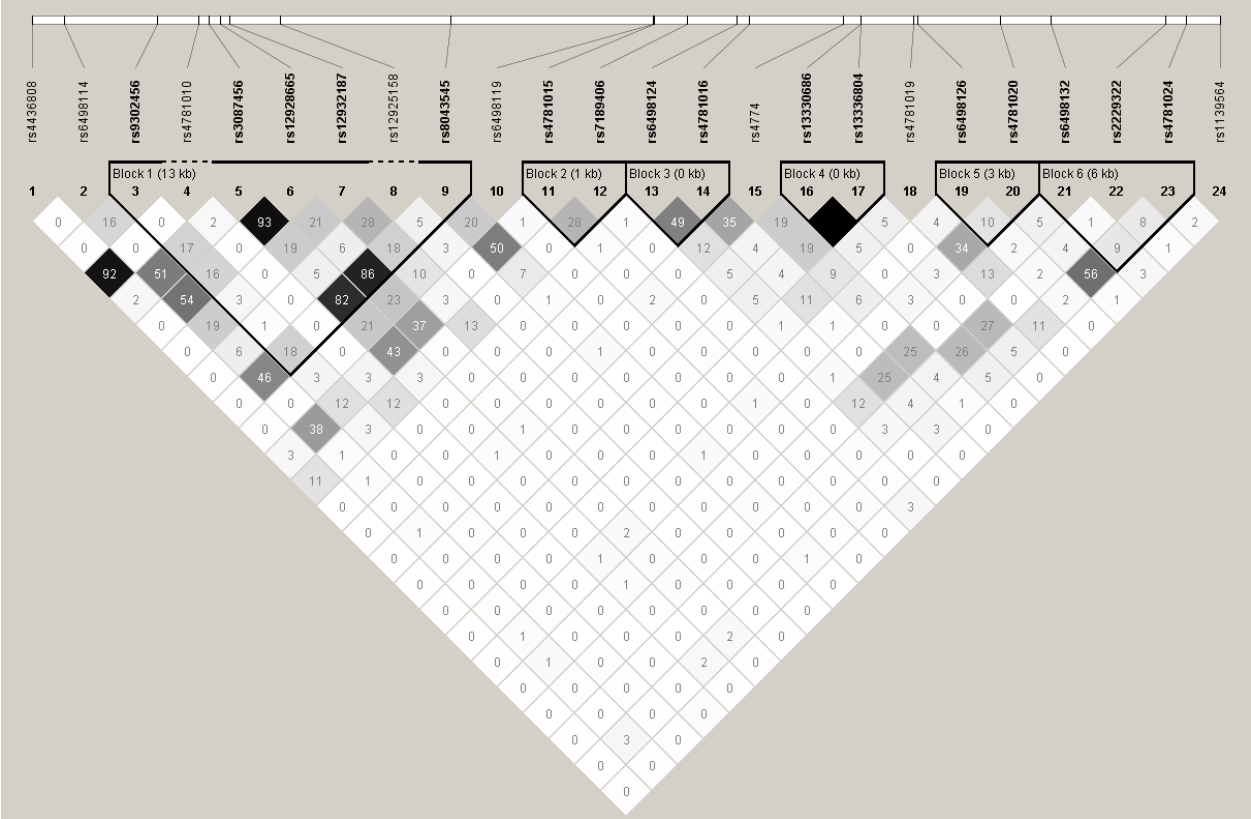
Table 9. Additional 19 *CIITA* SNP variants captured in this analysis, based on HapMap CEU ($r^2 \geq 0.8$).

Marker	Position	Alleles	Region
rs12922863	10872627	A/G	Promoter
rs6498115	10873012	C/T	Promoter
rs6416647	10873098	C/T	Promoter
rs7404672	10873980	C/T	Promoter
rs6498116	10876783	A/T	Promoter
rs12596540	10877977	A/G	Promoter
rs2071170	10878128	C/G	Promoter
rs4781011	10882812	G/T	Intron
rs11074934	10886941	C/T	Intron
rs8063850	10899122	A/T	Intron
rs8048002	10899489	C/T	Intron
rs11647384	10904790	A/G	Intron
rs4780333	10905983	C/T	Intron
rs4781018	10909634	C/G	Intron
rs4780334	10910127	A/G	Intron
rs4780335	10911829	C/G	Intron
rs12598246	10912233	C/T	Intron
rs11074939	10919210	A/G	Intron
rs9788916	10919229	A/G	Intron

Table 10. 8 *CIITA* SNP variants not captured in this analysis, based on HapMap CEU ($r^2 \geq 0.8$).

Marker	Position	Alleles	Type
rs6498122	10901683	A/G	Intron
rs7204799	10912050	C/G	Intron
rs11074938	10914044	A/G	Intron
rs6498130	10917651	G/T	Intron
rs6498131	10918127	C/T	Intron
rs7404786	10920051	C/G	Intron
rs8056269	10920068	C/G	Intron

Figure 1. R^2 plot illustrating the LD structure of *CIITA* SNP variants in healthy controls; darker gray indicates higher r^2 between pairs of SNPs.



CHAPTER SIX

A candidate gene study of *CLEC16A* does not provide evidence of association with risk for anti-CCP positive rheumatoid arthritis

In Press: Bronson PG, Ramsay PP, Seldin MF, Gregersen PK, Criswell LA, Barcellos LF. (2010) A candidate gene study of CLEC16A does not provide evidence of association with risk for anti-CCP positive rheumatoid arthritis. Genes Immun. doi:10.1038/gene.2010.7

ABSTRACT

CLEC16A, a putative immunoreceptor, was recently established as a susceptibility locus for type 1 diabetes and multiple sclerosis. Subsequently, associations between *CLEC16A* and rheumatoid arthritis (RA), Addison's disease and Crohn's disease have been reported. A large comprehensive and independent investigation of *CLEC16A* variation in RA was pursued. The current study tested 251 *CLEC16A* SNPs in 2542 RA cases (85% anti-cyclic citrullinated peptide [anti-CCP] positive) and 2210 controls ($N = 4752$). All individuals were of European ancestry, as determined by ancestry informative genetic markers. No evidence for significant association between *CLEC16A* variation and RA was observed. This is the first study to fully characterize common genetic variation in *CLEC16A* including assessment of haplotypes and gender-specific effects. The previously reported association between RA and rs6498169 was not replicated. Results demonstrate that *CLEC16A* does not play a prominent role in susceptibility to anti-CCP positive RA.

BACKGROUND

Rheumatoid arthritis (RA) is the most common systemic autoimmune disease with a prevalence of one percent (1). This chronic inflammatory disease can cause substantial disability from the erosive and deforming processes in joints, and is associated with increased mortality (2). RA has a strong genetic component, as demonstrated by twin and other family studies; however the etiology is unknown (3). Major histocompatibility complex (MHC) genes, particularly HLA class II, are strongly associated with risk of developing RA. However, MHC genes only account for a portion of the genetic risk. Several non-MHC genes have recently been associated with risk for RA, including *PTPN22*, *STAT4* and *TNFAIP3* (4-6). Results from recent genome-wide association (GWA) studies underscore the overlap of replicated findings across complex diseases, including autoimmune conditions (7, 8). Variants within some confirmed genetic risk loci for RA also confer risk for other autoimmune diseases. These include *CTLA4* in type 1 diabetes (T1D), *IL2* in T1D and Celiac disease, *PTPN22* in systemic lupus erythematosus (SLE), T1D and autoimmune thyroid disease, *STAT4* in SLE and TNFAIP3 in SLE, T1D, Celiac disease and Crohn's disease (5, 9-19).

The C-type lectin domain family 16, member A gene (*CLEC16A*, previously called *KIAA0350*) spans 237.7 kb and encodes a sugar-binding receptor that contains a putative immunoreceptor tyrosine-based activation motif (10). C-type lectin receptors can be expressed on dendritic cells to distinguish between self and non-self glycoproteins, and may be involved in immune activation and peripheral tolerance (20, 21). These sugar binding receptors have been shown to play a role in multiple animal models for RA (22-25). For example, in rats, C-type lectin-like receptors are encoded by the antigen-presenting lectin-like receptor gene complex (*APLEC*), which have been shown to influence susceptibility to arthritis (oil-, collagen-, squalene- and pristine-induced), autoimmune phenotypes (autoantibody levels) and clinical phenotypes (day of disease onset, maximal severity, severity over time, body weight loss, arthritis symptoms) (24). The effect of *APLEC* variation on susceptibility to arthritis and clinical phenotypes varied by gender (24).

Recently, GWA studies have identified the sugar binding receptor gene *CLEC16A* as a novel risk locus for T1D and MS, and this association has since been replicated in independent samples (10, 26-31). *CLEC16A* is located on 16p13, a region that has been implicated in RA linkage studies (32). The purpose of this study was to perform a comprehensive haplotype-based investigation of *CLEC16A* as a candidate RA gene. This study sample consisted of 682 RA cases and 752 controls collected by the North American RA Consortium (NARAC) [RA1], 1860 RA cases collected by the Wellcome Trust Case Control Consortium (WTCCC) RA Group in the U.K. and 1458 controls collected by the WTCCC from the U.K. Blood Services [RA2] (total N = 4752) (Table 1).

METHODS

Patients

RA cases met the American College of Rheumatology classification criteria for RA (33). RA2 controls were a subset of the WTCCC T1D GWA study controls (19). RA1 controls were frequency matched by age and gender to the cases. RA2 controls were frequency matched by geographical region and gender to the 1958 Birth cohort (which included all births in England,

Wales and Scotland, during one week in 1958) so as to be nationally representative. Based on the available genetic ancestry data for all individuals, and to apply the most stringent criteria possible for genetic analysis of *CLEC16A*, only RA1 subjects with ≥90% N. European ancestry and RA2 subjects with European ancestry were analyzed. Previous GWA studies provided genotyping data for 64 *CLEC16A* SNPs in RA1 derived from the Illumina HumanHap550 Genotyping BeadChip (San Diego, CA, USA) at the Feinstein Institute for Medical Research and 49 *CLEC16A* SNPs in RA2 from the Affymetrix GeneChips Mapping 500K Array Set (Santa Clara, CA, USA) as previously described (19, 34, 35).

Statistical analysis

European ancestry was estimated in RA1 using a Bayesian clustering algorithm (Structure v. 2.0) and data for 112 European and 246 Northern European ancestry informative markers (36, 37). For RA2, European ancestry was estimated by principal components analysis (19). Three SNPs in RA1 and 6 SNPs in RA2 were excluded from analysis due to low minor allele frequency (MAF) (<0.01). Deviation from Hardy-Weinberg equilibrium (HWE) was examined in controls separately for each cohort using the exact test (PLINK v. 1.05) (38, 39). Three SNPs from RA1 with evidence for deviation from HWE in the controls (P<0.001) were omitted from further analyses.

Sufficient power for this study was confirmed with PGA v.2.0 (two-sided $\alpha = 0.05$) (40). Haplotype blocks were estimated in RA1 and RA2 controls and CEU separately (Haploview v.4.1) (41). Percent of *CLEC16A* variation captured was based on $r^2 \geq 0.8$ in CEU using two- and three-marker haplotypes (HAPLOVIEW).

Allelic association was tested by creating *2x2* contingency tables and estimating ORs with Fisher's exact test (PLINK). Haplotypes were estimated with the expectation-maximization (EM) algorithm (HAPLOVIEW). Maximum likelihood estimates of haplotype probabilities were computed with the EM algorithm and score statistics were used for global haplotype association tests, assuming a dominant genetic model (HaploStats v.1.4.3, R v.2.6) (42). Haplotypes with inferred frequencies <5% were excluded. A significance threshold of $P = 1.1 \times 10^{-3}$ was set using a Bonferroni correction for the number of *CLEC16A* haplotype blocks (10) and SNPs that were not located in haplotype blocks (34), based on CEU. Empirical *P*-values based on 10, 000 simulations were reported for all allelic and haplotype tests. Allelic and haplotype empirical *P*-values were estimated in PLINK (max(T) permutation procedure) and HaploStats, respectively, by permuting the ordering of the disease status, counting the number of times the permuted test was greater than the observed test, and dividing by the total number of simulations (10,000).(38, 42). Because there is no evidence of an association of age or gender with the polymorphisms of interest we decided not to adjust for either.

In order to conduct a combined analysis of RA1 + RA2, missing genotypes were imputed for 38 SNPs in RA1, 53 SNPs in RA2 and 171 SNPs in the combined RA sample. A hidden Markov Model based algorithm was used to infer missing genotypes from known haplotypes (IMPUTE (v.0.5.0 (Oxford, UK)).(43) The robustness of the imputation accuracy rate for this standard imputation method has been demonstrated (44). Known haplotypes were obtained from publicly available genotype data for CEU, using observed linkage disequilibrium patterns ($r^2 \geq 0.8$) in two 500 kb regions adjacent to each side of *CLEC16A*.(43) Association tests of

imputed genotypes accounted for the uncertainty of imputed genotypes in missing data likelihood score tests, using the frequentist proper option and a dominant genetic model in SNPTEST (v.1.1.5 (Oxford, UK)).(43) Imputed genotypes with <90% probability were omitted. After omitting 12 SNPs with evidence for deviation from HWE in the controls and 4 SNPs with low MAF from further analyses, 251 SNPs in RA1+RA2 were tested for allelic association.

RESULTS

We conducted allelic tests of association for 58 SNPs and global haplotype tests (12 haplotype blocks encompassing 53 SNPs) in 682 anti-cyclic citrullinated peptide positive (anti-CCP positive) RA cases and 752 controls (N = 1434 [RA1]) (Figure 1). All results were negative after correcting for multiple testing (Figure 2, Table 2). Next, we conducted allelic tests of 43 SNPs and global haplotype tests (7 haplotype blocks encompassing 37 SNPs) in the second RA dataset comprised of 1860 RA cases and 1458 controls (N = 3318 [RA2]). No evidence for association was present (Figure 2, Table 2). Furthermore, allelic tests of 251 imputed SNPs within *CLEC16A* derived for the combined RA sample (2542 cases and 2210 controls, total N = 4752 [RA1+RA2]) revealed no evidence for disease association (Figure 2, Table 2).

DISCUSSION

The six *CLEC16A* SNPs shown to be associated with T1D and/or MS are intronic and were either genotyped or tagged (r^2>0.95 based on the Caucasian HapMap population [CEU]) in the current study. Similar to the current study, candidate gene investigations of *CLEC16A* in Grave's disease, Celiac disease and ulcerative colitis have been negative, but associations have been reported with Addison's disease, Crohn's disease and for RA in other datasets (10, 29, 45-48). A case-control study by Martinez et al. examined three *CLEC16A* SNPs and reported that rs6498169**G*, a variant associated with MS, was over-represented in RA cases (38%) compared to controls (32%) ($P = 8x10^{-3}$, odds ratio (OR) = 1.27, 95% confidence interval (CI) = 1.06-1.51) (29). Although our study was well-powered to detect such an effect size, with 80% power to detect an OR as low as 1.13, the association between RA and rs6498169 was not replicated. The rs6498169**G* allele frequency did not differ between RA cases (33.6%) and healthy controls in the current study (32.9%) (P = 0.45, OR = 1.03, 95% CI = 0.95-1.11).

It is also important to note that recent studies have revealed the presence of different MHC associations in anti-CCP positive and negative RA cases when considered separately (49-51). It is possible that this phenotypic difference may also be important for other RA genetic susceptibility loci. The well-established *PTPN22* RA locus appears to be associated only with anti-CCP positive RA, although some studies have reported association with both anti-CCP positive and negative RA (52-55). Anti-CCP autoantibodies and shared epitope alleles are also markers for increased RA severity, particularly when both are present (56). In the current study, 85% of RA cases were anti-CCP positive, compared to only 50% in the Martinez et al. study. This difference may have contributed to the observed disparity between results. Indeed, Skinningsrud et al. have recently examined three *CLEC16A* SNPs and reported that the rs6498169**G* variant was over-represented in anti-CCP negative RA cases (44%) compared to anti-CCP positive RA cases (37.7%) (P = 0.016, OR = 1.3, 95% CI = 1.05-1.61) and controls (35.9%) ($P = 2x10^{-4}$, OR = 1.4, 95% CI = 1.18-1.68) (48). Martinez et al. did not observe

differences between cases and controls after stratifying for anti-CCP status or presence/absence of shared epitope alleles, but this may be due to a lack of statistical power. Although all of our RA1 cases were anti-CCP positive, only 80% of RA2 cases were anti-CCP positive and this information was not publicly available for the RA2 cases. Therefore, we were not able to stratify RA2 or RA1+RA2 by anti-CCP status for analyses of *CLEC16A* SNPs.

Because animal models suggest that C-type lectin receptor genes may have gender-specific effects on autoimmunity, we conducted gender-stratified allelic tests and gender-adjusted global haplotype tests of *CLEC16A* within RA1 and RA2 (24). The rs3960630 *A* variant was under-represented in female RA1 cases (20%) compared to female controls (25%) (OR = 0.71, 95% CI = 0.59-0.86, $P = 4x10^{-4}$). This intronic SNP was not present in or captured by RA2 data and therefore could not be tested in the larger combined dataset. Given the number of multiple tests performed, these results should be interpreted with caution. Results did not differ when global haplotype tests were adjusted by gender (data not shown). Animal models of RA also indicate that it may be worthwhile to stratify cases by clinical phenotypes in future genetic studies of C-type lectin receptors and autoimmunity (24).

Although rare variants in *CLEC16A* were not directly investigated here, for the first time all common genetic variation within *CLEC16A* was interrogated for a role in RA susceptibility. Even without imputed genotypes, the RA1 dataset ($N = 58$ SNPs) captured 93%, RA2 ($N = 43$ SNPs) captured 80%, and both datasets combined ($N = 96$ SNPs) captured 96% of the common variation based on CEU data from HapMap. The data used in this study were taken from GWA studies that did not identify *CLEC16A* as a risk locus for RA based on stringent genome-wide significance. A focused candidate gene study that captures a larger portion of genetic variation compared to initial GWA studies is a useful and complementary strategy.

CONCLUSION

In conclusion, this is the first candidate gene study of *CLEC16A* to fully characterize common genetic variation in *CLEC16A* including assessment of haplotypes and gender-specific effects. We did not replicate the association between RA and rs6498169 reported by other studies. Results convincingly demonstrate that variation within *CLEC16A* does not play a prominent role in susceptibility to anti-CCP positive RA.

ACKNOWLEDGEMENTS

We would like to thank Farren Briggs, Benjamin Goldstein, Alan Hubbard and Ira Tager, as well as study participants. This work was supported by an Abbott Graduate Student Achievement Award (ACR REF), grants R01 AI065841, R01 AI059829 and F31 AI075609 (NIH/NIAID), and grants RO1 AR44422, NO1 AR22263, R01 AR050267, K24 AR02175 (NIH/NIAMS). The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the NIH, NIAID or NIAMS. This study makes use of data generated by the WTCCC; see www.wtccc.org.uk for a full list of the investigators who contributed to the generation of the data; funding was provided by the Wellcome Trust under award 076113. These studies were performed in part in the General Clinical Research Center, Moffitt Hospital, UCSF, with funds provided by NCRR, 5 M01 RR-00079, U.S. PHS.

REFERENCES

1. Gabriel SE. The epidemiology of rheumatoid arthritis. Rheum Dis Clin North Am 2001;27(2):269-81.
2. Gabriel SE, Crowson CS, Kremers HM, Doran MF, Turesson C, O'Fallon WM, et al. Survival in rheumatoid arthritis: a population-based analysis of trends over 40 years. Arthritis Rheum 2003;48(1):54-8.
3. MacGregor AJ, Snieder H, Rigby AS, Koskenvuo M, Kaprio J, Aho K, et al. Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins. Arthritis Rheum 2000;43(1):30-7.
4. Begovich AB, Carlton VE, Honigberg LA, Schrodi SJ, Chokkalingam AP, Alexander HC, et al. A missense single-nucleotide polymorphism in a gene encoding a protein tyrosine phosphatase (PTPN22) is associated with rheumatoid arthritis. Am J Hum Genet 2004;75(2):330-7.
5. Remmers EF, Plenge RM, Lee AT, Graham RR, Hom G, Behrens TW, et al. STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus. N Engl J Med 2007;357(10):977-86.
6. Plenge RM, Cotsapas C, Davies L, Price AL, de Bakker PIW, Maller J, et al. Two independent alleles at 6q23 associated with risk of rheumatoid arthritis. Nat Genet 2007;39(12):1477-1482.
7. Altshuler D, Daly MJ, Lander ES. Genetic mapping in human disease. Science 2008;322(5903):881-8.
8. Lettre G, Rioux JD. Autoimmune diseases: insights from genome-wide association studies. Hum Mol Genet 2008;17(R2):R116-21.
9. Ueda H, Howson JM, Esposito L, Heward J, Snook H, Chamberlain G, et al. Association of the T-cell regulatory gene CTLA4 with susceptibility to autoimmune disease. Nature 2003;423(6939):506-11.
10. Todd JA, Walker NM, Cooper JD, Smyth DJ, Downes K, Plagnol V, et al. Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes. Nat Genet 2007;39(7):857-864.
11. van Heel DA, Franke L, Hunt KA, Gwilliam R, Zhernakova A, Inouye M, et al. A genome-wide association study for celiac disease identifies risk variants in the region harboring IL2 and IL21. Nat Genet 2007;39(7):827-9.
12. Kyogoku C, Langefeld CD, Ortmann WA, Lee A, Selby S, Carlton VE, et al. Genetic association of the R620W polymorphism of protein tyrosine phosphatase PTPN22 with human SLE. Am J Hum Genet 2004;75(3):504-7.
13. Bottini N, Musumeci L, Alonso A, Rahmouni S, Nika K, Rostamkhani M, et al. A functional variant of lymphoid tyrosine phosphatase is associated with type I diabetes. Nat Genet 2004;36(4):337-8.
14. Criswell LA, Pfeiffer KA, Lum RF, Gonzales B, Novitzke J, Kern M, et al. Analysis of families in the multiple autoimmune disease genetics consortium (MADGC) collection: the PTPN22 620W allele associates with multiple autoimmune phenotypes. Am J Hum Genet 2005;76(4):561-71.
15. Musone SL, Taylor KE, Lu TT, Nititham J, Ferreira RC, Ortmann W, et al. Multiple polymorphisms in the TNFAIP3 region are independently associated with systemic lupus erythematosus. Nat Genet 2008;40(9):1062-4.

