To my parents, Vicki and John Zuvich,

my biggest fans, for their infinite support.

In loving memory of my grandmother, Vera Laudermilch,

who was a woman ahead of her time,

for her great example and encouragements.

ACKNOWLEDGEMENTS

This work would not be possible without the financial support of the Vanderbilt University Center for Human Genetics Research (CHGR) and NIH Grant R01NS32830. I would like to thank the Multiple Sclerosis Genetics Group (MSGG) and the International Multiple Sclerosis Genetics Consortium (IMSGC) for their financial support and resources. I would like to thank Dr. Phil De Jager for his compilation of various datasets for follow-up studies. Also, I would like to thank all the study participants without which none of this would be possible.

I am extremely indebted to my mentor, Dr. Jonathan Haines, who has helped mold me into a competent scientist while still allowing me the independence to be myself. He has supported me and provided guidance despite my non-traditional ways. Dr. Haines also has allowed me to explore other interests outside the realm of traditional biomedical research. He has shown me that you can reach the top through collaborative, rather than cutthroat, efforts. I am very grateful for his encouragements and for his example in showing me how to be the best scientist and person I can be. I respect his opinion and look forward to working with him in the future.

I am also very thankful to the rest of my committee. Dr. Marylyn Ritchie, the chair of my committee, has been a great support system in all aspects of life. She has shown me how to properly balance personal and professional facets. I am thankful to Dr. Dana Crawford for her expertise in population genetics. I am thankful to Dr. Michael Rock for his knowledge of immunology and his friendly reminders to avoid the genetics jargon. And finally, I am grateful to Dr. Subramaniam Sriram for his devotion to his MS patients. I appreciate the quality time he took to help

navigate me around the MS clinic and for the opportunity to meet with his patients to get a better understanding of the personal effects of this disabling disease.

I appreciate all the work of the Computational Genomes Core within the CHGR, in particular Yuki Bradford, Justin Giles, Lana Olson, and David Sexton. Yuki has helped me with various aspects of the MS project. Justin has been very helpful with helping me troubleshoot various statistic packages. Lana patiently taught me how to use the software and answered my numerous questions. Finally, the Core would not run so smoothly without David's leadership and organizational skills. I would also like to thank Dr. Doug Mortlock for his help with the molecular aspects of this project. I also am very grateful to Dr. Jeff Canter who has been a great inspiration to me and provided endless humor, even though most of it revolves around picking on me!

I also want to thank the members of the Haines Lab. Dr. Nathalie Schnetz-Boutaud has been like a mother figure in the lab. Whenever I had personal or professional problems, I could always turn to her for encouragement and advice. I am very grateful for her leadership, and she will surely be a lifelong friend. Ping Mayo, our lab manager, has been a great example of someone with a kind heart and a strong work ethic. I greatly admire her dedication. I am very grateful for Melissa Allen. She has always provided me with a steady mind when I overreacted, which she can tell you happens a lot! I am also thankful to have her as a lunch buddy, a crossword puzzle enthusiast, and a very special friend. Special thanks to the other graduate students in the Haines Lab: Anna (Davis) Cummings, Mary Davis, and Olivia Veatch. Anna has been a great friend and bay-mate over the past 2 and a half years. I am especially grateful for her kind words, her level head, and her encouragement. Mary and Olivia, the newest members of the lab, have already

become great friends in such a short time. Their ambition and zest for research has helped renew my enthusiasm over the stressful times towards the end of my graduate career. Finally, two other lab members (in spirit), Dr. Deborah Murdock and Laura Hunt, have been great friends and provided countless laughs through the years. I am especially grateful for all these wonderful women who have made it a pleasure coming into work day after day. And I'm amazed and privileged that in a male-dominated field that I could be surrounded by such brilliant, beautiful, well-rounded women. I am so thankful for each of every one of them and will surely never find such a supportive working environment elsewhere.