16. Graham RR, Cotsapas C, Davies L, Hackett R, Lessard CJ, Leon JM, et al. Genetic variants near TNFAIP3 on 6q23 are associated with systemic lupus erythematosus. Nat Genet 2008;40(9):1059-61.
17. Fung EY, Smyth DJ, Howson JM, Cooper JD, Walker NM, Stevens H, et al. Analysis of 17 autoimmune disease-associated variants in type 1 diabetes identifies 6q23/TNFAIP3 as a susceptibility locus. Genes Immun 2009;10(2):188-91.
18. Trynka G, Zhernakova A, Romanos J, Franke L, Hunt KA, Turner G, et al. Coeliac disease-associated risk variants in TNFAIP3 and REL implicate altered NF-kappaB signalling. Gut 2009;58(8):1078-83.
19. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007;447(7145):661-678.
20. Bates EE, Fournier N, Garcia E, Valladeau J, Durand I, Pin JJ, et al. APCs express DCIR, a novel C-type lectin surface receptor containing an immunoreceptor tyrosine-based inhibitory motif. J Immunol 1999;163(4):1973-83.
21. Geijtenbeek TB, van Vliet SJ, Engering A, t Hart BA, van Kooyk Y. Self- and nonself-recognition by C-type lectins on dendritic cells. Annu Rev Immunol 2004;22:33-54.
22. Fujikado N, Saijo S, Iwakura Y. Identification of arthritis-related gene clusters by microarray analysis of two independent mouse models for rheumatoid arthritis. Arthritis Res Ther 2006;8(4):R100.
23. Fujikado N, Saijo S, Yonezawa T, Shimamori K, Ishii A, Sugai S, et al. Dcir deficiency causes development of autoimmune diseases in mice due to excess expansion of dendritic cells. Nature Medicine 2008;14(2):176-180.
24. Guo JP, Backdahl L, Marta M, Mathsson L, Ronnelid J, Lorentzen JC. Profound and paradoxical impact on arthritis and autoimmunity of the rat antigen-presenting lectin-like receptor complex. Arthritis and Rheumatism 2008;58(5):1343-1353.
25. Lorentzen JC, Flornes L, Eklow C, Backdahl L, Ribbhammar U, Guo JP, et al. Association of arthritis with a gene complex encoding C-type lectin-like receptors. Arthritis and Rheumatism 2007;56(8):2620-2632.
26. Hakonarson H, Grant SFA, Bradfield JP, Marchand L, Kim CE, Glessner JT, et al. A genome-wide association study identifies KIAA0350 as a type 1 diabetes gene. Nature 2007;448(7153):591-594.
27. Hafler DA, Compston A, Sawcer S, Lander ES, Daly MJ, De Jager PL, et al. Risk alleles for multiple sclerosis identified by a genomewide study. N Engl J Med 2007;357(9):851-62.
28. Zoledziewska M, Costa G, Pitzalis M, Cocco E, Melis C, Moi L, et al. Variation within the CLEC16A gene shows consistent disease association with both multiple sclerosis and type 1 diabetes in Sardinia. Genes Immun 2009;10(1):15-7.
29. Martinez A, Perdigones N, Cenit MC, Espino L, Varade J, Lamas JR, et al. Chromosomal region 16p13: further evidence of increased predisposition to immune diseases. Ann Rheum Dis 2010;69(1):309-11.
30. Barrett JC, Clayton DG, Concannon P, Akolkar B, Cooper JD, Erlich HA, et al. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. Nat Genet 2009;41(6):703-707.
31. Rubio JP, Stankovich J, Field J, Tubridy N, Marriott M, Chapman C, et al. Replication of KIAA0350, IL2RA, RPL5 and CD58 as multiple sclerosis susceptibility genes in Australians. Genes Immun 2008;9(7):624-30.

32. Etzel CJ, Chen WV, Shepard N, Jawaheer D, Cornelis F, Seldin MF, et al. Genome-wide meta-analysis for rheumatoid arthritis. Hum Genet 2006;119(6):634-41.
33. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988;31(3):315-324.
34. Plenge RM, Seielstad M, Padyukov L, Lee AT, Remmers EF, Ding B, et al. TRAF1-C5 as a Risk Locus for Rheumatoid Arthritis -- A Genomewide Study. N Engl J Med 2007;357(12):1199-1209.
35. Hom G, Graham RR, Modrek B, Taylor KE, Ortmann W, Garnier S, et al. Association of Systemic Lupus Erythematosus with C8orf13-BLK and ITGAM-ITGAX. N Engl J Med 2008;358(9):900-909.
36. Falush D, Stephens M, Pritchard JK. Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. Genetics 2003;164(4):1567-1587.
37. Tian C, Plenge RM, Ransom M, Lee A, Villoslada P, Selmi C, et al. Analysis and Application of European Genetic Substructure Using 300 K SNP Information. PLoS Genetics 2008;4(1):29-39.
38. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81(3):559-75.
39. Wigginton JE, Cutler DJ, Abecasis GR. A note on exact tests of Hardy-Weinberg equilibrium. Am J Hum Genet 2005;76(5):887-93.
40. Menashe I, Rosenberg PS, Chen BE. PGA: power calculator for case-control genetic association analyses. BMC Genet 2008;9:36.
41. Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, et al. The structure of haplotype blocks in the human genome. Science 2002;296(5576):2225-9.
42. Schaid DJ, Rowland CM, Tines DE, Jacobson RM, Poland GA. Score tests for association between traits and haplotypes when linkage phase is ambiguous. Am J Hum Genet 2002;70(2):425-34.
43. Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. Nat Genet 2007;39(7):906-913.
44. Pei Y-F, Li J, Zhang L, Papasian CJ, Deng H-W. Analyses and Comparison of Accuracy of Different Genotype Imputation Methods. PLoS ONE 2008;3(10):e3551.
45. Dema B, Martinez A, Fernandez-Arquero M, Maluenda C, Polanco I, Angeles Figueredo M, et al. Autoimmune disease association signals in CIITA and KIAA0350 are not involved in celiac disease susceptibility. Tissue Antigens 2009;73(4):326-9.
46. Marquez A, Varade J, Robledo G, Martinez A, Mendoza JL, Taxonera C, et al. Specific association of a CLEC16A/KIAA0350 polymorphism with NOD2/CARD15(-) Crohn's disease patients. Eur J Hum Genet 2009; e-pub ahead of print 1 April 2009; doi: 10.1038/ejhg.2009.50.
47. Skinningsrud B, Husebye ES, Pearce SH, McDonald DO, Brandal K, Wolff AB, et al. Polymorphisms in CLEC16A and CIITA at 16p13 are associated with primary adrenal insufficiency. J Clin Endocrinol Metab 2008;93(9):3310-7.
48. Skinningsrud B, Lie BA, Husebye ES, Kvien TK, Forre O, Flato B, et al. A CLEC16A variant confers risk for juvenile idiopathic arthritis and anti-CCP negative rheumatoid arthritis. Ann Rheum Dis 2009.
49. Ding B, Padyukov L, Lundstrom E, Seielstad M, Plenge RM, Oksenberg JR, et al. Different patterns of associations with anti-citrullinated protein antibody-positive and anti-

citrullinated protein antibody-negative rheumatoid arthritis in the extended major histocompatibility complex region. Arthritis Rheum 2009;60(1):30-8.
50. Irigoyen P, Lee AT, Wener MH, Li W, Kern M, Batliwalla F, et al. Regulation of anti-cyclic citrullinated peptide antibodies in rheumatoid arthritis: contrasting effects of HLA-DR3 and the shared epitope alleles. Arthritis Rheum 2005;52(12):3813-8.
51. Verpoort KN, van Gaalen FA, van der Helm-van Mil AH, Schreuder GM, Breedveld FC, Huizinga TW, et al. Association of HLA-DR3 with anti-cyclic citrullinated peptide antibody-negative rheumatoid arthritis. Arthritis Rheum 2005;52(10):3058-62.
52. Wesoly J, van der Helm-van Mil AH, Toes RE, Chokkalingam AP, Carlton VE, Begovich AB, et al. Association of the PTPN22 C1858T single-nucleotide polymorphism with rheumatoid arthritis phenotypes in an inception cohort. Arthritis Rheum 2005;52(9):2948-50.
53. Kallberg H, Padyukov L, Plenge RM, Ronnelid J, Gregersen PK, van der Helm-van Mil AH, et al. Gene-gene and gene-environment interactions involving HLA-DRB1, PTPN22, and smoking in two subsets of rheumatoid arthritis. Am J Hum Genet 2007;80(5):867-75.
54. Pierer M, Kaltenhauser S, Arnold S, Wahle M, Baerwald C, Hantzschel H, et al. Association of PTPN22 1858 single-nucleotide polymorphism with rheumatoid arthritis in a German cohort: higher frequency of the risk allele in male compared to female patients. Arthritis Res Ther 2006;8(3):R75.
55. Farago B, Talian GC, Komlosi K, Nagy G, Berki T, Gyetvai A, et al. Protein tyrosine phosphatase gene C1858T allele confers risk for rheumatoid arthritis in Hungarian subjects. Rheumatol Int 2009;29(7):793-6.
56. Van Gaalen FA, Van Aken J, Huizinga TW, Schreuder GM, Breedveld FC, Zanelli E, et al. Association between HLA class II genes and autoantibodies to cyclic citrullinated peptides (CCPs) influences the severity of rheumatoid arthritis. Arthritis Rheum 2004;50(7):2113-21.

Table 1. RA study cohorts utilized for *CLEC16A* analyses.

	RA1	**Controls**	**RA2**	**Controls**
N	682	752	1860	1458
Site	N.A.	N.A.	U.K.	U.K.
Mean age, years	56.2	48.5	-	-
Age range, years	21-87	30-82	-	<70
Female, *N (%)*	503 (73.7)	525 (69.8)	1390 (74.7)	753 (51.6)
Mean age-at-onset, years	45.7		-	
Rheumatoid factor positive, *N (%)*	580 (85)		1310 (83.9)	
Shared epitope[a] (no. of copies), *N (%)*				
0	15 (2.3)	401 (53.3)	286 (20.7)	
1	362 (56.5)	301 (40)	680 (49.2)	
2	264 (41.2)	50 (6.6)	416 (30.1)	
Erosions, *N (%)*	211 (66.6)		-	
Anti-CCP Positive, *N (%)*	681 (100)		884 (79.8)	

[1]*HLA-DRB1*0101, *0102, *0104, *0401, *0404, *0405, *0408, *0413, *0416, *1001* alleles.

Table 2. MAF, OR, 95% CI and *P*-values from allelic tests and *P*-values from global haplotype tests in healthy controls and RA cases.

				MAF		Allelic test		Haplotype test	
Marker	**Location**	**Function**	**Sample**	**Cases**	**Controls**	**OR (95% CI)**	***P***	**Block**	***P***
rs3743976	T10946325C	intron	RA1+RA2	0.212	0.207	1.03 (0.94-1.13)	0.49		
rs16957801	G10946703A	intron	RA1+RA2	0.021	0.021	1.00 (0.77-1.28)	0.96		
rs12598008	G10947194C	intron	RA1+RA2	0.412	0.399	1.06 (0.98-1.15)	0.16		
			RA2	0.410	0.412	0.99 (0.90-1.10)	0.88		
rs12922318	A10948127T	intron	RA1+RA2	0.168	0.173	0.97 (0.88-1.07)	0.49		
rs16957807	A10948834G	intron	RA1+RA2	0.180	0.190	0.94 (0.85-1.03)	0.15		
rs8051196	C10948870G	intron	RA1+RA2	0.037	0.039	0.95 (0.79-1.15)	0.63		
rs8051340	G10948914C	intron	RA1+RA2	0.101	0.104	0.97 (0.86-1.09)	0.61		
rs12931878	G10949695A	intron	RA1+RA2	0.156	0.146	1.08 (0.97-1.20)	0.18		
			RA1	0.192	0.182	1.07 (0.89-1.29)	0.49	1	$5.4x10^{-2}$
rs8055533	A10949740G	intron	RA1+RA2	0.291	0.304	0.94 (0.86-1.03)	0.17		
			RA1	0.313	0.331	0.92 (0.79-1.08)	0.32	1	$5.4x10^{-2}$
rs8055876	A10949895G	intron	RA1+RA2	0.121	0.129	0.93 (0.83-1.04)	0.20		
			RA1	0.114	0.135	0.83 (0.66-1.04)	0.10	2	0.25
rs6498137	A10950451G	intron	RA1+RA2	0.037	0.040	0.93 (0.77-1.12)	0.47		
rs3813754	T10952434A	intron	RA1+RA2	0.180	0.190	0.93 (0.85-1.02)	0.13		
rs6498138	T10952779A	intron	RA1+RA2	0.137	0.140	0.98 (0.88-1.09)	0.67		
			RA2	0.151	0.153	0.99 (0.86-1.13)	0.86	1	0.89
rs10492842	A10953094T	intron	RA1+RA2	0.025	0.025	1.00 (0.79-1.25)	0.96		
rs8055123	G10953220A	intron	RA1+RA2	0.179	0.190	0.93 (0.85-1.02)	0.11		
			RA1	0.174	0.191	0.89 (0.74-1.08)	0.25	2	0.25
rs8057540	A10953343G	intron	RA1+RA2	0.036	0.038	0.95 (0.77-1.18)	0.66		
			RA1	0.154	0.149	1.04 (0.85-1.27)	0.74	2	0.25
rs8059156	T10953493C	intron	RA1+RA2	0.119	0.121	0.98 (0.88-1.10)	0.74		
rs12443971	G10954327A	intron	RA1+RA2	0.406	0.395	1.04 (0.97-1.12)	0.24		
rs6498139	A10955424G	intron	RA1+RA2	0.151	0.152	0.99 (0.90-1.10)	0.92		
rs7185978	A10955775T	intron	RA1+RA2	0.151	0.151	1.00 (0.90-1.10)	0.96		
rs1035089	A10955851G	intron	RA1+RA2	0.448	0.440	1.03 (0.96-1.11)	0.38		
			RA1	0.442	0.438	1.02 (0.88-1.18)	0.85	2	0.25
rs1035088	G10955875A	intron	RA1+RA2	0.215	0.212	1.02 (0.93-1.11)	0.68		
rs7186264	A10956451C	intron	RA1+RA2	0.179	0.190	0.93 (0.85-1.02)	0.13		
rs7192171	C10956620T	intron	RA1+RA2	0.119	0.121	0.98 (0.88-1.10)	0.74		
rs16957835	T10959556G	intron	RA1+RA2	0.122	0.124	0.98 (0.88-1.09)	0.71		
			RA2	0.117	0.124	0.94 (0.81-1.09)	0.41	1	0.89
rs4781027	A10959949G	intron	RA1+RA2	0.406	0.396	1.04 (0.97-1.12)	0.24		
rs7404554	C10960425T	intron	RA1+RA2	0.410	0.400	1.04 (0.97-1.12)	0.26		
			RA1	0.416	0.394	1.10 (0.94-1.27)	0.25	2	0.25
rs16957836	G10960587A	intron	RA1+RA2	0.119	0.122	0.98 (0.88-1.09)	0.72		
rs16957839	T10963879C	syn-cod[1]	RA1+RA2	0.111	0.113	0.98 (0.88-1.10)	0.73		
rs1700820	A10964119G	intron	RA1+RA2	0.419	0.407	1.05 (0.98-1.13)	0.17		
			RA2	0.413	0.408	1.02 (0.93-1.13)	0.66	1	0.89
rs1700818	C10964770G	intron	RA1+RA2	0.415	0.404	1.05 (0.97-1.13)	0.20		
rs4781028	G10966864C	intron	RA1+RA2	0.209	0.218	0.95 (0.87-1.03)	0.21		
rs9302457	A10967338G	intron	RA1+RA2	0.331	0.342	0.95 (0.88-1.02)	0.19		
rs8059260	G10967652A	intron	RA1+RA2	0.151	0.152	0.99 (0.90-1.10)	0.90		
			RA1	0.155	0.148	1.06 (0.86-1.30)	0.60	2	0.25
rs16957843	T10968706C	intron	RA1+RA2	0.182	0.192	0.94 (0.85-1.03)	0.17		
			RA2	0.182	0.188	0.96 (0.84-1.08)	0.48	1	0.89
rs12923849	A10969498G	intron	RA1+RA2	0.171	0.175	0.98 (0.89-1.07)	0.65		
			RA1	0.164	0.164	1.00 (0.82-1.22)	0.98	2	0.25
rs8049278	T10970124G	intron	RA1+RA2	0.120	0.123	0.98 (0.88-1.09)	0.68		
rs17229044	T10970437C	intron	RA1+RA2	0.205	0.204	1.01 (0.92-1.10)	0.87		
			RA1	0.212	0.205	1.04 (0.87-1.25)	0.66	2	0.25
			RA2	0.204	0.199	1.03 (0.91-1.16)	0.66	1	0.89
rs2302558	T10970512C	intron	RA1+RA2	0.120	0.123	0.98 (0.87-1.09)	0.65		
			RA1	0.122	0.119	1.04 (0.83-1.30)	0.78	2	0.25

rs12921922	C10971822T	intron	RA1+RA2	0.206	0.204	1.01 (0.93-1.11)	0.79		
rs16957849	C10972808T	intron	RA1+RA2	0.211	0.221	0.95 (0.87-1.03)	0.21		
			RA1	0.209	0.227	0.90 (0.75-1.07)	0.25	2	0.25
rs17803698	G10977790A	intron	RA1+RA2	0.154	0.152	1.01 (0.92-1.12)	0.76		
			RA2	0.155	0.155	1.00 (0.88-1.15)	0.97	1	0.89
rs16957854	T10977824C	intron	RA1+RA2	0.099	0.106	0.93 (0.83-1.05)	0.24		
			RA2	0.106	0.110	0.95 (0.82-1.12)	0.57	1	0.89
rs8055893	T10979296C	intron	RA1+RA2	0.120	0.124	0.96 (0.86-1.08)	0.52		
			RA1	0.121	0.112	1.10 (0.87-1.38)	0.43	2	0.25
			RA2	0.116	0.125	0.93 (0.80-1.08)	0.32	1	0.89
rs12708713	T10979960C	intron	RA1+RA2	0.171	0.173	0.99 (0.90-1.08)	0.76		
			RA2	0.172	0.175	0.98 (0.86-1.11)	0.74	1	0.89
rs16957864	T10983777C	intron	RA1+RA2	0.071	0.072	0.98 (0.85-1.13)	0.79		
			RA1	0.060	0.077	0.77 (0.57-1.03)	0.08	3	$4.5x10^{-2}$
rs7201845	G10985839A	intron	RA1+RA2	0.214	0.211	1.02 (0.94-1.11)	0.63		
			RA1	0.221	0.209	1.07 (0.89-1.28)	0.48	3	$4.5x10^{-2}$
rs17803907	A10985998G	intron	RA1+RA2	0.023	0.016	1.50 (1.15-1.97)	$2.9x10^{-3}$		
			RA1	0.037	0.031	1.17 (0.78-1.75)	0.45		
rs7202408	A10986116G	intron	RA1+RA2	0.121	0.124	0.97 (0.87-1.08)	0.59		
rs9652599	A10986662G	intron	RA1+RA2	0.215	0.211	1.02 (0.94-1.12)	0.58		
rs12448611	C10987952T	intron	RA1+RA2	0.485	0.475	1.04 (0.97-1.12)	0.27		
rs7197758	A10988583G	intron	RA1+RA2	0.300	0.314	0.94 (0.87-1.01)	0.11		
			RA1	0.293	0.308	0.93 (0.79-1.09)	0.38	3	$4.5x10^{-2}$
rs6498142	C10988750G	intron	RA1+RA2	0.181	0.181	1.00 (0.91-1.09)	0.94		
			RA2	0.180	0.184	0.97 (0.86-1.10)	0.67	1	0.89
rs16957872	T10990193C	intron	RA1+RA2	0.119	0.124	0.96 (0.86-1.07)	0.46		
rs4781031	T10990369C	intron	RA1+RA2	0.247	0.254	0.97 (0.89-1.06)	0.46		
rs8055968	A10992204G	intron	RA1+RA2	0.215	0.211	1.02 (0.94-1.12)	0.60		
rs7403919	C10993469T	intron	RA1+RA2	0.353	0.357	0.98 (0.91-1.06)	0.64		
			RA1	0.356	0.345	1.05 (0.90-1.22)	0.56	4	0.60
rs4781033	C10994403T	intron	RA1+RA2	0.493	0.482	1.05 (0.97-1.12)	0.21		
rs4781034	A10994441G	intron	RA1+RA2	0.169	0.181	0.92 (0.84-1.01)	0.09		
rs16957883	T10996146G	intron	RA1+RA2	0.120	0.124	0.96 (0.86-1.07)	0.49		
rs13339285	T10996195C	intron	RA1+RA2	0.206	0.203	1.02 (0.93-1.11)	0.73		
rs13330041	A10996309G	intron	RA1+RA2	0.206	0.203	1.02 (0.93-1.11)	0.70		
			RA1	0.211	0.203	1.05 (0.88-1.26)	0.61	4	0.60
rs9935445	T10996890C	intron	RA1+RA2	0.493	0.482	1.05 (0.97-1.12)	0.21		
rs8050144	C10997588G	intron	RA1+RA2	0.421	0.408	1.05 (0.98-1.13)	0.17		
rs16957894	G10997859A	intron	RA1+RA2	0.119	0.124	0.96 (0.86-1.07)	0.46		
rs16957895	C10997912T	intron	RA1+RA2	0.299	0.313	0.93 (0.86-1.01)	0.08		
rs8054758	G10998627C	intron	RA1+RA2	0.493	0.482	1.05 (0.97-1.13)	0.20		
rs8055544	G10999062T	intron	RA1+RA2	0.421	0.411	1.04 (0.97-1.12)	0.26		
			RA2	0.419	0.413	1.03 (0.93-1.13)	0.63	1	0.89
rs8062322	A10999820C	intron	RA1+RA2	0.292	0.298	0.97 (0.90-1.05)	0.51		
			RA1	0.297	0.292	1.02 (0.87-1.20)	0.79	5	0.17
rs8061306	A10999979G	intron	RA1+RA2	0.176	0.189	0.92 (0.84-1.01)	0.08		
rs9926367	C11000680T	intron	RA1+RA2	0.292	0.299	0.97 (0.90-1.05)	0.46		
rs16957899	G11000775A	intron	RA1+RA2	0.174	0.184	0.94 (0.85-1.03)	0.16		
			RA2	0.181	0.189	0.94 (0.83-1.07)	0.37	1	0.89
rs9934231	G11000822C	intron	RA1+RA2	0.474	0.491	0.94 (0.87-1.01)	0.07		
			RA2	0.469	0.493	0.91 (0.83-1.00)	0.06		
rs13337334	C11001919G	intron	RA1+RA2	0.278	0.274	1.02 (0.94-1.11)	0.61		
rs7200623	G11002546C	intron	RA1+RA2	0.120	0.125	0.96 (0.86-1.07)	0.44		
rs9940155	G11003007C	intron	RA1+RA2	0.106	0.100	1.07 (0.95-1.20)	0.27		
rs9925481	T11004980C	intron	RA1+RA2	0.072	0.071	1.02 (0.89-1.17)	0.75		
			RA2	0.078	0.074	1.07 (0.89-1.28)	0.52		
rs11074944	A11005054G	intron	RA1+RA2	0.073	0.072	1.01 (0.88-1.16)	0.91		
rs723586	G11005880C	intron	RA1+RA2	0.276	0.271	1.03 (0.95-1.11)	0.54		
rs7192695	T11006753C	intron	RA1+RA2	0.124	0.128	0.96 (0.86-1.07)	0.49		
rs7194305	G11007208A	intron	RA1+RA2	0.402	0.401	1.00 (0.93-1.08)	0.90		
rs17804470	C11007469G	intron	RA1+RA2	0.402	0.401	1.00 (0.93-1.08)	0.91		
rs7199305	G11007834C	intron	RA1+RA2	0.124	0.128	0.96 (0.86-1.07)	0.48		

rs9926862	T11010932G	intron	RA1+RA2	0.107	0.099	1.09 (0.97-1.22)	0.16		
			RA2	0.112	0.095	1.21 (1.03-1.42)	$2.2x10^{-2}$		
rs12934193	C11011226T	intron	RA1+RA2	0.167	0.169	0.99 (0.90-1.09)	0.86		
			RA1	0.179	0.170	1.06 (0.87-1.29)	0.56	5	0.17
rs9932895	G11012539C	intron	RA1+RA2	0.108	0.101	1.07 (0.96-1.21)	0.22		
rs7186166	G11012935C	intron	RA1+RA2	0.198	0.200	0.99 (0.90-1.08)	0.75		
rs9925833	C11013255T	intron	RA1+RA2	0.073	0.072	1.02 (0.89-1.17)	0.76		
rs11074945	A11013559G	intron	RA1+RA2	0.043	0.034	1.26 (1.00-1.58)	$4.2x10^{-2}$		
			RA1	0.098	0.073	1.37 (1.05-1.79)	$2.1x10^{-2}$	5	0.17
rs6498146	T11014208C	intron	RA1+RA2	0.195	0.196	0.99 (0.90-1.08)	0.83		
			RA2	0.208	0.205	1.01 (0.90-1.14)	0.85	2	0.58
rs7203687	C11014421T	intron	RA1+RA2	0.198	0.200	0.99 (0.90-1.08)	0.79		
rs9935174	T11018848C	intron	RA1+RA2	0.401	0.400	1.01 (0.93-1.08)	0.89		
			RA1	0.402	0.403	0.99 (0.85-1.15)	0.93	5	0.17
rs6498148	C11019732A	intron	RA1+RA2	0.124	0.128	0.96 (0.86-1.07)	0.48		
rs2286975	A11021507G	intron	RA1+RA2	0.277	0.271	1.03 (0.95-1.12)	0.47		
			RA1	0.275	0.278	0.99 (0.84-1.16)	0.87		
rs1003603	G11022124A	intron	RA1+RA2	0.401	0.400	1.00 (0.93-1.08)	0.93		
			RA1	0.402	0.402	1.00 (0.86-1.16)	0.98	6	0.58
rs1985372	T11022340C	intron	RA1+RA2	0.402	0.400	1.00 (0.93-1.08)	0.90		
rs1861548	C11026001G	intron	RA1+RA2	0.402	0.400	1.00 (0.93-1.08)	0.89		
rs8045749	G11027154A	intron	RA1+RA2	0.402	0.400	1.00 (0.93-1.08)	0.90		
rs7194545	T11031797C	intron	RA1+RA2	0.183	0.187	0.97 (0.89-1.07)	0.57		
			RA1	0.176	0.187	0.93 (0.77-1.12)	0.46	6	0.58
rs17804800	C11032794T	intron	RA2	0.054	0.049	1.10 (0.88-1.37)	0.40		
rs11074946	A11033404G	intron	RA1+RA2	0.069	0.067	1.03 (0.90-1.19)	0.65		
			RA2	0.078	0.069	1.15 (0.95-1.38)	0.16	2	0.58
rs7196077	C11034556T	intron	RA1+RA2	0.092	0.094	0.98 (0.86-1.11)	0.70		
			RA2	0.116	0.125	0.92 (0.79-1.07)	0.28	2	0.58
rs11860777	A11035076C	intron	RA1+RA2	0.087	0.088	1.00 (0.87-1.14)	0.99		
rs9923455	C11039365G	intron	RA1+RA2	0.388	0.386	1.01 (0.94-1.09)	0.82		
rs16957957	T11039653C	intron	RA1+RA2	0.091	0.093	0.98 (0.86-1.11)	0.73		
			RA2	0.109	0.116	0.93 (0.80-1.09)	0.36	2	0.58
rs9932114	A11042404G	intron	RA1+RA2	0.070	0.069	1.02 (0.89-1.18)	0.73		
			RA2	0.078	0.071	1.10 (0.92-1.33)	0.33	3	0.57
rs7185202	C11042565G	intron	RA1+RA2	0.437	0.428	1.03 (0.96-1.11)	0.39		
			RA2	0.421	0.415	1.02 (0.93-1.13)	0.65	3	0.57
rs8052325	G11045479A	intron	RA1+RA2	0.118	0.123	0.96 (0.86-1.07)	0.46		
rs8056098	A11046313G	intron	RA1+RA2	0.407	0.409	0.99 (0.92-1.07)	0.87		
			RA1	0.397	0.401	0.99 (0.85-1.15)	0.85	6	0.58
rs9940096	G11047095A	intron	RA1+RA2	0.118	0.122	0.96 (0.86-1.07)	0.48		
			RA1	0.113	0.115	0.97 (0.77-1.23)	0.83	7	0.53
rs16957966	C11047638T	intron	RA1+RA2	0.119	0.123	0.96 (0.86-1.07)	0.50		
rs11859648	G11048153A	intron	RA1+RA2	0.109	0.105	1.04 (0.92-1.17)	0.52		
rs16957976	C11048822T	intron	RA1+RA2	0.104	0.100	1.05 (0.94-1.19)	0.37		
			RA1	0.102	0.117	0.86 (0.68-1.09)	0.23	7	0.53
			RA2	0.115	0.102	1.14 (0.97-1.33)	0.10	4	0.40
rs11865480	A11049503G	intron	RA1+RA2	0.075	0.076	0.99 (0.87-1.14)	0.95		
rs17230818	A11049638T	intron	RA1+RA2	0.213	0.209	1.02 (0.93-1.11)	0.68		
rs3901386	C11050221T	intron	RA1+RA2	0.386	0.382	1.02 (0.94-1.10)	0.64		
			RA1	0.404	0.410	0.98 (0.84-1.13)	0.77	7	0.53
			RA2	0.409	0.407	1.01 (0.91-1.11)	0.87	4	0.40
rs876457	A11051227G	intron	RA1+RA2	0.073	0.075	0.97 (0.85-1.12)	0.69		
rs9302458	G11051710C	intron	RA1+RA2	0.118	0.122	0.96 (0.86-1.07)	0.47		
			RA2	0.120	0.127	0.94 (0.81-1.09)	0.43	4	0.40
rs16957984	C11052463T	intron	RA1+RA2	0.119	0.123	0.96 (0.86-1.07)	0.50		
rs741177	T11053149C	intron	RA1+RA2	0.119	0.123	0.96 (0.86-1.07)	0.50		
rs9939298	C11053220T	intron	RA1+RA2	0.110	0.106	1.04 (0.93-1.16)	0.53		
rs741176	G11053311A	intron	RA1+RA2	0.119	0.122	0.96 (0.86-1.08)	0.51		
			RA1	0.114	0.119	0.95 (0.76-1.19)	0.67	7	0.53
rs11863415	G11054595C	intron	RA1+RA2	0.118	0.122	0.96 (0.86-1.07)	0.47		
			RA2	0.121	0.127	0.94 (0.82-1.09)	0.45	4	0.40

rs12102345	G11055221A	intron	RA1+RA2	0.120	0.125	0.96 (0.86-1.07)	0.47		
rs876476	A11057749G	intron	RA1+RA2	0.288	0.286	1.01 (0.93-1.09)	0.81		
			RA1	0.288	0.286	1.01 (0.86-1.19)	0.92	8	0.07
rs2286973	A11062271G	syn-cod[1]	RA1+RA2	0.409	0.412	0.99 (0.92-1.06)	0.74		
rs2286972	G11062499T	intron	RA1+RA2	0.125	0.129	0.96 (0.86-1.08)	0.51		
rs2160042	A11065733G	missense	RA1+RA2	0.115	0.120	0.95 (0.85-1.07)	0.38		
rs887864	G11066386A	intron	RA1+RA2	0.342	0.345	0.99 (0.91-1.07)	0.73		
rs741175	C11067186T	intron	RA1+RA2	0.421	0.423	0.99 (0.92-1.07)	0.86		
rs741174	T11067339C	intron	RA1+RA2	0.421	0.423	0.99 (0.92-1.07)	0.85		
rs741173	A11067420C	intron	RA1+RA2	0.421	0.423	0.99 (0.92-1.07)	0.84		
rs8061043	T11068430G	intron	RA1+RA2	0.073	0.076	0.95 (0.82-1.10)	0.48		
rs8062923	A11068467C	intron	RA1+RA2	0.245	0.244	1.00 (0.92-1.09)	0.94		
rs4781035	A11068679G	intron	RA1+RA2	0.210	0.214	0.98 (0.89-1.07)	0.60		
rs7195452	G11070827C	intron	RA1+RA2	0.082	0.081	1.02 (0.89-1.17)	0.81		
rs7200940	G11072068C	intron	RA1+RA2	0.316	0.319	0.98 (0.91-1.06)	0.66		
rs11860603	C11072518T	intron	RA1+RA2	0.306	0.309	0.99 (0.91-1.07)	0.73		
rs11865121	A11074189C	intron	RA1+RA2	0.306	0.309	0.99 (0.91-1.07)	0.73		
rs7198621	G11074959C	intron	RA1+RA2	0.306	0.309	0.99 (0.91-1.07)	0.74		
			RA1	0.201	0.182	1.13 (0.94-1.36)	0.21	8	0.07
rs725613	G11077184T	intron	RA1+RA2	0.340	0.339	1.00 (0.93-1.08)	0.92		
			RA1	0.370	0.341	1.13 (0.97-1.32)	0.12	8	0.07
rs16958021	G11077548A	intron	RA1+RA2	0.082	0.081	1.02 (0.89-1.17)	0.84		
			RA1	0.120	0.119	1.01 (0.80-1.26)	0.96	8	0.07
rs12444495	T11077956C	intron	RA1+RA2	0.209	0.212	0.98 (0.90-1.08)	0.77		
rs7184491	C11078262T	intron	RA1+RA2	0.232	0.235	0.98 (0.89-1.07)	0.67		
rs12925642	G11079103A	intron	RA1+RA2	0.306	0.309	0.99 (0.91-1.07)	0.74		
rs9652600	A11081515G	intron	RA1+RA2	0.090	0.089	1.01 (0.89-1.16)	0.83		
			RA1	0.121	0.116	1.05 (0.83-1.32)	0.72	8	0.07
rs9652601	A11081866G	intron	RA1+RA2	0.314	0.318	0.98 (0.91-1.06)	0.67		
rs9652582	A11082065G	intron	RA1+RA2	0.309	0.313	0.98 (0.91-1.06)	0.68		
rs2041670	A11082153G	intron	RA1+RA2	0.314	0.318	0.98 (0.91-1.06)	0.62		
			RA1	0.322	0.301	1.10 (0.94-1.29)	0.24	8	0.07
rs16958028	T11084169A	intron	RA1+RA2	0.067	0.068	0.98 (0.85-1.14)	0.84		
rs7200786	A11085302G	intron	RA1	0.470	0.453	1.07 (0.93-1.24)	0.36	8	0.07
rs12932833	G11085571C	intron	RA1+RA2	0.202	0.203	0.99 (0.91-1.09)	0.92		
rs9929994	G11085746A	intron	RA1+RA2	0.348	0.347	1.01 (0.93-1.08)	0.87		
rs12708716	G11087374A	intron	RA1+RA2	0.342	0.341	1.00 (0.93-1.09)	0.90		
			RA2	0.343	0.350	0.97 (0.87-1.07)	0.51	4	0.40
rs12926153	G11088263C	intron	RA1+RA2	0.093	0.092	1.01 (0.89-1.16)	0.84		
rs9888908	A11088745C	intron	RA1+RA2	0.314	0.318	0.98 (0.91-1.06)	0.67		
rs7204099	C11089257T	intron	RA1+RA2	0.348	0.347	1.01 (0.93-1.08)	0.87		
rs11642009	G11090394T	intron	RA1+RA2	0.266	0.258	1.04 (0.95-1.14)	0.39		
rs11861236	C11091127T	intron	RA1+RA2	0.424	0.424	1.00 (0.93-1.08)	0.93		
rs17805769	G11093374A	intron	RA1+RA2	0.424	0.424	1.00 (0.93-1.08)	0.93		
rs12448240	C11094719T	intron	RA1+RA2	0.015	0.010	1.49 (1.07-2.09)	$2.1x10^{-2}$		
rs16958033	T11094916C	intron	RA1+RA2	0.108	0.104	1.05 (0.93-1.18)	0.41		
rs16958036	C11095237T	intron	RA1+RA2	0.108	0.104	1.05 (0.93-1.18)	0.42		
rs12924729	A11095284G	intron	RA1+RA2	0.318	0.321	0.98 (0.91-1.06)	0.69		
			RA2	0.315	0.327	0.95 (0.85-1.05)	0.31	4	0.40
rs12917656	C11095363T	intron	RA1+RA2	0.424	0.424	1.00 (0.93-1.08)	0.93		
rs12919083	C11096431A	intron	RA1+RA2	0.316	0.319	0.98 (0.91-1.06)	0.68		
rs12917716	C11096649G	intron	RA1+RA2	0.424	0.424	1.00 (0.93-1.08)	0.93		
rs7205916	A11096757C	intron	RA1+RA2	0.091	0.090	1.00 (0.88-1.14)	0.97		
rs12599402	C11097389T	intron	RA1+RA2	0.423	0.423	1.00 (0.93-1.07)	0.99		
			RA1	0.422	0.417	1.02 (0.88-1.19)	0.80	8	0.07
rs11649025	C11098431T	intron	RA1+RA2	0.107	0.104	1.04 (0.93-1.17)	0.48		
			RA1	0.088	0.112	0.77 (0.60-0.99)	$4.9x10^{-2}$	8	0.07
rs16958051	T11098720A	intron	RA1+RA2	0.091	0.091	1.00 (0.88-1.14)	0.97		
rs12928537	A11098901G	intron	RA1+RA2	0.316	0.319	0.99 (0.91-1.06)	0.73		
rs8061826	G11100288A	intron	RA1+RA2	0.424	0.423	1.00 (0.93-1.08)	0.92		
rs3893661	G11101381C	intron	RA1+RA2	0.424	0.423	1.00 (0.93-1.08)	0.93		
rs3893660	G11101431A	intron	RA1+RA2	0.422	0.422	1.00 (0.93-1.07)	0.99		