There are also a few men associated with the Haines lab, past and present. I am thankful for Brent Anderson, our former lab manager, for his countless stories. He always kept working in the lab fun. Dr. Jacob McCauley, a former postdoctoral fellow and current faculty at University of Miami, has been a wonderful competitor, collaborator, and friend. I am thankful for Jacob's healthy competitive nature and his ability to keep me on my toes. I have enjoyed working with him on various MS projects, and I hope our professional career paths cross in the future. Finally, Dr. Will Bush, a current postdoctoral fellow, has provided me with vast insight and has been an enormous source of inspiration to me. I have enjoyed his help throughout my graduate work.

I am also very thankful to the students (past and present) of the CHGR for their encouragement. They have helped support my growth into the scientist I have become. I am also grateful for my friends from my Interdisciplinary Graduate Program (IGP) class who will likely be lifelong friends. I am especially grateful for my fellow Human Genetics classmates, Lauren Walters and Meaghan Neill. They have been the best friends a person could ask for. I know I will see these ladies at

every major milestone in my life and am especially lucky to have found them. Despite Lauren's devotion to the Buckeyes, we became friends, and she has always provided me with constructive criticism. I am thankful for her ability to speak the truth no matter what. Meaghan's bubbly personality and gentle ways made us instant friends. I am so thankful to have found someone so full of life. I appreciate her ability to calm me down in any situation. These ladies came along at a time in my life when I needed them the most, and I will forever be grateful and loyal to them.

Finally, and most importantly, I want to thank my family, especially my parents. When I moved 800+ miles away from home, I never knew how much I would miss my parents. My mom has truly become a best friend to me. They are my biggest fans, my cheerleaders. They have believed in me when it was hard for me to believe in myself. I am so lucky to have amazingly supportive parents who have encouraged me every step of the way. They are smart, wonderful individuals who have given me the skills to navigate my own path in this world and who have also given me the means to do just that. I will forever be indebted to them for this.

TABLE OF CONTENTS

DEDICATION ii

ACKNOWLEDGEMENTS iii

LIST OF TABLES x

LIST OF FIGURES xvi

LIST OF ABBREVIATIONS xvii

LIST OF WEBSITES xxiv

Chapter

I. BACKGROUND ON MULTIPLE SCLEROSIS 1

Basics of Multiple Sclerosis 1
Pathogenesis 4
Risk Factors 5
Diagnosis 8
Genetics- Early Success 13
Genetic Challenges Beyond HLA 14
New Genetic Approaches 16
Looking At Common Versus Rare Variation 18
Gene-gene And Gene-environment Interactions 20

II. CANDIDATE PATHWAY APPROACH 22

Precedent For Pathway Approach 22
Selection Process For Pathway Genes 23

III. METHODS AND RESULTS 34

Introduction 34
SNP Selection 34
Samples 35
Genotyping Techniques 39
Quality Control (QC) 40
Analysis Plan 42
Main Results (Armitage Trend Test) 45

Genotypic Analysis 83
Haplotype Analysis and MDR Analysis 83

IV. SECONDARY ANALYSES 89

Introduction 89
HLA Stratification 89
Justifications 89
HLA+ Stratum Results 91
HLA- Stratum Results 96
Discussion 102
Gender Stratification 106
Justifications 106
Female Stratum Results 106
Male Stratum Results 112
Discussion 121

Other Stratifications 125

IL7RA Stratification 125
Justifications 125
IL7RA+ Stratum Results 126
IL7RA- Stratum Results 131
Discussion 132
RRMS Stratification 134
Justifications 134
RRMS/SPMS Stratum Results 134
Discussion 139
MSSS Stratification 139
Justification 139
Results 140

Conclusions 143

V. CHROMOSOME 16, A CLOSER LOOK 149

Introduction 149
Genotyping Techniques 151
Analysis Plan 151
Results 152
Genomic Annotation 166