			RA2	0.426	0.426	1.00 (0.91-1.11)	0.98	4	0.40
rs3862468	C11101519G	intron	RA1+RA2	0.424	0.423	1.00 (0.93-1.08)	0.93		
rs3862469	T11101581C	intron	RA1+RA2	0.330	0.328	1.01 (0.93-1.09)	0.87		
rs12927355	T11102272C	intron	RA1+RA2	0.314	0.318	0.98 (0.91-1.06)	0.66		
rs9941107	A11103542G	intron	RA1+RA2	0.423	0.423	1.00 (0.93-1.07)	0.96		
			RA2	0.422	0.420	1.01 (0.91-1.11)	0.90	4	0.40
rs12929596	C11106853T	intron	RA1+RA2	0.094	0.093	1.02 (0.89-1.16)	0.80		
rs998592	T11107179C	intron	RA1+RA2	0.423	0.423	1.00 (0.93-1.07)	0.95		
			RA1	0.413	0.411	1.01 (0.87-1.17)	0.94	8	0.07
rs17806299	A11107481G	intron	RA1+RA2	0.201	0.203	0.99 (0.90-1.09)	0.88		
			RA1	0.204	0.182	1.15 (0.95-1.38)	0.16	8	0.07
rs9933507	C11108929T	intron	RA1+RA2	0.421	0.423	0.99 (0.92-1.07)	0.88		
			RA1	0.421	0.419	1.01 (0.87-1.17)	0.96	8	0.07
rs9926078	C11111066G	intron	RA1+RA2	0.422	0.423	1.00 (0.93-1.07)	0.93		
rs12103174	G11111231A	intron	RA1+RA2	0.389	0.386	1.01 (0.94-1.09)	0.73		
			RA1	0.415	0.415	1.00 (0.86-1.16)	1.00	8	0.07
rs767448	G11111722A	intron	RA1+RA2	0.422	0.423	1.00 (0.93-1.07)	0.94		
rs11644969	A11113265G	intron	RA1+RA2	0.073	0.074	0.97 (0.85-1.12)	0.71		
rs7198004	G11115118A	intron	RA1+RA2	0.422	0.423	1.00 (0.93-1.07)	0.94		
			RA2	0.424	0.424	1.00 (0.91-1.10)	1.00	4	0.40
rs7203150	C11115223T	intron	RA1+RA2	0.422	0.423	1.00 (0.93-1.07)	0.94		
			RA2	0.423	0.426	0.99 (0.90-1.09)	0.84	4	0.40
rs9746695	C11115395T	intron	RA1+RA2	0.314	0.319	0.98 (0.91-1.06)	0.61		
			RA2	0.311	0.325	0.93 (0.84-1.04)	0.21	5	$2.4x10^{-2}$
rs11647011	G11115591A	intron	RA1+RA2	0.107	0.104	1.04 (0.92-1.17)	0.52		
			RA2	0.113	0.100	1.15 (0.98-1.34)	0.09	5	$2.4x10^{-2}$
rs12924985	C11115823G	intron	RA1+RA2	0.422	0.423	1.00 (0.93-1.07)	0.93		
rs12917947	G11117744C	intron	RA1+RA2	0.086	0.087	0.98 (0.86-1.12)	0.80		
rs9923856	C11117916T	intron	RA1+RA2	0.352	0.353	1.00 (0.92-1.09)	0.97		
			RA1	0.299	0.295	1.02 (0.87-1.20)	0.83		
rs12935413	A11117948G	intron	RA1+RA2	0.342	0.341	1.00 (0.93-1.08)	0.93		
rs9931657	T11118094C	intron	RA1+RA2	0.069	0.070	0.98 (0.85-1.13)	0.80		
rs2080272	A11119054G	intron	RA1+RA2	0.344	0.334	1.04 (0.97-1.13)	0.26		
			RA1	0.354	0.350	1.02 (0.87-1.19)	0.84	9	0.08
rs1861198	C11119084T	intron	RA1+RA2	0.344	0.335	1.04 (0.96-1.12)	0.31		
rs9806963	T11119640C	intron	RA1+RA2	0.301	0.310	0.96 (0.89-1.04)	0.31		
rs9927527	G11120182A	intron	RA1+RA2	0.342	0.342	1.00 (0.92-1.08)	0.98		
rs12925474	G11122156T	intron	RA1+RA2	0.055	0.056	0.97 (0.82-1.14)	0.73		
			RA1	0.078	0.092	0.84 (0.64-1.09)	0.19	9	0.08
rs17604868	T11122528G	intron	RA1+RA2	0.337	0.340	0.99 (0.92-1.07)	0.80		
			RA1	0.326	0.352	0.89 (0.76-1.04)	0.15	9	0.08
rs10852330	A11123559G	intron	RA1+RA2	0.348	0.339	1.04 (0.96-1.12)	0.31		
rs11074952	A11126107G	intron	RA1+RA2	0.327	0.336	0.96 (0.88-1.04)	0.30		
			RA1	0.271	0.315	0.81 (0.69-0.95)	$1x10^{-2}$	10	$2.5x10^{-2}$
rs17604903	A11126530G	intron	RA1+RA2	0.035	0.028	1.24 (1.01-1.53)	$3.9x10^{-2}$		
rs12935657	A11126542G	intron	RA1+RA2	0.248	0.244	1.02 (0.94-1.11)	0.62		
rs2003400	A11126822G	intron	RA1+RA2	0.326	0.336	0.96 (0.89-1.04)	0.32		
rs2241099	G11132565C	intron	RA1+RA2	0.252	0.246	1.03 (0.95-1.12)	0.47		
rs2867879	A11134257G	intron	RA1+RA2	0.305	0.312	0.97 (0.89-1.05)	0.44		
			RA2	0.315	0.309	1.03 (0.93-1.14)	0.63	6	0.59
rs7184083	A11135415G	intron	RA1+RA2	0.344	0.338	1.03 (0.95-1.11)	0.47		
			RA2	0.344	0.342	1.01 (0.91-1.12)	0.85	6	0.59
rs2041733	A11137090G	intron	RA1	0.448	0.456	0.97 (0.83-1.12)	0.67	10	$2.5x10^{-2}$
rs7203459	C11138204T	intron	RA1+RA2	0.256	0.249	1.03 (0.95-1.13)	0.44		
rs2867880	A11139358G	intron	RA1+RA2	0.342	0.340	1.01 (0.94-1.09)	0.79		
rs9923175	T11140048G	intron	RA1+RA2	0.389	0.401	0.95 (0.88-1.02)	0.18		
rs16958089	T11140680C	intron	RA1+RA2	0.075	0.078	0.95 (0.83-1.09)	0.45		
rs17684919	T11142346C	intron	RA1+RA2	0.090	0.093	0.97 (0.85-1.10)	0.65		
			RA2	0.103	0.094	1.11 (0.94-1.31)	0.23	6	0.59
rs1078328	T11144685A	intron	RA1+RA2	0.343	0.336	1.03 (0.95-1.11)	0.45		
rs2903692	A11146284G	intron	RA1+RA2	0.338	0.335	1.02 (0.94-1.10)	0.67		
			RA1	0.312	0.312	1.00 (0.85-1.18)	0.97		

rs12917893	T11147479A	intron	RA1+RA2	0.342	0.340	1.01 (0.93-1.09)	0.81		
rs11863016	T11148059G	intron	RA1+RA2	0.167	0.179	0.92 (0.83-1.02)	0.10		
rs13331231	G11148504A	intron	RA1+RA2	0.167	0.179	0.92 (0.83-1.01)	0.10		
rs17673553	G11149407A	intron	RA1+RA2	0.251	0.243	1.04 (0.96-1.14)	0.35		
			RA1	0.258	0.218	1.25 (1.05-1.49)	$1.3x10^{-2}$	10	$2.5x10^{-2}$
rs12927046	T11149692G	intron	RA1+RA2	0.071	0.075	0.95 (0.82-1.09)	0.48		
rs794423	A11149998C	intron	RA1+RA2	0.039	0.037	1.05 (0.87-1.28)	0.61		
rs794424	G11150315A	intron	RA1+RA2	0.342	0.336	1.02 (0.95-1.11)	0.53		
rs794425	G11150348A	intron	RA1+RA2	0.342	0.336	1.02 (0.95-1.11)	0.53		
rs794426	C11150359A	intron	RA1+RA2	0.342	0.336	1.02 (0.95-1.11)	0.53		
rs11864680	G11151695T	intron	RA1+RA2	0.250	0.242	1.05 (0.96-1.14)	0.28		
rs7185300	G11151739A	intron	RA1+RA2	0.036	0.030	1.20 (0.98-1.47)	0.09		
			RA1	0.044	0.033	1.34 (0.91-1.96)	0.14		
rs7206912	G11152159C	intron	RA1+RA2	0.385	0.373	1.05 (0.98-1.14)	0.17		
rs9937607	C11152291T	intron	RA1+RA2	0.037	0.030	1.23 (1.00-1.50)	$4.7x10^{-2}$		
rs42469	T11152900C	intron	RA1+RA2	0.315	0.306	1.04 (0.96-1.13)	0.32		
rs17763452	G11153873A	intron	RA1+RA2	0.092	0.093	0.98 (0.87-1.11)	0.78		
rs7204935	A11155214C	intron	RA1+RA2	0.272	0.287	0.93 (0.85-1.01)	0.08		
rs171593	T11155872C	intron	RA1+RA2	0.342	0.336	1.03 (0.95-1.11)	0.48		
rs6498169	G11156830A	intron	RA1+RA2	0.336	0.329	1.03 (0.95-1.11)	0.45		
			RA2	0.349	0.346	1.01 (0.92-1.12)	0.81	6	0.59
rs28087	C11160330T	intron	RA2	0.338	0.338	1.00 (0.90-1.11)	1.00	6	0.59
rs27838	G11160984A	intron	RA1+RA2	0.338	0.332	1.03 (0.95-1.11)	0.44		
rs27839	T11161176C	intron	RA1+RA2	0.338	0.332	1.03 (0.95-1.11)	0.44		
rs41367	G11161832A	intron	RA1+RA2	0.338	0.331	1.03 (0.96-1.11)	0.42		
rs8063318	C11164093G	intron	RA1+RA2	0.383	0.373	1.04 (0.97-1.12)	0.27		
rs767019	T11164437G	intron	RA1+RA2	0.442	0.440	1.01 (0.93-1.09)	0.86		
			RA2	0.431	0.445	0.95 (0.86-1.04)	0.27	7	0.47
rs27908	A11164602G	intron	RA1+RA2	0.351	0.341	1.04 (0.97-1.13)	0.28		
			RA2	0.341	0.339	1.01 (0.91-1.12)	0.89	7	0.47
rs27836	A11164620G	intron	RA1+RA2	0.340	0.334	1.03 (0.95-1.11)	0.46		
rs41370	G11164830A	intron	RA1+RA2	0.340	0.334	1.03 (0.95-1.11)	0.46		
rs11641347	A11165234C	intron	RA1+RA2	0.013	0.011	1.18 (0.84-1.65)	0.34		
rs11643123	G11167941A	intron	RA1+RA2	0.109	0.107	1.02 (0.91-1.15)	0.76		
rs11640376	T11169778G	intron	RA1+RA2	0.107	0.105	1.02 (0.91-1.16)	0.69		
rs27965	G11170284A	intron	RA1+RA2	0.365	0.362	1.01 (0.94-1.09)	0.73		
rs16958108	G11171357T	intron	RA1+RA2	0.109	0.104	1.06 (0.93-1.20)	0.38		
			RA1	0.112	0.117	0.95 (0.75-1.20)	0.69	11	0.84
rs42369	G11172113A	intron	RA1+RA2	0.370	0.366	1.02 (0.93-1.11)	0.73		
			RA1	0.371	0.370	1.00 (0.86-1.17)	0.96	11	0.84
			RA2	0.363	0.359	1.02 (0.92-1.13)	0.75		
rs794428	A11175932G	intron	RA1+RA2	0.353	0.353	1.00 (0.92-1.08)	0.93		
			RA1	0.365	0.367	0.99 (0.85-1.15)	0.89	11	0.84
rs3960630	A11178405C	intron	RA1+RA2	0.219	0.221	0.99 (0.90-1.09)	0.85		
			RA1	0.209	0.245	0.81 (0.68-0.97)	$2.5x10^{-2}$	12	0.11
rs11647285	A11180073G	utr-3[2]	RA1+RA2	0.039	0.033	1.19 (0.97-1.44)	0.10		
rs2040	T11183414C	utr-3[2]	RA1+RA2	0.091	0.092	0.99 (0.87-1.12)	0.86		
			RA1	0.080	0.097	0.81 (0.62-1.05)	0.12	12	0.11

[1]Synonymous coding
[2]Untranslated region

Figure 1. Schematic of our analysis strategy in stages (**a**) 1, (**b**) 2 and (**c**) 3.

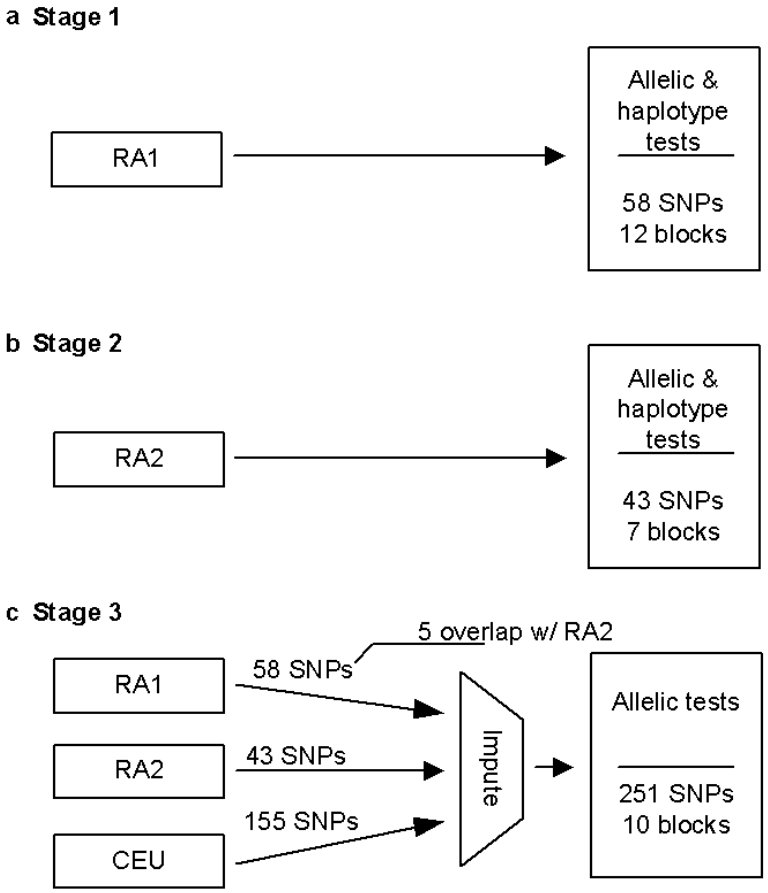
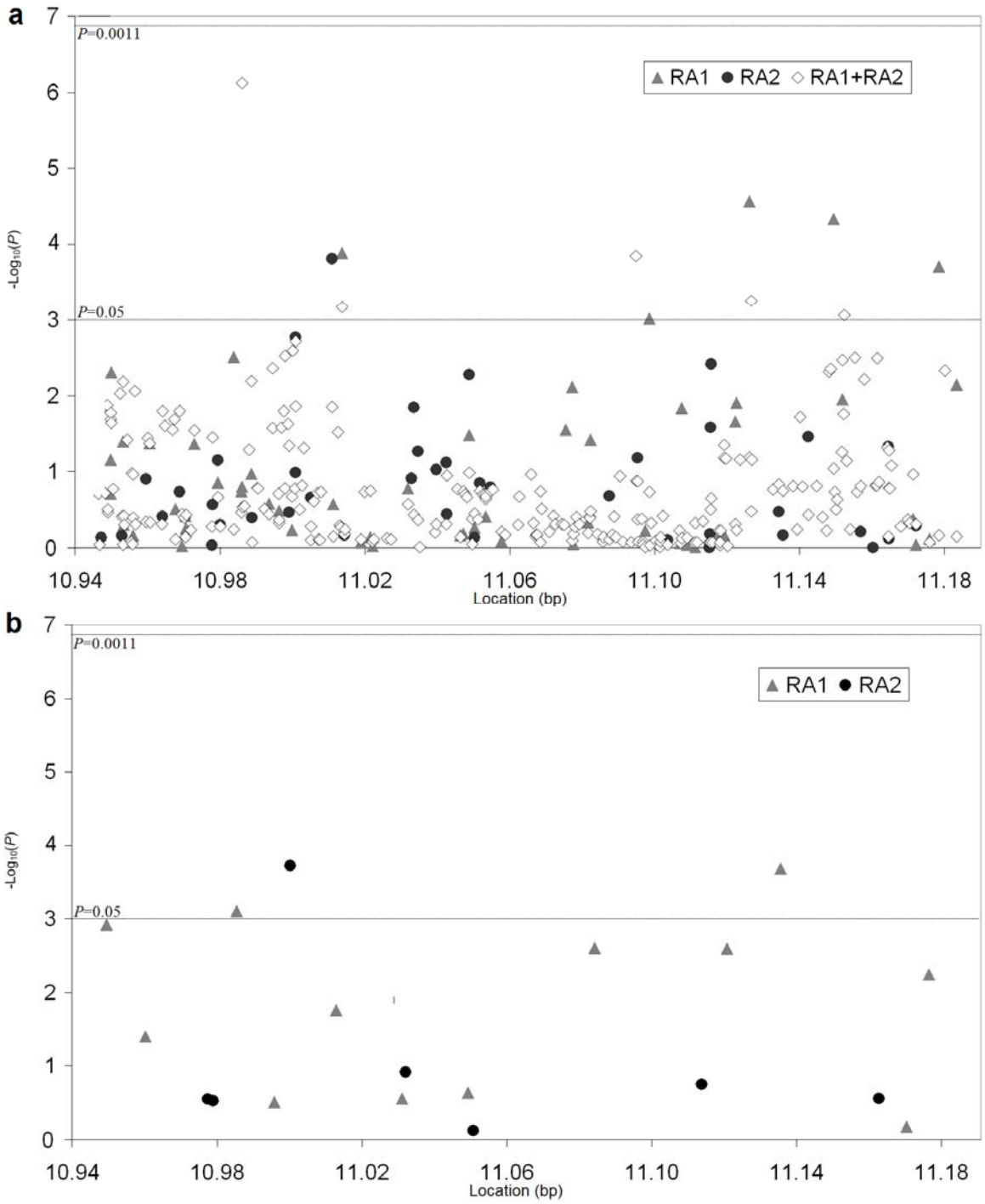


Figure 2. *P*-values from (**a**) allelic and (**b**) haplotype tests of *CLEC16A* SNPs in RA.



CHAPTER SEVEN

Analysis of Maternal-Offspring HLA Compatibility, Parent-of-Origin and Noninherited Effects for *HLA-DRB1* in Systemic Lupus Erythematosus

Published: Bronson PG, Komorowski LK, Ramsay PP, May S, Noble J, Lane JA, Thomson G, Claas FH, Seldin MF, Kelly JA, Harley JB, Moser KL, Gaffney PM, Behrens T, Criswell LA, Barcellos LF. (2010) Analysis of maternal-offspring HLA compatibility, parent-of-origin and noninherited effects for HLA-DRB1 in systemic lupus erythematosus. Arthritis Rheum 62(6):1712-7.

ABSTRACT

Genetic susceptibility to systemic lupus erythematosus (SLE) is well-established, with HLA class II *DRB1* and *DQB1* loci demonstrating the strongest association. However, HLA may also influence SLE through novel biological mechanisms, in addition to genetic transmission of risk alleles. Evidence for increased maternal-offspring HLA class II compatibility in SLE and differences in maternal vs. paternal transmission rates (parent-of-origin effects) and nontransmission rates (noninherited maternal antigen [NIMA] effects) in other autoimmune diseases have been reported. Thus, we investigated maternal-offspring HLA compatibility, parent-of-origin and NIMA effects at *DRB1* in SLE. The cohort consisted of 707 SLE families and 188 independent healthy maternal-offspring pairs (N = 2,497 individuals). Family-based association tests were conducted to compare transmitted vs. nontransmitted alleles (TDT) and both maternal vs. paternal transmitted (parent-of-origin) and nontransmitted alleles (NIMA) (χ^2 test of heterogeneity). Analyses were stratified by offspring gender. Maternal-affected offspring *DRB1* compatibility in SLE families was compared to paternal-affected offspring compatibility and to independent control maternal-offspring pairs (Fisher's test), and restricted to male and nulligravid female SLE offspring. As expected, *DRB1* was associated with SLE in the overall ($P < 1x10^{-4}$) and gender-stratified analyses (females: $P < 1x10^{-4}$; males: $P < 0.05$). However, mothers of SLE patients had similar transmission and nontransmission frequencies for *DRB1* alleles when compared to fathers, including for known SLE risk alleles *DRB1*0301*, **1501* and **0801*. There was no association between maternal-offspring compatibility and SLE. In conclusion, maternal-offspring HLA compatibility, parent-of-origin and NIMA effects at *DRB1* are unlikely to play a role in SLE.

BACKGROUND

Systemic lupus erythematosus (SLE) is a systemic autoimmune disease characterized by autoantibodies to nuclear and cell surface antigens. Although the etiology remains unknown, evidence for genetic susceptibility is well established. The human leukocyte antigen (HLA) class II alleles *DRB1*1501*, **0301* and **0801* and class I alleles *A*01* and *B*08* in the major histocompatibility complex (MHC) region are consistently associated with SLE.

HLA loci may also influence SLE through additional inherited or noninherited mechanisms. Differences in maternal and paternal transmission rates, or parent-of-origin effects, have not been previously examined in SLE. One potential mechanism influencing disease susceptibility is 'genomic imprinting', due to epigenetic modification of the genome. This modification results in unequal transcription of parental alleles and subsequent allele expression, depending on whether alleles were transmitted maternally or paternally.

Increased HLA compatibility between a mother and her offspring is hypothesized to contribute to risk for autoimmune disease. Maternal-offspring effects can present as excess HLA compatibility between the mother and affected offspring (Figure 1). In mice, HLA similarity between mother and fetus has been shown to promote the persistence of maternal cells in the fetus (maternal microchimerism) following pregnancy (1).

The developing immune system of the fetus is also directly exposed to noninherited maternal antigens (NIMA) in utero (2). Exposure to NIMA can have a lifelong influence on the immune system and has been theorized to tolerize or predispose to autoimmune reactions. A tolerogenic effect may explain the longer survival of renal transplants from sibling donors expressing NIMA vs. noninherited paternal HLA. Decreased B cell responses to HLA class I NIMA in humans have been reported (3). Recent data suggests that fetuses may also develop T cell tolerance to NIMA in utero through tolerogenic fetal regulatory T cells which are maintained throughout the lifetime (4).

This study tested hypotheses that the *DRB1* locus influences SLE through these novel biological mechanisms, in addition to genetic transmission of particular HLA risk alleles.

METHODS

Patients

The cohort consisted of 707 European-American trio families (N = 2,121 individuals) with one SLE-affected child and two parents and 188 Dutch healthy maternal-offspring pairs (N = 376 individuals; total N = 2,497). The child was female in 93% of the trios (N = 661) and 60% of the healthy maternal-offspring pairs (N = 111). All SLE patients met the American College of Rheumatology criteria (5). Families were enrolled through the University of California, San Francisco (UCSF) (N = 314), University of Minnesota (N = 233), and Oklahoma Medical Research Foundation (N = 160) as previously described (6). Ages of the cases ranged from 15 to 72 years (mean 45.4 ± 11.2 years) and the mean age-at-onset was 33.1 ± 13.2 years. 40% of SLE patients had renal disease or antibodies to double-stranded DNA. Data were collected by questionnaire and chart review. Parity data were collected by questionnaire for cases recruited

through UCSF, of which half were nulligravid (never pregnant at, or prior to, diagnosis). Independent controls consisted of healthy maternal-offspring pairs enrolled through Leiden University Medical Center as previously described, and also have been utilized to study NIMA in rheumatoid arthritis (RA) (7, 8). *DRB1* genotypes were generated using PCR-SSO methodology. Four-digit *DRB1* resolution was available for the SLE trios, and two-digit *DRB1* resolution was available for the healthy maternal-offspring pairs.

Although healthy maternal-offspring pairs were Dutch and SLE families were North American, *HLA-DRB1* allele frequencies derived from SLE families (non-transmitted alleles or 'controls') were statistically indistinguishable ($P = 0.2$) to those in Dutch control mothers, providing strong evidence that population frequencies for both groups were also very similar (data not shown). Further, the family-based nature of all analyses used in the current study greatly reduced the impact of population stratification on HLA compatibility, parent-of-origin and NIMA analyses.

Statistical analysis

The transmission disequilibrium test (TDT) (PLINK v1.02, http://pngu.mgh.harvard.edu/~purcell/plink) was used to examine differences between transmitted and nontransmitted *DRB1* alleles in affected offspring. For compatibility analyses based on *DRB1* genotypes, we defined maternal-offspring compatibility categorically: (1) unidirectional offspring-to-mother compatible, (2) unidirectional mother-to-offspring compatible, (3) bidirectional and (4) no increased compatibility (Figure 1). We created *2x2* contingency tables to test maternal-offspring *DRB1* compatibility with Fisher's exact test (R v2.6, http://cran.r-project.org) using paternal-offspring compatibility as controls. The test was conducted using histocompatibility estimates based on four-digit *DRB1* resolution (data not shown) and two-digit *DRB1* resolution. These analyses were repeated with any *DRB1* compatibility as the exposure (categories 1-3 above were combined). Similarly, we created *2x2* contingency tables to test maternal-offspring compatibility with Fisher's exact test using the *DRB1* compatibility from the independent healthy mother-offspring pairs as controls. Here, based on availability of data, two-digit *DRB1* resolution was used to determine histocompatibility estimates for all analyses. Male and nulligravid females were examined as a separate subgroup. Further, we restricted all above analyses to pairs where offspring were carrying SLE-associated *DRB1*03* or *DRB1*15* (for two-digit *DRB1* resolution), or *DRB1*0301* or *DRB1*1501* (for four-digit *DRB1* resolution) alleles (data not shown). Finally, we checked for an excess of *DRB1* homozygotes in the mothers of SLE patients by examining deviation from Hardy-Weinberg proportions in the pooled set of homozygous genotypes using the χ^2 goodness-of-fit test (PyPop v0.6, http://www.pypop.org).

For the parent-of-origin analyses, frequencies of maternal vs. paternal transmitted alleles were first derived from trio families using pedigree information and then compared using a χ^2 contingency table test for heterogeneity (AFBAC v1.13, http://www.pypop.org). The same test was used in the NIMA analyses to compare frequencies of maternal vs. paternal nontransmitted alleles. Both global and allele-specific analyses (when appropriate) were performed. Parent-of-origin and NIMA analyses were stratified by offspring gender.

Statistical power was estimated for parent-of-origin and NIMA analyses (Quanto v1.2.4, http://hydra.usc.edu/gxe/) assuming a two-sided 5% type I error rate and using control frequencies derived from paternal transmitted and non-transmitted, respectively, *DRB1*0301* and *DRB1*1501* frequencies.

RESULTS

As expected, *DRB1* was strongly associated with SLE ($P < 1x10^{-4}$). *DRB1*0301* exhibited the strongest association with SLE (odds ratio [OR] = 2.2, 95% confidence interval [CI] 1.8-2.7, $P = 9x10^{-14}$). *DRB1*1501* was also associated with SLE (OR = 1.4, 95% CI = 1.1-1.7, $P = 0.003$). Results for the previously identified SLE-associated *DRB1*0801* allele were not significant. *DRB1* compatibility was not associated with SLE in the overall sample, or in the subset restricted to males and nulligravid females (Table 1). Evidence for excess homozygosity at the *DRB1* locus in SLE mothers was not present.

There was no evidence for parent-of-origin and noninherited maternal HLA (NIMA) effects in SLE (data not shown), even when specific SLE risk alleles (*DRB1*1501*, **0301* and **0801*) were considered separately. Comparison of paternal vs. maternal transmitted and paternal vs. maternal nontransmitted SLE risk *DRB1* alleles in the overall dataset, and for male and female patients analyzed separately, demonstrated that frequencies did not differ (Figure 2).

DISCUSSION

The MHC confers the strongest known genetic effect in SLE to date; associations are well established for MHC class I and II loci, particularly for *HLA-DRB1*0301* and **1501* associated haplotypes. HLA loci may also influence SLE susceptibility through additional inherited or noninherited mechanisms. These hypotheses were tested using a large, well-characterized dataset of SLE and control families.

The current study is the largest study, to date, to examine maternal-offspring HLA compatibility in SLE. Several biological hypotheses have been proposed, where increased compatibility could result in a small number of non-host cells that could 1) cause dysregulation among host cells, 2) lead to presentation of non-host peptides by host cells to other host cells, 3) inactivate T lymphocytes upon interaction, or 4) undergo differentiation and become targets of a later immune response (9). Evidence for increased maternal-offspring HLA class II compatibility has been previously reported for both SLE and systemic sclerosis (SSc), suggesting that HLA class II loci may be involved through an undefined pathway dependent on maternal-offspring compatibility (10, 11). A recent study reported that maternal-offspring HLA compatibility does not influence risk for type 1 diabetes (T1D) (12).

Stevens et al. reported evidence for increased *DRB1* bidirectional compatibility for *DRB1* allele groups in 30 maternal-SLE son pairs when compared to 76 independent, healthy maternal-son pairs (OR = 5.0, 95% CI = 1.6-15.7, $P = 0.006$) (10). When analyses were restricted to sons carrying *DRB1*03* or *DRB1*15/16*, results were stronger (OR = 7.2, 95% CI = 1.6-32.8, $P = 0.01$), and remained significant when non-European Americans were excluded. In contrast, our results indicate that maternal-offspring *DRB1* compatibility does not influence SLE

susceptibility. We observed some weak evidence for decreased bidirectional compatibility in male and nulligravid female maternal-offspring pairs compared to paternal-offspring pairs, however, this result did not reach statistical significance (OR = 0.56, 95% CI = 0.29-1.04, P = 0.06). Larger studies will be required to exclude the possibility of very modest *DRB1* compatibility effects on risk for SLE.

Several factors may have contributed to the disparity between findings reported by Stevens et al. and our results. The current study was larger, and used both independent and family-based controls for all histocompatibility analyses. In addition, male and nulligravid (never pregnant at, or prior to, diagnosis) female cases were analyzed separately to account for the potential contribution of fetal microchimerism. In contrast, Stevens et al. excluded a role for fetal microchimerism by including only maternal-son pairs and used an independent control group for comparison. Whereas our study was limited to European-Americans, the previous study also included African- and Asian-American individuals. Finally, SLE cases in our study were derived from trio families, whereas Stevens et al. utilized cases from families with multiple affected individuals. It is possible that one or more of these factors, or an undetermined difference in clinical phenotype represented by both groups, may help explain the observed differences. For example, disease differences attributed to familial SLE or subgroups defined by gender, race/ethnicity or other clinical features such as presence of particular autoantibodies and/or lupus nephritis may be relevant to studies of histocompatibility and SLE. Finally, the analysis of male and nulligravid female cases conceived to nulligravid mothers separately would account for the potential contribution of sibling microchimerism, and could be the focus of a future study.

Parent-of-origin effects, potentially operating through imprinting, have been reported for multiple sclerosis (MS) with respect to the inheritance of HLA class II alleles (13). A similar HLA study in T1D reported negative findings (12). Likewise, results from our study do not support a role for *DRB1*-associated parent-of-origin effects in SLE, even for known risk alleles *DRB1*1501*, **0301* and **0801*. Additional classical HLA loci were not the focus of the current study and should be included in future studies. Although strong linkage disequilibrium is present between *DRB1*, *DQB1* and *DQA1* loci, association between particular alleles on haplotypes is not complete, and therefore, more may be learned by including additional class II HLA loci, as well as all class I HLA loci, in larger SLE studies.

There is evidence that HLA alleles may act as environmental risk factors. Exposure to HLA NIMA may therefore shape the immune repertoire of the offspring and either predispose to or protect against future immune reactions. In addition to maternal-offspring cell trafficking and oral exposure through breast milk, NIMA effects may occur through maternal microchimerism. Both risk and protective NIMA effects have been reported for RA (8, 14). NIMA effects do not appear to play a strong role in T1D, though some evidence has been reported (12, 15). We tested the hypothesis that maternal histocompatibility antigens, specifically those for *HLA-DRB1*, may contribute to risk for SLE. Our study did not reveal any evidence for NIMA effects in SLE at the *DRB1* locus, even for established risk alleles. The current study had 80% power to detect a modest association (OR $\geq$ 1.5) for parent-of-origin or NIMA effects conferred by SLE risk alleles *DRB1*0301* or **1501*. A role for other class I or II NIMA in risk for SLE cannot be excluded.

CONCLUSION

In conclusion, this large study of SLE families and healthy maternal-offspring pairs does not support a major role for *DRB1* in disease susceptibility mediated through maternal-offspring compatibility, parent-of-origin or NIMA effects. Future studies should examine additional classical HLA loci.

ACKNOWLEDGMENTS

We thank Benjamin Goldstein, Farren Briggs, Ira Tager and Gary Artim. Support was provided by the Kirkland Scholar Award (LAC, JBH) from the Mary Kirkland Center for Lupus Research, the Graduate Student Achievement Award (PGB) from the American College of Rheumatology, the Alliance for Lupus Research (JBH, KLM), the U.S. Department of Veterans Affairs (JBH), Lupus Foundation Minnesota (KLM) and NIH grants R01-AR22804 (LAC), R01-AR052300 (LAC), K24-AR02175 (LAC), P01 AR049084 (JBH), R01 AR42460 (JBH), R01 AR043274-14 (KLM) and N01 AR62277 (JBH) from NIAMS, R37 AI24717 (JBH), P01 AI083194 (JBH), 5R01 A1063274-06 (PMG) and F31-AI075609 (PGB) from NIAID, and P20-RR020143 (JBH) from NCRR. This study was performed in part in the General Clinical Research Center, Moffitt Hospital, UCSF, with funds provided by the National Center for Research Resources, 5-M01-RR-00079, U.S. Public Health Service. The content is solely the responsibility of the authors and does not necessarily represent the official views of NIH/NIAID.

REFERENCES

1. Kaplan J, Land S. Influence of maternal-fetal histocompatibility and MHC zygosity on maternal microchimerism. J Immunol 2005;174(11):7123-8.
2. Hall JM, Lingenfelter P, Adams SL, Lasser D, Hansen JA, Bean MA. Detection of maternal cells in human umbilical cord blood using fluorescence in situ hybridization. Blood 1995;86(7):2829-32.
3. Claas FH, Gijbels Y, van der Velden-de Munck J, van Rood JJ. Induction of B cell unresponsiveness to noninherited maternal HLA antigens during fetal life. Science 1988;241(4874):1815-7.
4. Mold JE, Michaelsson J, Burt TD, Muench MO, Beckerman KP, Busch MP, et al. Maternal alloantigens promote the development of tolerogenic fetal regulatory T cells in utero. Science 2008;322(5907):1562-5.
5. Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum 1997;40(9):1725.
6. Barcellos LF, May SL, Ramsay PP, Quach HL, Lane JA, Nititham J, et al. High-Density SNP Screening of the Major Histocompatibility Complex in Systemic Lupus Erythematosus Demonstrates Strong Evidence for Independent Susceptibility Regions. PLoS Genet 2009;5(10):e1000696.
7. Dankers MKA, Roelen DL, Korfage N, de Lange P, Witvliet M, Sandkuijl L, et al. Differential immunogenicity of paternal HLA Class I antigens in pregnant women. Human Immunology 2003;64(6):600-606.
8. Feitsma AL, Worthington J, van der Helm-van Mil AHM, Plant D, Thomson W, Ursum J, et al. Protective effect of noninherited maternal HLA-DR antigens on rheumatoid arthritis development. Proceedings of the National Academy of Sciences 2007;104(50):19966-19970.
9. Nelson JL. Microchimerism and human autoimmune diseases. Lupus 2002;11(10):651-654.
10. Stevens AM, Tsao BP, Hahn BH, Guthrie K, Lambert NC, Porter AJ, et al. Maternal HLA class II compatibility in men with systemic lupus erythematosus. Arthritis Rheum 2005;52(9):2768-73.
11. Artlett CM, Welsh KI, Black CM, Jimenez SA. Fetal-maternal HLA compatibility confers susceptibility to systemic sclerosis. Immunogenetics 1997;47(1):17-22.
12. Bronson PG, Ramsay PP, Thomson G, Barcellos LF. Analysis of maternal–offspring HLA compatibility, parent-of-origin and non-inherited maternal effects for the classical HLA loci in type 1 diabetes. Diabetes, Obesity and Metabolism 2009;11(s1):74-83.
13. Ramagopalan SV, Herrera BM, Bell JT, Dyment DA, DeLuca GC, Lincoln MR, et al. Parental transmission of HLA-DRB1*15 in multiple sclerosis. Human Genetics 2008;122(6):661-663.
14. Guthrie KA, Tishkevich NR, Nelson JL. Non-inherited maternal human leukocyte antigen alleles in susceptibility to familial rheumatoid arthritis. Ann Rheum Dis 2009;68(1):107-9.
15. Åkesson K, Carlsson A, Ivarsson SA, Johansson C, Weidby BM, Ludvigsson J, et al. The non-inherited maternal HLA haplotype affects the risk for type 1 diabetes. International Journal of Immunogenetics 2009;36(1):1-8.

Table 1. OR, 95% CI and *P*-values from Fisher's exact tests of maternal-offspring *DRB1* compatibility[1] in SLE families compared to both paternal-offspring compatibility (father controls) and to independent healthy maternal-offspring pairs (healthy controls) in the overall sample[2] as well as restricted to male and nulligravid female SLE offspring[3].

	SLE	**Father controls**			**Healthy controls**		
***HLA-DRB1* Compatibility**	**N (%)**	**N (%)**	**OR (95% CI)**	***P***	**N (%)**	**OR (95% CI)**	***P***
Overall[2]							
Unidirectional child-to-parent	96 (13.6)	93 (13.2)	1.01 (0.73-1.41)	0.94			
	96 (13.6)				17 (9.0)	1.32 (0.91-2.98)	0.10
Unidirectional parent-to-child	74 (10.5)	72 (10.2)	1.01 (0.70-1.46)	0.99			
	74 (10.5)				22 (11.7)	0.85 (0.56-1.68)	0.89
Bidirectional	87 (12.3)	99 (14.0)	0.87 (0.62-1.20)	0.38			
	87 (12.3)				21 (11.2)	0.65 (0.69-2.08)	0.61
Not increased	450 (63.6)	443 (62.7)	Reference				
	450 (63.6)				128 (68.1)	Reference	
Male and nulligravid female SLE offspring[3]							
Unidirectional child-to-parent	27 (13.9)	25 (12.9)	0.96 (0.50-1.82)	0.88			
	27 (13.9)				17 (9.0)	1.56 (0.78-3.21)	0.19
Unidirectional parent-to-child	15 (7.7)	19 (9.8)	0.70 (0.31-1.53)	0.36			
	15 (7.7)				22 (11.7)	0.67 (0.31-1.43)	0.29
Bidirectional	22 (11.3)	35 (18.0)	0.56 (0.29-1.04)	0.06			
	22 (11.3)				21 (11.2)	1.03 (0.51-2.08)	0.99
Not increased	130 (67.0)	115 (59.3)	Reference				
	130 (67.0)				128 (68.1)	Reference	

[1]Histocompatibility was estimated using two-digit *DRB1* typing resolution.

[2]The overall sample compared *HLA-DRB1* compatibility of 707 maternal-affected offspring pairs to 707 paternal-affected offspring pairs as well as 188 independent healthy controls.

[3]The male and nulligravid female sample compared *HLA-DRB1* compatibility of 194 maternal-affected offspring pairs where the affected offspring was either male or nulligravid (never pregnant at, or prior to, diagnosis) female to 194 paternal-affected offspring pairs as well as 188 independent healthy controls.

Figure 1. Maternal-offspring HLA compatibility relationships.

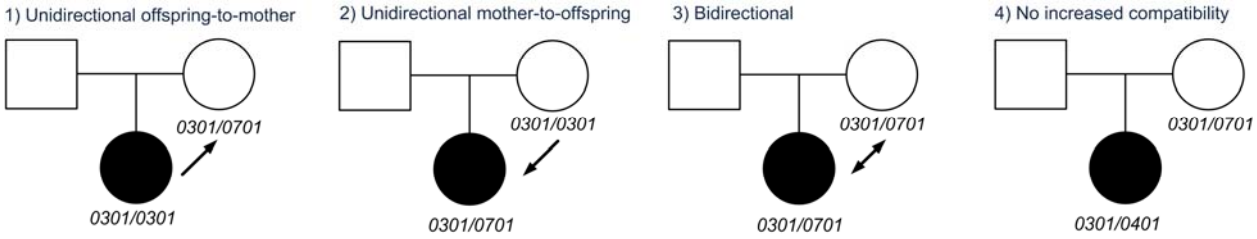
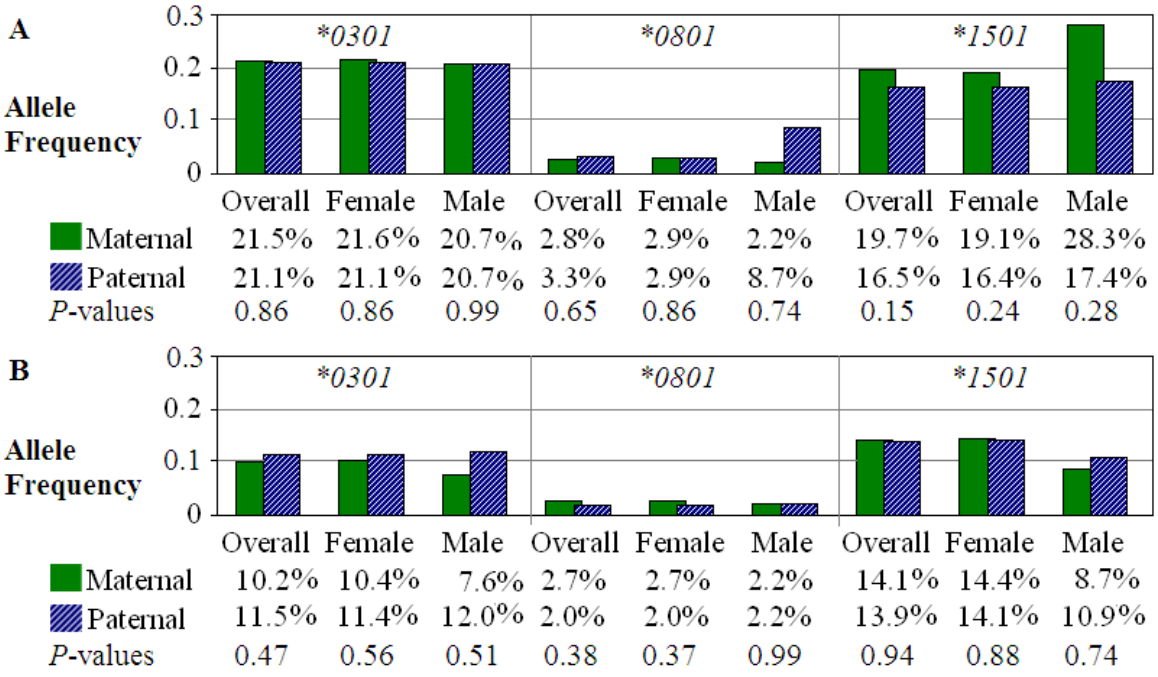


Figure 2. *HLA-DRB1* susceptibility allele frequencies in SLE families from tests of (A) parent-of-origin effects (maternal vs. paternal transmitted alleles); (B) noninherited maternal antigen (NIMA) effects (maternal vs. paternal nontransmitted alleles).



CHAPTER EIGHT

Analysis of Maternal-Offspring HLA Compatibility, Parent-of-Origin and Noninherited Maternal Effects for the Classical HLA Loci in Type 1 Diabetes

Published: Bronson PG, Ramsay PP, Thomson G, Barcellos LF. (2009) Analysis of HLA compatibility, parent-of-origin, and non-inherited effects in type 1 diabetes families. Diabetes Obes Metab 11(s1): 74-83.

ABSTRACT

Type 1 diabetes (T1D) is a complex trait for which variation in the classical human leukocyte antigen (HLA) loci within the major histocompatibility complex (MHC) significantly influences disease risk. To date, HLA class II DR-DQ genes confer the strongest known genetic effect in T1D. HLA loci may also influence T1D through additional inherited or noninherited effects. Evidence for the role of increased maternal-offspring HLA compatibility, and both parent-of-origin (POO) and noninherited maternal HLA (NIMA) effects in autoimmune disease has been previously established. The current study tested hypotheses that classical HLA loci influence T1D through POO and NIMA effects, in addition to genetic transmission of particular risk alleles. The T1D Genetics Consortium cohort was of European descent and consisted of 2,271 affected sib pair families (total N =11,023 individuals). Class I genes *HLA-A, C* and *B,* and class II genes *HLA-DRB1, DQA1, DQB1, DPA1* and *DPB1* were studied. The pedigree disequilibrium test was used to examine transmission of HLA alleles to individuals with T1D. Conditional logistic regression was used to model compatibility relationships between mother-offspring and father-offspring for all HLA loci. POO and NIMA effects were investigated by comparing frequencies of maternal and paternal transmitted and non-transmitted HLA alleles for each locus. Analyses were also stratified by gender of T1D-affected offspring. Strong associations were observed for all classical HLA loci except for *DPA1,* as expected. Compatibility differences between mother-offspring and father-offspring were not observed for any HLA loci. Further, POO and NIMA HLA effects influencing T1D were not present. In conclusion, maternal-offspring HLA compatibility, POO and NIMA effects for eight classical HLA loci were investigated. Results suggest that these HLA related effects are unlikely to play a major role in the development of T1D.

BACKGROUND

Type 1 diabetes (T1D) is an autoimmune disease characterized by chronic T-cell mediated destruction of pancreatic insulin-producing β-cells (1). While age-at-onset peaks in late childhood, adults also develop this disorder, and incidence rates for females and males are similar (2). Incidence in the United States is estimated to be ~15 in 100,000 children per year; however, it varies widely around the world and has been increasing over the past decade (2). Although the etiology of T1D remains unknown, evidence for genetic susceptibility is well established (3, 4). Concordance for T1D in monozygotic twins is 70% compared to just 13% in dizygotic twins; the relative risk for sibs (λ_s) is approximately 15 in Americans of European descent (5).

The human leukocyte antigen (HLA) class II genes *HLA-DRB1, DQA1* and *DQB1* in the major histocompatibility complex (MHC) region (6p21) are directly involved; the HLA region accounts for 40-50% of the genetic susceptibility in individuals of Northern European descent (6). The majority of individuals of European descent with T1D carry the HLA-DR3 (*DRB1*0301-DQA1*0501-DQB1*0201*) or DR4 (*DRB1*04-DQA1*0301-DQB1*0302*) class II haplotype, and approximately 30-50% are *DR3/DR4* heterozygotes (7). *DR3/DR4* heterozygosity confers the highest diabetes risk (8). Different class II HLA associations with T1D are present in non-European populations (9). Class I *HLA-B* has also been associated with T1D risk, specifically the *B*39* and *B*18* alleles (10, 11). Interestingly, the class II HLA-DR2 *(DRB1*1501-DQB1*0602*) haplotype is protective in all populations studied to date (12). Additional non-MHC genetic risk factors for T1D include *PTPN22* (1p13), *CTLA4* (2q33) and *IDDM2* (11p15) (13-15). Environmental factors have also been strongly implicated in both pathogenesis and outcome of T1D (16).

HLA loci may also influence T1D through additional inherited or noninherited effects. Differences in maternal and paternal transmission rates, or parent-of-origin (POO) effects, have been observed in T1D. One potential mechanism is 'genomic imprinting', an epigenetic modification of the genome that results in unequal transcription of parental alleles and subsequent allele expression, depending on whether alleles were transmitted maternally or paternally.

HLA compatibility between a mother and her offspring may also contribute to susceptibility to autoimmunity, possibly because HLA similarity between the mother and fetus may promote the persistence of fetal cells in the host or perhaps through specific exposure to noninherited maternal HLA (NIMA) risk or protective alleles. Risk for an autoimmune disease would be potentially increased in either mother or offspring. Maternal-offspring HLA compatibility that increases disease risk in the mother could explain: 1) increased prevalence of autoimmune diseases in women following their childbearing years, and 2) clinical similarities between scleroderma (systemic sclerosis [SSc] and graft-versus-host disease (17). With regard to T1D, maternal-offspring HLA compatibility could affect risk in the offspring. Maternal-offspring cell trafficking is common and bidirectional; maternal nucleated cell and plasma DNA transfers into fetal circulation in 24% and 30% of offspring, respectively (18). Maternal-offspring effects can present as excess HLA compatibility between the mother and affected offspring or excess maternal homozygosity. Possible maternal-offspring HLA compatibility relationships are illustrated in Figure 1.

Finally, non-host exposure during fetal development and potential long-term persistence of maternal cells in offspring may play a role in T1D pathogenesis (19-22). The developing immune system of the fetus is exposed to NIMA in utero (21, 23, 24). Decreased B cell responses and cytotoxic T cell activity to HLA class I NIMA have been reported (25-29). NIMA can have a lifelong influence on the immune system and may tolerize or predispose to autoimmune reactions (30-32).

The current study tested hypotheses that classical HLA loci influence T1D through these additional biological mechanisms in addition to genetic transmission of particular risk alleles.

METHODS

Subjects

The T1D Genetics Consortium cohort (release 2007.02.MHC) was of European descent and consisted of 2,271 affected sibling pair (ASP) families with 11,023 individuals. This research resource has been previously described (33). Briefly, the families were derived from multiple cohorts: Asia Pacific (177), BDA-Warren (422), Danish (147), European (475), HBDI (421), Joslin (117), North American (321), Sardinian (77) and UK (114). The mean number of individuals per family was 5 and ranged from 3 to 26. The mean number of generations per family was 2 and ranged from 1 to 4.

A subset of 1,780 ASP families had at least two-digit classical HLA genotypes available. Analyses were conducted on trio families consisting of two parents and one affected offspring; one affected offspring per family was chosen randomly from all affected offspring with *HLA-DRB1* genotypes. We created two additional trio family samples for gender-stratified analyses. For the male sample we randomly selected one male affected offspring per family. There were 1,376 families with at least one male affected offspring. For the female sample we randomly selected one female affected offspring per family. There were 1,291 families with at least one female affected offspring. Parental genotypes were available to determine transmission for all affected offspring. A total of 6,227 individuals were analyzed (Table 1).

For the HLA compatibility analyses, we limited the samples described in the previous paragraph to 1,213 ASP families (5,804 individuals) with complete four-digit classical HLA genotypes available. There were 954 families and 876 families in the male and female samples, respectively. A total of 4,256 individuals were utilized for the analyses of compatibility patterns (Table 2).

The families were collected for both linkage and association studies; the sample size was designed to be sufficiently large to detect associations as well as secondary gene effects in a region such as HLA. The data were generated as part of a high-density screen of the MHC designed for association and haplotype analysis and to detect genes in the HLA region additional to the well-documented HLA DR-DQ effect. The nuclear family study design is advantageous because it precludes potential confounding from ethnic mismatching between patients and randomly ascertained controls due to population stratification, migration, or admixture. This design also reduces the potential of misclassification error from genotyping because we can check the data for pedigree inconsistencies.

Classical HLA genotyping

The T1DGC protocol has been previously described (34).

Statistical analyses

We used PedCheck 1.1 to identify pedigree inconsistencies in our overall sample of 1,780 trio families (35). For any pedigrees with an inconsistency we zeroed out genotypes for the entire family at that specific locus only, assuming a genotyping error. Mega2 3.0 R12 was used to manipulate data (36). The pedigree disequilibrium test (PDT 6.0 build 5) was used to examine frequencies of transmitted vs. non-transmitted alleles for each HLA locus (37). The PDT is a powerful analytical method that uses genetic data from related nuclear families and discordant sibships within extended pedigrees.

We defined maternal-offspring compatibility categorically: (1) unidirectional child-to-mother compatible, (2) unidirectional mother-to-offspring compatible, (3) bidirectional and (4) no increased compatibility. We modeled maternal-offspring compatibility in R (2.6) using conditional logistic regression and for controls the compatibility of the affected child to the father was used (38). The analysis was restricted to trios from the previously described overall, male and female samples that had complete four-digit genotyping information available.

We pair-matched on family in a matched case-control analysis using conditional logistic regression; parent's gender was the outcome (39).

$$\text{logit } P(Y) = \beta_0 + \beta_1 \text{unidirectional p.} + \beta_2 \text{unidirectional c.} + \beta_3 \text{bidirectional} + \sum_{i=1}^{n} \gamma_i V_i$$

In the above formula, *n* refers to the number of families in the matched analysis, unidirectional p. is unidirectional parent-to-child compatibility and unidirectional c. is unidirectional child-to-parent compatibility. These analyses were repeated with HLA compatibility as a binary exposure (any compatibility vs. no increased compatibility).

$$\text{logit } P(Y) = \beta_0 + \beta_1 \text{any compatibility} + \sum_{i=1}^{n} \gamma_i V_i$$

For the parent-of-origin (POO) analyses, frequencies of maternal vs. paternal transmitted alleles were compared using a chi-square contingency table test for heterogeneity in AFBAC (1.13) (40). The same test was used in the NIMA analyses to compare frequencies of maternal vs. paternal nontransmitted alleles. Both global and allele-specific analyses (when appropriate) were performed. Based on a Bonferroni correction for the total number of tests performed in this study ($N = 24$), a criterion of $P<0.002$ was set for statistical significance.

RESULTS

There were no pedigree inconsistencies for *HLA-DQA1, DPA1* or *DPB1*. There were one, six and five families with pedigree inconsistencies at the *HLA-DRB1, DQB1* and *C* loci, respectively. None of the families had more than one pedigree inconsistency, indicating genotyping error rather than non-paternity. As expected, the majority of classical HLA loci were strongly associated with T1D in both the overall and gender-stratified analyses. The only exception was the class II gene *HLA-DPA1,* which did not show evidence for association. Table

3 displays global *P*-values. *HLA-DR* and *DQ* are the primary disease genes and we have not investigated whether the associations at the other loci are due to linkage disequilibrium with *DR-DQ*. Refer to Supplementary Table 1 for observed and expected transmitted and nontransmitted allele frequencies with the test statistic and odds ratio (OR).

Unidirectional offspring-to-mother compatibility at the class I genes *HLA-A, C* and *B* and the class II genes *HLA-DRB1, DQA1, DQB1, DPA1* and *DPB1* was not associated with T1D. The corresponding OR, 95% CIs and *P*-values are listed in Table 4. Unidirectional mother-to-offspring compatibility, bidirectional compatibility and any compatibility did not demonstrate association with T1D (data not shown). Refer to Supplementary Table 2 for frequencies of maternal-offspring and paternal-offspring HLA compatibility relationships.

In addition, parent-of-origin (POO) and noninherited maternal HLA (NIMA) effects were also examined for each classical HLA locus; evidence for involvement in T1D was not present (Table 5), even when specific T1D risk alleles were examined separately, including *DRB1*0401, DRB1*0301* and *DQB1*0302*. Supplementary Table 3 shows observed and expected allele-specific transmitted maternal and paternal allele frequencies with ORs. Supplementary Table 4 shows observed and expected allele-specific nontransmitted maternal and paternal allele frequencies with ORs. Comparison of paternal and maternal transmitted and paternal and maternal non-transmitted T1D risk HLA alleles in the overall dataset, and for male and female cases analyzed separately, revealed nearly identical frequencies (Figure 2).

DISCUSSION

To date, the MHC confers the strongest known genetic effect in T1D; associations are well established for class I and class II loci, particularly for the class II *HLA-DRB1*0301* and **04* associated haplotypes. HLA loci may also influence T1D through additional inherited or noninherited effects. Evidence for increased maternal-offspring HLA class II compatibility has been reported for systemic lupus erythematosus (SLE) and systemic sclerosis (SSc). Compared to controls, male SLE patients were more likely to have HLA class II genotypes identical to their mothers (41). In addition, compared to controls, SSc patients exhibited increased HLA class II compatibility with their offspring, or with their offspring or mother (42, 43). Taken together, these results suggest that HLA class II loci may be involved in etiology of both SLE and SSc through undefined phenomena dependent on maternal-offspring compatibility. Several biological hypotheses have been proposed, where increased compatibility could result in a small number of non-host cells that could ultimately 1) cause dysregulation among host cells, 2) lead to presentation of non-host peptides by host cells to other host cells, 3) inactivate T lymphocytes upon interaction, or 4) undergo differentiation and become targets of a later immune response (20, 44-46).

The current study is the first to examine maternal-offspring HLA compatibility in T1D. Our results indicate that maternal-offspring compatibility at the MHC class I genes *HLA-A, C* and *B* and class II genes *HLA-DRB1, DQA1, DQB1, DPA1* and *DPB1* does not influence T1D. Our study used paternal-offspring HLA compatibility as the controls in an analysis matched on family. Future studies would ideally test for differences in patterns of compatibility using an independent controls sample of mother-offspring pairs without T1D.

Parent-of-origin (POO) effects, potentially operating through the phenomenon of imprinting, have been observed previously in T1D and multiple sclerosis (MS) with respect to the inheritance of HLA class II alleles but results have been inconsistent. Excess paternal inheritance of the *DR3* risk haplotype has been reported in female Sardinian MS patients (47). More recently, excess maternal inheritance of *HLA-DRB1*15* was observed in a larger study of 1,515 MS families ($P = 0.005$) (48). An early study of 107 T1D families reported increased paternal transmission of DR4 to affected and unaffected offspring (72.1%) compared to maternal transmission (55.6%) (49). A study of 28 Japanese T1D families reported that the *DQA1*0301-DQB1*0302* haplotype exhibited preferential maternal transmission and strong transmission disequilibrium with T1D positive for antibodies to glutamic acid decarboxylase (50). Bain et al. (1994), Undlien et al. (1995) and Martin-Villa et al. (1990) have reported a lack of evidence for POO effects in HLA class II alleles; these studies examined 282, 61 and 108 T1D families, respectively. Many of these studies suffered from relatively small sample sizes and did not account for multiple testing when reporting statistical significance. In contrast, this study is the largest (Table 2), to date, to examine POO effects in the classical HLA loci and account for multiple statistical tests. Results do not support a role for HLA-associated POO effects in T1D, even for T1D risk alleles *DRB1*0401*, *DQB1*0302* and *DRB1*0301*, and indeed are in agreement with others (51-53).

Interestingly, POO effects have also been examined in T1D for non-MHC genetic risk factors. There were no POO effects observed for *PTPN22* in a study of 341 T1D families (54). A study of the *CTLA4* exon 1 polymorphism (49 A/G) in 70 T1D families showed increased maternal allele transmission of the *G* allele to T1D-affected offspring: 71% vs. the random 50% observed in unaffected offspring ($P<0.03$). This distortion was stronger in T1D offspring with maternal inheritance of *HLA-DRB1*03* (80%, $P<0.01$) or variable number tandem repeats at the *IDDM2* locus (80%, $P<0.02$) (55). A paternal origin effect has been observed for *IDDM8* (6q27) and a maternal origin effect has been observed in 404 parent-offspring T1D trios for the *IGF2R* locus (6q26) (3, 56). Further research is needed to confirm these findings.

There is evidence that HLA alleles may also act as environmental risk factors. This current study tested the hypothesis that cells and antigens of the mother may modulate the antigen-specific reactivity of the fetal immune system. Exposure to noninherited maternal HLA (NIMA) via several different mechanisms may therefore shape the immune repertoire of the offspring and either predispose to or protect against future immune reactions. A tolerogenic effect may explain the longer survival of renal transplants from sibling donors expressing NIMA vs. noninherited paternal antigens (NIPA) (57). In the pre-cyclosporine era, breastfeeding exposure was associated with improved graft survival in recipients of maternal kidney transplants (58, 59). The role of breast milk in this observation was confirmed by a highly immunogenic heart allograft mouse model in which both in utero exposure and milk feeding were required for the NIMA effect (60).

In addition to maternal-offspring cell trafficking and oral exposure through breast milk, another potential mechanism for NIMA is maternal microchimerism, when a small population of cells or DNA in an individual is derived from their mother. Maternal cells have been detected in offspring several decades following birth (61). Compared to healthy women, female SSc patients have increased frequencies of maternal cells in their peripheral blood cells (62).

A NIMA effect on risk for rheumatoid arthritis (RA) has been explored in several studies. An early study reported association between NIMA and RA for *HLA-DR4* alleles (63). Negative findings were later reported by a study of familial RA: frequencies of *HLA-DRB1*04*, **0401/*0404*, and shared epitope (SE) positive NIMA compared to NIPA were not increased in RA patients lacking these susceptibility alleles (64). A later study reported an excess of *DRB1*04* and SE NIMA ($P = 0.05$) compared with NIPA; a combined analysis with previous studies showed that mothers were more likely to carry a noninherited *DRB1*04* and SE alleles (65). Recently, in the largest study of NIMA to date, the first evidence for a protective NIMA effect was reported: a mother carrying the protective amino acid sequence DERAA (*HLA-DRB1*0103, *0402, *1103, *1301* and **1304*) at the SE may transfer protection against RA to her DERAA-negative offspring (66).

To date, only one study of NIMA effects in T1D has been reported. T1D patients who did not carry any high-risk HLA alleles presented HLA DR3-DQ2 and DR4-DQ8 risk haplotypes more frequently as NIMA compared to NIPA (67). The results from the current study, however, do not support a NIMA effect in T1D. A global test for NIMA effects, and a specific examination of the known T1D risk alleles *DRB1*0401*, *DQB1*0302* and *DRB1*0301* revealed negative results.

CONCLUSION

In conclusion, the largest study of T1D families, to date, does not support a major role for the classical HLA loci in disease susceptibility mediated through maternal-offspring compatibility, POO or NIMA effects.

ACKNOWLEDGEMENTS

The project described was supported by Grant Number F31AI075609 from the National Institute of Allergy and Infectious Diseases. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Allergy and Infectious Diseases or the National Institutes of Health. This research utilizes resources provided by the Type 1 Diabetes Genetics Consortium, a collaborative clinical study sponsored by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institute of Allergy and Infectious Diseases (NIAID), National Human Genome Research Institute (NHGRI), National Institute of Child Health and Human Development (NICHD), and Juvenile Diabetes Research Foundation International (JDRF) and supported by U01 DK062418.

REFERENCES

1. Eisenbarth GS. Type I diabetes mellitus. A chronic autoimmune disease. N Engl J Med 1986;314(21):1360-8.
2. Incidence and trends of childhood Type 1 diabetes worldwide 1990-1999. Diabet Med 2006;23(8):857-866.
3. McCann JA, Xu YQ, Frechette R, Guazzarotti L, Polychronakos C. The Insulin-Like Growth Factor-II Receptor Gene Is Associated with Type 1 Diabetes: Evidence of a Maternal Effect. J Clin Endocrinol Metab 2004;89(11):5700-5706.
4. Lorenzen T, Pociot F, Hougaard P, Nerup J. Long-term risk of IDDM in first-degree relatives of patients with IDDM. Diabetologia 1994;37(3):321-7.
5. Jahromi MM, Eisenbarth GS. Genetic Determinants of Type 1 Diabetes Across Populations. Ann NY Acad Sci 2006;1079(1):289-299.
6. Noble JA, Valdes AM, Cook M, Klitz W, Thomson G, Erlich HA. The role of HLA class II genes in insulin-dependent diabetes mellitus: molecular analysis of 180 Caucasian, multiplex families. Am J Hum Genet 1996;59(5):1134-48.
7. Rotter JI, Anderson CE, Rubin R, Congleton JE, Terasaki PI, Rimoin DL. HLA Genotypic Study of Insulin-Dependent Diabetes - the Excess of Dr3/Dr4 Heterozygotes Allows Rejection of the Recessive Hypothesis. Diabetes 1983;32(2):169-174.
8. Thomson G, Robinson WP, Kuhner MK, Joe S, Macdonald MJ, Gottschall JL, et al. Genetic-Heterogeneity, Modes of Inheritance, and Risk Estimates for a Joint Study of Caucasians with Insulin-Dependent Diabetes-Mellitus. Am J Hum Genet 1988;43(6):799-816.
9. Karvonen M, Tuomilehto J, Libman I, Laporte R. A Review of the Recent Epidemiologic Data on the Worldwide Incidence of Type-1 (Insulin-Dependent) Diabetes-Mellitus. Diabetologia 1993;36(10):883-892.
10. Valdes AM, Erlich HA, Noble JA. Human leukocyte antigen class I B and C loci contribute to Type 1 Diabetes (T1D) susceptibility and age at T1D onset. Hum Immunol 2005;66(3):301-313.
11. Nejentsev S, Howson JMM, Walker NM, Szeszko J, Field SF, Stevens HE, et al. Localization of type 1 diabetes susceptibility to the MHC class I genes HLA-B and HLA-A. Nature 2007;450(7171):887-892.
12. Pugliese A. Genetic protection from insulin-dependent diabetes mellitus. Diabetes Nutr Metab 1997;10(4):169-179.
13. Bottini N, Musumeci L, Alonso A, Rahmouni S, Nika K, Rostamkhani M, et al. A functional variant of lymphoid tyrosine phosphatase is associated with type I diabetes. Nat Genet 2004;36(4):337-8.
14. Nistico L, Buzzetti R, Pritchard LE, Van der Auwera B, Giovannini C, Bosi E, et al. The CTLA-4 gene region of chromosome 2q33 is linked to, and associated with, type 1 diabetes. Belgian Diabetes Registry. Hum Mol Genet 1996;5(7):1075-80.
15. Bennett ST, Lucassen AM, Gough SC, Powell EE, Undlien DE, Pritchard LE, et al. Susceptibility to human type 1 diabetes at IDDM2 is determined by tandem repeat variation at the insulin gene minisatellite locus. Nat Genet 1995;9(3):284-92.
16. Adeghate E, Schattner P, Dunn E. An update on the etiology and epidemiology of diabetes mellitus. Ann NY Acad Sci 2006;1084:1-29.
17. Bianchi DW. Fetomaternal cell trafficking: A new cause of disease? American Journal of Medical Genetics 2000;91(1):22-28.

18. Lo YMD, Lau TK, Chan LYS, Leung TN, Chang AMZ. Quantitative Analysis of the Bidirectional Fetomaternal Transfer of Nucleated Cells and Plasma DNA. Clin Chem 2000;46(9):1301-1309.
19. Lambert NC, Stevens AM, Tylee TS, Erickson TD, Furst DE, Nelson JL. From the simple detection of microchimerism in patients with autoimmune diseases to its implication in pathogenesis. Ann NY Acad Sci 2001;945:164-71.
20. Nelson JL. Microchimerism and human autoimmune diseases. Lupus 2002;11(10):651-654.
21. Nelson JJL. Microchimerism and HLA relationships of pregnancy: implications for autoimmune diseases. Curr Rheumatol Rep 2001;3(3):222-229.
22. Holzgreve W, Hahn S, Zhong XY, Lapaire O, Hosli I, Tercanli S, et al. Genetic communication between fetus and mother: short- and long-term consequences. Am J Obstet Gynecol 2007;196(4):372-381.
23. Bianchi DDW, Zickwolf GGK, Weil GGJ, Sylvester SS, DeMaria MMA. Male fetal progenitor cells persist in maternal blood for as long as 27 years postpartum. Proc Natl Acad Sci USA 1996;93(2):705-708.
24. Hall JM, Lingenfelter P, Adams SL, Lasser D, Hansen JA, Bean MA. Detection of maternal cells in human umbilical cord blood using fluorescence in situ hybridization. Blood 1995;86(7):2829-32.
25. Claas FH, van Leeuwen A, van Rood JJ. Hyperimmunized patients do not need to wait for an HLA identical donor. Tissue Antigens 1989;34(1):23-9.
26. Claas FH, Gijbels Y, van der Velden-de Munck J, van Rood JJ. Induction of B cell unresponsiveness to noninherited maternal HLA antigens during fetal life. Science 1988;241(4874):1815-7.
27. Phelan D, Hadley G, Duffy B, Mohanam S, Mohanakumar T. Antiidiotypic antibodies to HLA class I alloantibodies in normal individuals: a mechanism of tolerance to noninherited maternal HLA antigens. Hum Immunol 1991;31(1):1-6.
28. Zhang L, van Rood JJ, Claas FH. The T-cell repertoire is not dictated by self antigens alone. Res Immunol 1991;142(5-6):441-5.
29. Roelen DL, van Bree FP, van Beelen E, van Rood JJ, Claas FH. No evidence of an influence of the noninherited maternal HLA antigens on the alloreactive T cell repertoire in healthy individuals. Transplantation 1995;59(12):1728-33.
30. van Rood JJ, Claas F. Noninherited maternal HLA antigens: a proposal to elucidate their role in the immune response. Hum Immunol 2000;61(12):1390-4.
31. van Rood JJ, Claas F. Both self and non-inherited maternal HLA antigens influence the immune response. Immunol Today 2000;21(6):269-73.
32. Stevens AM. Microchimeric cells in systemic lupus erythematosus: targets or innocent bystanders? Lupus 2006;15(11):820-6.
33. Rich SS, Concannon P, Erlich H, Julier C, Morahan G, Nerup J, et al. The Type 1 Diabetes Genetics Consortium. Ann NY Acad Sci 2006;1079(1):1-8.
34. Erlich H, Valdes AM, Noble J, Carlson JA, Varney M, Concannon P, et al. HLA DR-DQ haplotypes and genotypes and type 1 diabetes risk: analysis of the type 1 diabetes genetics consortium families. Diabetes 2008;57(4):1084-92.
35. O'Connell JR, Weeks DE. PedCheck: a program for identification of genotype incompatibilities in linkage analysis. Am J Hum Genet 1998;63(1):259-66.

36. Mukhopadhyay N, Almasy L, Schroeder M, Mulvihill WP, Weeks DE. Mega2: data-handling for facilitating genetic linkage and association analyses. Bioinformatics 2005;21(10):2556-7.
37. Martin ER, Monks SA, Warren LL, Kaplan NL. A test for linkage and association in general pedigrees: the pedigree disequilibrium test. Am J Hum Genet 2000;67(1):146-54.
38. R Development Core Team. R: A language and environment for statistical computing. In. Vienna: R Foundation for Statistical Computing; 2007.
39. Kleinbaum D, Klein M. Logistic regression: a self-learning text. 2 ed. New York: Springer; 2002.
40. Thomson G. Mapping disease genes: family-based association studies. Am J Hum Genet 1995;57(2):487-98.
41. Stevens AM, Tsao BP, Hahn BH, Guthrie K, Lambert NC, Porter AJ, et al. Maternal HLA class II compatibility in men with systemic lupus erythematosus. Arthritis Rheum 2005;52(9):2768-73.
42. Artlett CM, Welsh KI, Black CM, Jimenez SA. Fetal-maternal HLA compatibility confers susceptibility to systemic sclerosis. Immunogenetics 1997;47(1):17-22.
43. Nelson JL, Furst DE, Maloney S, Gooley T, Evans PC, Smith A, et al. Microchimerism and HLA-compatible relationships of pregnancy in scleroderma. Lancet 1998;351(9102):559-562.
44. Srivatsa B, Srivatsa S, Johnson KL, Samura O, Lee SL, Bianchi DW. Microchimerism of presumed fetal origin in thyroid specimens from women: a case-control study. Lancet 2001;358(9298):2034-2038.
45. Toda I, Kuwana M, Tsubota K, Kawakami Y. Lack of evidence for an increased microchimerism in the circulation of patients with Sjogren's syndrome. Ann Rheum Dis 2001;60(3):248-253.
46. Tanaka A, Lindor K, Gish R, Batts K, Shiratori Y, Omata M, et al. Fetal microchimerism alone does not contribute to the induction of primary biliary cirrhosis. Hepatology 1999;30(4):833-838.
47. Marrosu MG, Sardu C, Cocco E, Costa G, Murru MR, Mancosu C, et al. Bias in parental transmission of the HLA-DR3 allele in Sardinian multiple sclerosis. Neurology 2004;63(6):1084-1086.
48. Ramagopalan SV, Herrera BM, Bell JT, Dyment DA, DeLuca GC, Lincoln MR, et al. Parental transmission of HLA-DRB1*15 in multiple sclerosis. Hum Genet 2008;122(6):661-663.
49. Vadheim CM, Rotter JI, Maclaren NK, Riley WJ, Anderson CE. Preferential transmission of diabetic alleles within the HLA gene complex. N Engl J Med 1986;315(21):1314-8.
50. Sasaki T, Nemoto M, Yamasaki K, Tajima N. Preferential transmission of maternal allele with DQA1*0301-DQB1*0302 haplotype to affected offspring in families with type 1 diabetes. J Hum Genet 1999;44(5):318-322.
51. Martin-Villa JM, Vicario JL, Martinez-Laso J, Serrano-Rios M, Lledo G, Damiano A, et al. Lack of preferential transmission of diabetic HLA alleles by healthy parents to offspring in Spanish diabetic families. J Clin Endocrinol Metab 1990;70(2):346-8.
52. Bain SC, Rowe BR, Barnett AH, Todd JA. Parental origin of diabetes-associated HLA types in sibling pairs with type I diabetes. Diabetes 1994;43(12):1462-8.

53. Undlien DE, Akselsen HE, Joner G, Dahl-Jorgensen K, Aagenaes O, Sovik O, et al. No difference in the parental origin of susceptibility HLA class II haplotypes among Norwegian patients with insulin-dependent diabetes mellitus. Am J Hum Genet 1995;57(6):1511-4.
54. Ladner MB, Bottini N, Valdes AM, Noble JA. Association of the single nucleotide polymorphism C1858T of the PTPN22 gene with type 1 diabetes. Hum Immunol 2005;66(1):60-64.
55. Fajardy I, Vambergue A, Stuckens C, Weill J, Danze PM, Fontaine P. CTLA-4 49 A/G dimorphism and type 1 diabetes susceptibility: a French case-control study and segregation analysis. Evidence of a maternal effect. Eur J Immunogenet 2002;29(3):251-7.
56. Paterson AD, Petronis A. Sex of affected sibpairs and genetic linkage to type 1 diabetes. Am J Med Genet 1999;84(1):15-9.
57. Burlingham WJ, Grailer AP, Heisey DM, Claas FH, Norman D, Mohanakumar T, et al. The effect of tolerance to noninherited maternal HLA antigens on the survival of renal transplants from sibling donors. N Engl J Med 1998;339(23):1657-64.
58. Campbell DA, Jr., Lorber MI, Sweeton JC, Turcotte JG, Niederhuber JE, Beer AE. Breast feeding and maternal-donor renal allografts. Possibly the original donor-specific transfusion. Transplantation 1984;37(4):340-4.
59. Kois WE, Campbell DA, Lorber MI, Sweeton JC, Dafoe DC. Influence of Breast-Feeding on Subsequent Reactivity to a Related Renal-Allograft. J Surg Res 1984;37(2):89-93.
60. Andrassy J, Kusaka S, Jankowska-Gan E, Torrealba JR, Haynes LD, Marthaler BR, et al. Tolerance to noninherited maternal MHC antigens in mice. J Immunol 2003;171(10):5554-61.
61. Adams KM, Nelson JL. Microchimerism: an investigative frontier in autoimmunity and transplantation. JAMA-J Am Med Assoc 2004;291(9):1127-31.
62. Lambert NC, Erickson TD, Yan Z, Pang JM, Guthrie KA, Furst DE, et al. Quantification of maternal microchimerism by HLA-specific real-time polymerase chain reaction: studies of healthy women and women with scleroderma. Arthritis Rheum 2004;50(3):906-14.
63. van der Horst-Bruinsma IE, Hazes JM, Schreuder GM, Radstake TR, Barrera P, van de Putte LB, et al. Influence of non-inherited maternal HLA-DR antigens on susceptibility to rheumatoid arthritis. Ann Rheum Dis 1998;57(11):672-5.
64. Barrera P, Balsa A, Alves H, Westhovens R, Maenaut K, Cornelis F, et al. Noninherited maternal antigens do not increase the susceptibility for familial rheumatoid arthritis. European Consortium on Rheumatoid Arthritis Families (ECRAF). J Rheumatol 2001;28(5):968-74.
65. Harney S, Newton J, Milicic A, Brown MA, Wordsworth BP. Non-inherited maternal HLA alleles are associated with rheumatoid arthritis. Rheumatology 2003;42(1):171-4.
66. Feitsma AL, Worthington J, van der Helm-van Mil AHM, Plant D, Thomson W, Ursum J, et al. Protective effect of noninherited maternal HLA-DR antigens on rheumatoid arthritis development. Proc Natl Acad Sci USA 2007;104(50):19966-19970.
67. Pani MA, Van Autreve J, Van der Auwera BJ, Gorus FK, Badenhoop K. Non-transmitted maternal HLA DQ2 or DQ8 alleles and risk of Type I diabetes in offspring: the importance of foetal or post partum exposure to diabetogenic molecules. Diabetologia 2002;45(9):1340-3.

Figure 1. Illustration of possible maternal-offspring *HLA* compatibility relationships.

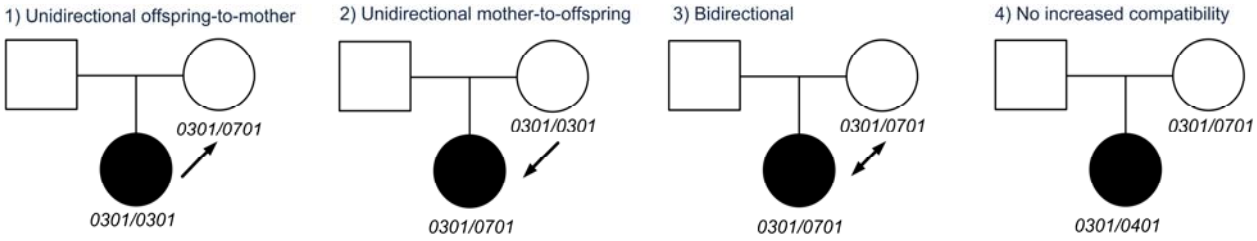
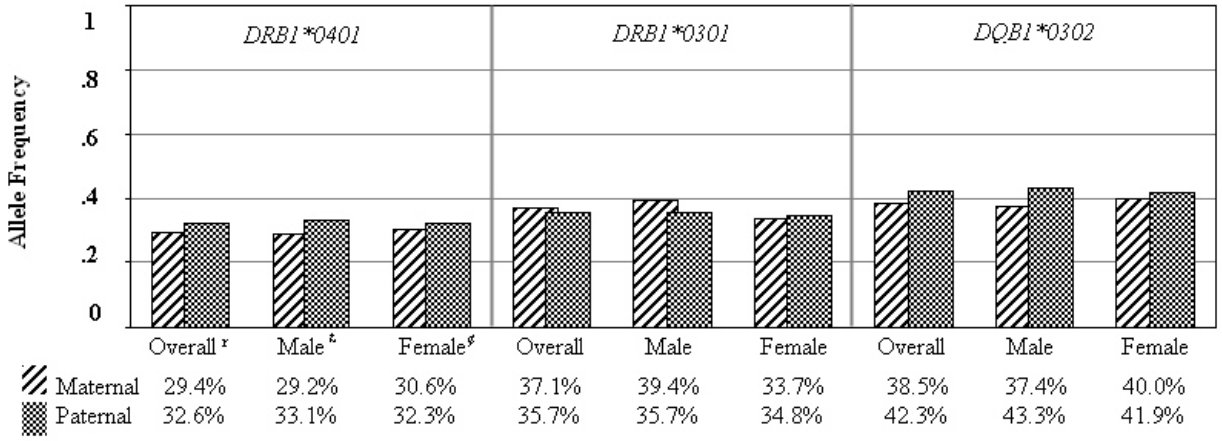
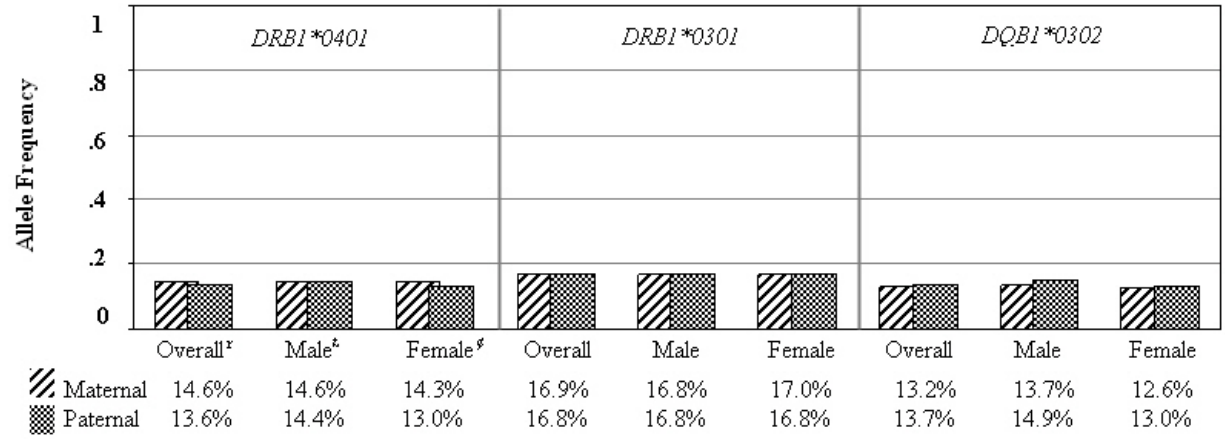


Figure 2. (A) Transmitted paternal and maternal allele frequencies for T1D susceptibility alleles from analyses of parent-of-origin (POO) effects in T1D families; (B) Nontransmitted paternal and maternal allele frequencies for T1D susceptibility alleles from analyses of noninherited maternal HLA (NIMA) effects in T1D families.

A



B



[r]The overall sample (N = 1,780) compares maternal vs. paternal nontransmitted alleles using one affected child per family.
[£]The male sample (N = 1,376) compares maternal vs. paternal nontransmitted alleles using one affected male offspring per family.
[¢]The female sample (N = 1,291) compares maternal vs. paternal nontransmitted alleles using one affected female offspring per family.

Table 1. T1D families with 2-digit genotyping available for the classical HLA loci used in analyses of parent-of-origin (POO) and noninherited maternal HLA (NIMA) effects.

Sample	Families	Families with affected offspring of one gender only	Families with affected offspring of both genders	Affected offspring in analysis	Individuals in analysis
Overall	1,780	893	887	1,780	5,340
Male	1,376	489	887	1,376	4,128
Female	1,291	404	887	1,291	3,873

Table 2. Subset of T1D families with 4-digit genotyping available for the classical HLA loci used in the HLA compatibility analyses.

Sample	Families	Families with affected offspring of one gender only	Families with affected offspring of both genders	Affected offspring in analysis	Individuals in analysis
Overall	1,213	596	617	1,213	3,639
Male	954	337	617	954	2,862
Female	876	259	617	876	2,628

Table 3. Results from pedigree disequilibrium test (PDT) analyses of transmission of HLA alleles to individuals with T1D.

	Global *P*-values		
***HLA* Locus**	**Overall**[1]	**Male**[2]	**Female**[3]
A	7.2×10^{-9}	2.3×10^{-8}	1.4×10^{-5}
C	7.2×10^{-11}	8.7×10^{-11}	4.8×10^{-11}
B	8.2×10^{-11}	8.9×10^{-11}	9.1×10^{-11}
DRB1	8.4×10^{-11}	9.2×10^{-11}	8.7×10^{-11}
DQA1	7.6×10^{-11}	8×10^{-11}	7.4×10^{-11}
DQB1	6.6×10^{-11}	8.4×10^{-11}	9.5×10^{-11}
DPA1	0.03	0.14	0.36
DPB1	4.9×10^{-11}	1.3×10^{-8}	6.6×10^{-11}

[1]The overall sample (N = 1,780) compares transmitted vs. nontransmitted HLA alleles using one affected offspring per family.
[2]The male sample (N = 1,376) compares transmitted vs. nontransmitted HLA alleles using one affected offspring per family.
[3]The female sample (N = 1,291) compares transmitted vs. nontransmitted HLA alleles using one affected offspring per family.

Table 4. Results from analyses of maternal vs. paternal unidirectional offspring-to-parent HLA compatibility in T1D families.

HLA Locus	Sample	OR (95% CI)	Global *P*-value
A	Overall[1]	0.74 (0.42-1.30)	0.31
	Male[2]	0.78 (0.41-1.46)	0.44
	Female[3]	1.20 (0.61-2.34)	0.59
C	Overall	0.87 (0.42-1.82)	0.71
	Male	0.60 (0.25-1.43)	0.25
	Female	0.76 (0.30-1.90)	0.57
B	Overall	0.41 (0.14-1.22)	0.11
	Male	0.18 (0.04-0.84)	0.03
	Female	0.38 (0.11-1.26)	0.11
DRB	Overall	0.77 (0.37-1.58)	0.48
	Male	0.64 (0.29-1.42)	0.28
	Female	0.60 (0.26-1.39)	0.22
DQA	Overall	0.99 (0.61-1.61)	0.96
	Male	0.85 (0.50-1.46)	0.56
	Female	0.83 (0.47-1.47)	0.53
DQB	Overall	0.78 (0.46-1.33)	0.35
	Male	0.70 (0.38-1.28)	0.25
	Female	0.63 (0.33-1.19)	0.15
DPA	Overall	0.97 (0.50-1.90)	0.93
	Male	0.66 (0.28-1.57)	0.36
	Female	2.30 (0.97-5.45)	0.05
DPB	Overall	1.04 (0.68-1.58)	0.86
	Male	0.89 (0.55-1.43)	0.63
	Female	0.83 (0.51-1.30)	0.46

[1]The overall sample (N = 1,213) compares maternal vs. paternal HLA compatibility using one affected offspring per family.
[2]The male sample (N = 954) compares maternal vs. paternal HLA compatibility using one affected male offspring per family.
[3]The female sample (N = 876) compares maternal vs. paternal HLA compatibility using one affected female offspring per family.

Table 5. Results from analyses of parent-of-origin (POO) HLA effects and noninherited maternal HLA (NIMA) effects (nontransmitted maternal vs. paternal alleles) in T1D families.

		Maternal vs. Paternal					
		Transmitted			Nontransmitted		
HLA Locus	**Sample**	χ^2	**df**	***P***	χ^2	**df**	***P***
A	Overall[1]	7.9	18	0.98	18.3	18	0.44
	Male[2]	12.4	17	0.77	17.8	17	0.40
	Female[3]	15.1	18	0.66	17.0	18	0.53
C	Overall	23.4	21	0.32	22.1	21	0.39
	Male	41.2	21	0.01	20.1	21	0.51
	Female	7.1	20	1.00	15.7	20	0.73
B	Overall	30.8	36	0.71	42.1	36	0.23
	Male	46.6	36	0.11	41.9	36	0.23
	Female	29.8	33	0.63	36.2	33	0.32
DRB	Overall	39.6	32	0.17	36.4	32	0.27
	Male	33.9	31	0.33	19.6	31	0.94
	Female	29.3	32	0.60	45.9	32	0.05
DQA	Overall	14.7	8	0.07	5.8	8	0.66
	Male	13.2	7	0.07	2.2	7	0.95
	Female	5.9	8	0.66	9.8	8	0.28
DQB	Overall	20.3	11	0.04	6.9	11	0.81
	Male	18.1	11	0.08	2.7	11	0.99
	Female	14.8	11	0.19	11.6	11	0.39
DPA	Overall	14.9	7	0.04	6.2	7	0.52
	Male	7.5	7	0.38	7.1	7	0.42
	Female	12.1	6	0.06	4.7	6	0.58
DPB	Overall	26.3	31	0.71	24.7	31	0.78
	Male	15.4	27	0.96	18.5	27	0.89
	Female	25.3	31	0.75	23.1	31	0.85

[1]The overall sample (N = 1,780) compares maternal vs. paternal transmitted or nontransmitted alleles using one affected offspring per family.
[2]The male sample (N = 1,376) compares maternal vs. paternal transmitted or nontransmitted alleles using one affected male offspring per family.
[3]The female sample (N = 1,291) compares maternal vs. paternal transmitted or nontransmitted alleles using one affected female offspring per family.

CHAPTER NINE

Summary and Directions for Future Research

SUMMARY

The investigation described in this dissertation had a dual focus. This was to test: 1) common genetic variation in two important autoimmune candidate genes and 2) human leukocyte antigen (HLA)-related risk factors for their involvement in disease susceptibility. Four different autoimmune diseases were studied.

This dissertation examined roles for the *CIITA* gene in rheumatoid arthritis (RA), systemic lupus erythematosus (SLE) and multiple sclerosis (MS) and the *CLEC16A* gene in RA. *CIITA* is a strong biological candidate for studies of autoimmunity because it encodes CIITA, a transcription factor required for HLA class II major histocompatibility complex (MHC)-restricted antigen presentation. Adjacent to *CIITA* on chromosome 16p13 lies *CLEC16A*, which was identified as a novel susceptibility locus for MS and type 1 diabetes (T1D) by recent genomewide association (GWA) studies. Both case-control and family-based study designs were utilized. Family-based studies are more robust, though not completely protected, against potential bias in the presence of population stratification. Case-control samples were limited to individuals of European ancestry to guard against potential spurious findings due to population stratification. Rather than rely solely on self-reported race and ethnicity, European ancestry was assessed through ancestry informative genetic markers whenever possible. This dissertation includes the first genetic study to fully characterize common genetic variation in *CIITA* and *CLEC16A*, including a comprehensive assessment of haplotypes.

This dissertation also examined HLA-related risk factors in SLE and T1D. HLA loci significantly influence disease risk for SLE and T1D, and might also influence risk through additional inherited or noninherited effects in addition to genetic transmission of particular risk alleles. Evidence for the role of increased maternal-offspring HLA compatibility, and both parent-of-origin and noninherited maternal HLA (NIMA) effects in autoimmune disease has been previously been reported. Conditional logistic regression was used to compare HLA compatibility relationships between mother-offspring and father-offspring. Family-based association tests were conducted to compare transmitted vs. nontransmitted alleles and both maternal vs. paternal transmitted (parent-of-origin) and nontransmitted alleles (NIMA). The study described in this dissertation is the largest, to date, to examine maternal-offspring HLA compatibility, parent-of-origin and NIMA effects in SLE and T1D.

Chapter 2 describes a meta-analysis of the *CIITA* -168A/G polymorphism and RA in 6,861 RA patients and 9,270 controls (N = 16,131 individuals) from ten case-control studies. Results from previous studies had yielded conflicted results. Results from this study did not reveal any evidence for association between the *CIITA* -168A/G polymorphism and RA. This suggested that future studies of *CIITA* and RA should cast a wider net rather than focusing solely on the promoter polymorphism, and incorporate clinical phenotypes such as the shared epitope and anti-cyclic citrullinated peptide (anti-CCP) positivity.

In Chapter 3, allelic and haplotype association tests of 31 *CIITA* SNPs in 2,542 RA cases and 3,690 controls (N = 6,232 individuals) were performed. Allelic association was tested by creating *2x2* contingency tables and estimating odds ratios (OR) with the Fisher's exact test. Haplotypes were estimated with the expectation maximization (EM) algorithm and maximum likelihood estimates (MLE) of haplotype probabilities were computed with the EM algorithm.

Score statistics were used for global haplotype association tests. A hidden Markov Model based algorithm was used to infer missing genotypes from known haplotypes. Results do not provide evidence that common variation in *CIITA* plays a role in susceptibility to RA.

Chapter 4 describes a study of 18 *CIITA* SNPs in 637 SLE trio families and 826 independent, unrelated SLE cases (*N* = 2,737 individuals). Family-based association tests were conducted to compare transmitted vs. nontransmitted alleles, and case-control association tests in the combined trios and unrelated cases (utilizing nontransmitted parental alleles as controls). Analyses were stratified by *DRB1*1501*, *DRB1*0301* and the secondary clinical phenotypes lupus nephritis, arthritis, serositis, neurological involvement, Sm and Ro autoantibody production. Case-only analyses of *DRB1*1501*, *DRB1*0301* and clinical phenotypes were conducted. No evidence for association was observed between *CIITA* and SLE in family-based and case-control analyses after correcting for multiple testing. Cases exhibited modest evidence for association between the rs11074938**G* variant and the presence of either *DRB1*1501* or **0301*. This is the first study of SLE to fully characterize common genetic variation in *CIITA*, including secondary phenotypes and HLA risk alleles. Results do not provide evidence that common variation in *CIITA* plays a role in susceptibility to SLE.

Chapter 5 includes a multi-stage investigation of variation within *CIITA, DRB1*1501* and MS in 6,108 individuals. In stage 1, 24 *CIITA* SNPs in 1,320 MS cases and 1,363 controls (*N* = 2,683 individuals) were tested for association utilizing a test of trend stratified by *DRB1*1501*. Rs4774 (missense +1614G/C; G500A) was associated with MS particularly in *DRB1*1501+* individuals. No association was observed for the -168A/G promoter variant. In stage 2, the rs4774 variant was tested in 973 extended families, utilizing the pedigree disequilibrium test (PDT) to compare transmitted vs. nontransmitted alleles. Rs4774**C* was associated with increased risk for MS in *DRB1*1501+* families. In stage 3, rs4774 was further tested in the cases and controls (stage 1) combined with one case per family (stage 2), for increased power. Rs4774**C* was associated with MS particularly in *DRB1*1501+* cases and controls. Results obtained from logistic regression analysis showed evidence for interaction between rs4774**C* and *DRB1*1501* associated with risk for MS when conditioned on the presence or absence of *CLEC16A* rs6498169**G*, a putative MS risk allele adjacent to *CIITA*. Strong evidence supporting a role for *CIITA* variation in MS risk was observed, which appears to depend on the presence of *DRB1*1501*.

In Chapter 6, a total of 251 *CLEC16A* SNPs were investigated for association in 2,542 RA cases (85% anti-CCP positive) and 2,210 controls (*N* = 4,752 individuals). Allelic association was tested by creating *2x2* contingency tables and estimating OR with the Fisher's exact test. Haplotypes were estimated with the EM algorithm and haplotype MLE probabilities were computed with the EM algorithm. Score statistics were used for global haplotype association tests. Evidence for association between *CLEC16A* variation and anti-CCP positive RA was not observed.

In Chapter 7, maternal-offspring HLA compatibility, parent-of-origin and NIMA effects at the *DRB1* locus was examined in 707 SLE-affected offspring trio families of self-reported European ancestry and 188 independent healthy maternal-offspring pairs of Dutch ancestry (*N* = 2,497 individuals). HLA maternal-offspring compatibility analyses were restricted to male and

nulligravid female SLE-affected offspring, and parent-of-origin and NIMA analyses were stratified by gender of SLE-affected offspring. Strong association was observed between SLE and the *DRB1* locus, as expected, in the overall and gender-stratified analyses. However, compatibility differences between mother-offspring and father-offspring were not observed. Further, parent-of-origin and NIMA HLA effects influencing SLE were not present: mothers of SLE patients had similar transmission and nontransmission frequencies for *DRB1* alleles when compared to fathers, including for known SLE risk alleles *DRB1*0301*, **1501* and **0801*. Results did not provide evidence to support a role for these HLA-related effects in SLE risk.

Finally, in Chapter 8, a study of maternal-offspring HLA compatibility, parent-of-origin and NIMA effects at the eight classical HLA loci in 2,271 affected sibling pair T1D families of self-reported European ancestry collected by the T1D Genetics Consortium (N =11,023 individuals) is described. Analyses were stratified by gender of T1D-affected offspring. Strong associations were observed for all classical HLA loci except for *DPA1*, as expected. Compatibility differences between mother-offspring and father-offspring were not observed for any HLA loci. Further, parent-of-origin and NIMA HLA effects influencing T1D were not present. Results did not provide evidence to support a role for these HLA-related effects in T1D risk.

DIRECTIONS FOR FUTURE RESEARCH

This dissertation investigated common variation in *CIITA* and *CLEC16A*. Future studies should consider rare variants in *CIITA* and *CLEC16A*, which were not directly investigated here. Based on the results from the study described in this dissertation, another study should be conducted to further replicate the association observed between the *CIITA* rs4774 missense (G/C) mutation (G500A) and MS in independent samples. Another important step is to conduct functional studies of the *CIITA* rs4774 variant. Rs4774 was originally discovered by Steimle et al. during their initial discovery of the *CIITA* gene, over a decade ago (1). Although Steimle et al. did not observe a functional difference between G500 and A500 in their study of bare lymphocyte syndrome (BLS), it would be worth repeating this functional study because minor differences may not have been detected (Steimle, personal communication). In MS, the relationship between variation at the *CIITA* locus and gene expression for both *CIITA* and MHC class II loci, as well as the resulting biological implications for the immune response and MS pathogenesis, are poorly understood. Large and comprehensive studies that can also fully explore clinical MS phenotypes and include environmental exposure data are needed.

This dissertation highlights the complementary nature of GWA studies, based on the International HapMap Project, and candidate gene studies (2). GWA studies have not previously identified *CIITA* as a susceptibility locus for MS (3-5). Further, results for *CIITA* analysis in the current study would not meet criteria for genome-wide significance. While GWA studies are attractive for many reasons, including that in principle they are ‘hypothesis free’, it is clear that experiments using current technology will be limited in their ability to identify the entire genetic contribution for most complex diseases, including MS (6). Nevertheless, candidate gene studies have historically failed to identify susceptibility loci with conclusive evidence. However, revisiting candidate gene studies with well-powered datasets and strong hypotheses based on

prior research remains an important strategy for disease gene identification. Investigators can incorporate both GWA and candidate gene study designs into their research (7).

GWA studies have recently confirmed the *CLEC16A* gene as a risk locus for MS (3, 8-10). Because *CLEC16A* is adjacent to *CIITA* on chromosome 16, linkage disequilibrium (LD) patterns between *CLEC16A* SNPs and *CIITA* SNPs were examined in various independent samples and did not demonstrate evidence of association between *CIITA* and *CLEC16A*. Based on patterns of LD derived from independent samples and results from comprehensive analyses, including logistic regression modeling, it does not appear that association observed between *CIITA* and MS, specifically the effect on disease risk conferred by rs4774, is due to *CLEC16A*. Nevertheless, it will be essential to account for *CLEC16A* in future studies of *CIITA* and MS.

Interestingly, *CIITA* has been reported to have reached genomewide significance in two recent GWA studies: ulcerative colitis (UC) and celiac disease (CD), both of which are autoimmune diseases (11, 12). McGovern et al. report evidence for association between the intronic, ancestral *CIITA* rs4781011**T* allele (minor allele) and risk of UC (OR = 1.23, $P = 3x10^{-6}$). Their initial study included 2,693 cases and 6,791 controls and their replication study included 2,009 cases and 1,580 controls, all of European descent (N = 13,073 individuals). The rs4781011 variant was not examined directly in the MS study described in this dissertation, but it was captured in the analysis (based on HapMap CEU, $r^2 \geq 0.8$). Animal models of colitis support a role for *CIITA* in UC. *CIITA* transgenic mice with over-expression of *CIITA* in helper T cells demonstrated aggravated oxazolone-induced colitis compared to wild type mice, due to elevated IL-4 production and Th2 inflammation (13).

Dubois et al. report evidence for association between the *CIITA-CLEC16A-SOCS1* gene region and increased risk of CD, with the strongest result being for the ancestral *C16orf75* rs12928822**C* allele (major allele, most likely intronic) and (OR = 1.16, 95% CI = 1.10 to 1.22, $P = 3x10^{-8}$). The *SOCS1* gene (suppressor of cytokine signaling 1) is adjacent to *CLEC16A*, as is *C16orf75*. Their initial study included 4,533 cases and 10,750 controls and their replication study included 4,918 cases and 5,684 controls, all of European descent (N = 25,885 individuals). Future studies could be expanded to examine the role of *CIITA* in additional autoimmune diseases (besides RA and SLE) and non-European populations. Animal studies could also shed some light on the function of *CIITA*.

To date almost 600 GWA studies have been published, and their contribution has been undoubtedly of great importance (14). Nevertheless recent critical evaluation of the field of genetic epidemiology points toward the need to also consider rare variants and structural variation and to continue conducting family and linkage studies in the future (15). Copy number variation (CNV) (sections of DNA ranging in size from kilo base pairs (kb) to mega base pairs (Mb) that vary in copy number) has been shown to vary considerably between individuals and has also been implicated in some autoimmune diseases, including SLE, RA, T1D, Crohn's disease, and psoriasis (16, 17). To date almost 58,000 CNVs (located in about 14,500 loci), 850 inversions and almost 31,000 insertion/deletions (100 base pairs (bp) to 1 kb) have been identified (18). The Database of Genomic Variants does not list any identified CNVs, inversions, insertions or deletions in *CIITA*, but it does list a large deletion (6,486 bp) in *CLEC16A* (rs36216218; 11,041,889-11,048,374 bp) and a deletion in *SOCS1* (11,256,545-11,258,445 bp) (19, 20).

Some effects such as parent-of-origin effects can only be studied in family studies. This dissertation examined HLA-related risk factors, including parent-of-origin effects, in a family study of SLE and T1D. It would be worthwhile for a future study of HLA-related risk factors in SLE to restrict analyses of maternal-offspring *HLA-DRB1* compatibility to SLE offspring that are either first-born males or first-born, nulligravid females. Analyses could also be expanded to examine all eight classical HLA loci. A future study of HLA-related risk factors in T1D study could test for differences in patterns of HLA compatibility in the classical HLA loci using an independent control sample of healthy mother-offspring pairs (without T1D), and also restrict analyses to T1D offspring that are either first-born males or first-born, nulligravid females. It would be worthwhile to further examine HLA-related risk factors in additional autoimmune diseases, such as MS.

Applying meta-analysis, candidate gene, epistatic (gene-gene interaction), epigenetic and pathway analysis to GWA studies can yield additional insight (21). For example, Eleftherohorinou et al. has examined inflammatory pathways for association with Crohn's disease, RA and T1D (22). They identified gene and SNP sets that seem to predict disease, and several genes that they identified were not significant in the original GWA study, but have since been confirmed in meta-analysis or candidate gene studies (22). A candidate gene analysis of a biological pathway suspected to be involved in a disease can also increase knowledge of disease pathogenesis, as demonstrated by Briggs et al. (23). Given the tendency for autoimmune diseases to cluster in individuals and families, another interesting area of research is to examine genetic associations with multiple autoimmune phenotypes rather than individual autoimmune disease phenotypes (24).

The study of RA and *CIITA* and *CLEC16A* described in this dissertation is a candidate gene study that utilized data from GWA studies. Because two genotyping platforms were combined (Affymetrix and Illumina) genotypes for untyped SNPs were imputed. Recent developments in genotype imputation methods allow the use of multiple control sets across different genotyping platforms and dense genome-wide haplotypes available from the 1,000 Genomes Project (25). The 1,000 Genome Projects is an international effort to sequence ≥2,000 individuals worldwide and offer a detailed catalogue of genetic variants (SNPs, CNVs, insertions and deletions) with frequencies of ≥1% (or 0.1-0.5% for genic variants) (http://www.1000genomes.org/). This will be useful for estimating population frequencies, LD and haplotype patterns, though stricter genetic antidiscrimination laws may need to be passed around the world to safeguard individual privacy (26).

REFERENCES

1. Steimle V, Otten LA, Zufferey M, Mach B. Complementation cloning of an MHC class II transactivator mutated in hereditary MHC class II deficiency (or bare lymphocyte syndrome). Cell 1993;75(1):135-46.
2. International_HapMap_Consortium. A haplotype map of the human genome. Nature 2005;437(7063):1299-320.
3. Hafler DA, Compston A, Sawcer S, Lander ES, Daly MJ, De Jager PL, et al. Risk alleles for multiple sclerosis identified by a genomewide study. N Engl J Med 2007;357(9):851-62.
4. ANZgene. Genome-wide association study identifies new multiple sclerosis susceptibility loci on chromosomes 12 and 20. Nat Genet 2009;41(7):824-8.
5. Baranzini SE, Wang J, Gibson RA, Galwey N, Naegelin Y, Barkhof F, et al. Genome-wide association analysis of susceptibility and clinical phenotype in multiple sclerosis. Hum Mol Genet 2009;18(4):767-78.
6. Altshuler D, Daly MJ, Lander ES. Genetic mapping in human disease. Science 2008;322(5903):881-8.
7. Musunuru K, Lettre G, Young T, Farlow DN, Pirruccello JP, Ejebe KG, et al. Candidate Gene Association Resource (CARe): Design, Methods, and Proof of Concept. Circ Cardiovasc Genet.
8. Rubio JP, Stankovich J, Field J, Tubridy N, Marriott M, Chapman C, et al. Replication of KIAA0350, IL2RA, RPL5 and CD58 as multiple sclerosis susceptibility genes in Australians. Genes Immun 2008;9(7):624-30.
9. Zoledziewska M, Costa G, Pitzalis M, Cocco E, Melis C, Moi L, et al. Variation within the CLEC16A gene shows consistent disease association with both multiple sclerosis and type 1 diabetes in Sardinia. Genes Immun 2009;10(1):15-7.
10. Hoppenbrouwers IA, Aulchenko YS, Janssens AC, Ramagopalan SV, Broer L, Kayser M, et al. Replication of CD58 and CLEC16A as genome-wide significant risk genes for multiple sclerosis. J Hum Genet 2009;54(11):676-680.
11. McGovern DP, Gardet A, Torkvist L, Goyette P, Essers J, Taylor KD, et al. Genome-wide association identifies multiple ulcerative colitis susceptibility loci. Nat Genet;42(4):332-7.
12. Dubois PCA, Trynka G, Franke L, Hunt KA, Romanos J, Curtotti A, et al. Multiple common variants for celiac disease influencing immune gene expression. Nat Genet;advance online publication.
13. Kim TW, Park HJ, Choi EY, Jung KC. Overexpression of CIITA in T cells aggravates Th2-mediated colitis in mice. J Korean Med Sci 2006;21(5):877-82.
14. Hindorff L.A., Junkins H.A., Hall P.N., Mehta J.P., T.A. M. A Catalog of Published Genome-Wide Association Studies. Available at www.genome.gov/gwastudies. Accessed 6/7/2010
In.
15. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, et al. Finding the missing heritability of complex diseases. Nature 2009;461(7265):747-53.
16. Sebat J, Lakshmi B, Troge J, Alexander J, Young J, Lundin P, et al. Large-scale copy number polymorphism in the human genome. Science 2004;305(5683):525-8.
17. Schaschl H, Aitman TJ, Vyse TJ. Copy number variation in the human genome and its implication in autoimmunity. Clin Exp Immunol 2009;156(1):12-6.
18. Iafrate AJ, Feuk L, Rivera MN, Listewnik ML, Donahoe PK, Qi Y, et al. Detection of large-scale variation in the human genome. Nat Genet 2004;36(9):949-51.

19. Mills RE, Luttig CT, Larkins CE, Beauchamp A, Tsui C, Pittard WS, et al. An initial map of insertion and deletion (INDEL) variation in the human genome. Genome Res 2006;16(9):1182-90.
20. Conrad DF, Pinto D, Redon R, Feuk L, Gokcumen O, Zhang Y, et al. Origins and functional impact of copy number variation in the human genome. Nature;464(7289):704-12.
21. Cantor RM, Lange K, Sinsheimer JS. Prioritizing GWAS Results: A Review of Statistical Methods and Recommendations for Their Application.86(1):6-22.
22. Eleftherohorinou H, Wright V, Hoggart C, Hartikainen AL, Jarvelin MR, Balding D, et al. Pathway analysis of GWAS provides new insights into genetic susceptibility to 3 inflammatory diseases. PLoS One 2009;4(11):e8068.
23. Briggs FB, Goldstein BA, McCauley JL, Zuvich RL, De Jager PL, Rioux JD, et al. Variation Within DNA Repair Pathway Genes and Risk of Multiple Sclerosis. Am J Epidemiol.
24. Barcellos LF, Kamdar BB, Ramsay PP, DeLoa C, Lincoln RR, Caillier S, et al. Clustering of autoimmune diseases in families with a high-risk for multiple sclerosis: a descriptive study. Lancet Neurology 2006;5(11):924-931.
25. Howie BN, Donnelly P, Marchini J. A Flexible and Accurate Genotype Imputation Method for the Next Generation of Genome-Wide Association Studies. PLoS Genet 2009;5(6):e1000529.
26. Via M, Gignoux C, Burchard EG. The 1000 Genomes Project: new opportunities for research and social challenges. Genome Med;2(1):3.