Discussion 167

VI. CONCLUSIONS AND FUTURE PLANS 170

Conclusions 170
Autoimmunity Disease Loci 175
Future Studies 178

REFERENCES 181

LIST OF TABLES

Table

1.1 EDSS measurements and correlated disability....11

1.2 Global Multiple Sclerosis Severity Scores 12

1.3 Nine replicated non-MHC genes associated with multiple sclerosis....18

2.1 Genes Related to the IL7/IL7RA Pathway....29

3.1 Demographics of Case-control dataset from initial screen using Illumina 60K BeadArray....37

3.2 Demographics of Case-control dataset from replication screen....39

3.3 Armitage Trend Test Results for Illumina screen....47

3.4 Armitage Trend Test Results for Replication Study....49

3.5 Armitage Trend Test Results for the Combined Datasets 51

3.6 Armitage Trend Test Results for Replication Study #2....53

3.7 Armitage Trend Test Results for all 578 significant SNPs from Illumina 60K screen....54

3.8 Genotypic chi square results for Illumina 60K screen....84

3.9 Haplotype Results for IL7 in the 60K Illumina screen....87

3.10 Armitage Trend Test for Most Significant SNP in the IL7 Haplotype 87

3.11 Haplotype Results for IL4R in the 60K Illumina screen 88

3.12 Armitage Trend Test for Most Significant SNP in the IL4R Haplotype 88

4.1 Most Significant SNPs from Armitage trend test for HLA+ stratum only 93

4.2 Most Significant genotypic chi square analysis results for HLA+ stratum only 94

4.3 Haplotype Results for SOCS1 in the HLA+ stratum 95

4.4 Armitage Trend Test for Most Significant SNP in the SOCS1 Haplotype in the HLA+ stratum 95

4.5 Most Significant SNPs from Armitage trend test for HLA- stratum only 97

4.6 Most Significant genotypic chi square analysis results for HLA- stratum only 99

4.7 Haplotype Results for SOCS2 in the HLA- stratum 100

4.8 Armitage Trend Test for Most Significant SNP in the SOCS2 Haplotype in the HLA- stratum 100

4.9 Stratified HLA data for the combined datasets 103

4.10 Second follow-up HLA stratified Armitage trend test results 105

4.11 Most Significant SNPs from Armitage trend test for female stratum only 108

4.12 Most Significant genotypic chi square analysis results for female stratum only 109

4.13 Haplotype Results for IL4R in the female stratum 110

4.14 Armitage Trend Test for Most Significant SNP in the IL4R Haplotype in the female stratum 110

4.15 Haplotype Results for IRS1 in the female stratum.....111

4.16 Armitage Trend Test for Most Significant SNP in the IRS1 Haplotype in the female stratum 111

4.17 Most Significant SNPs from Armitage trend test for male stratum only 113

4.18 Most Significant genotypic chi square analysis results for male stratum only 115

4.19 Haplotype Results for MAPK14 in the male stratum.....118

4.20 Armitage Trend Test for Most Significant SNP in the MAPK14 Haplotype in the male stratum 118

4.21 Haplotype Results for SKAP2 in the male stratum 119

4.22 Armitage Trend Test for Most Significant SNP in the SKAP2 Haplotype in the male stratum 119

4.23 Haplotype Results for PTK2B in the male stratum 120

4.24 Armitage Trend Test for Most Significant SNP in the PTK2B Haplotype in the male stratum 120

4.25 First follow-up gender stratified Armitage trend test results 122

4.26 Second follow-up gender stratified Armitage trend test results 124

4.27 Power calculation for replication of female only results 125

4.28 Power calculation for replication of male only results 125

4.29 Most Significant SNPs from Armitage trend test for IL7RA+ stratum only 127

4.30 Most Significant genotypic chi square analysis results for IL7RA+ stratum only 128

4.31 Haplotype Results for Stat5b in the IL7RA+ stratum 130

4.32 Armitage Trend Test for Most Significant SNP in the Stat5b Haplotype in the IL7RA+ stratum 130

4.33 Most Significant genotypic chi square analysis results for IL7RA- stratum only 131

4.34 Haplotype Results for NFATC2 in the IL7RA- stratum 133

4.35 Armitage Trend Test for Most Significant SNP in the NFATC2 Haplotype in the IL7RA- stratum.....133

4.36 Most Significant SNPs from Armitage trend test for RRMS/SPMS stratum only 136

4.37 Most Significant genotypic chi square analysis results for RRMS/SPMS stratum only 137

4.38 Haplotype Results for IL4R in the RRMS/SPMS stratum 138

4.39 Armitage Trend Test for Most Significant SNP in the IL4R Haplotype in the RRMS/SPMS stratum 138

4.40 Most significant SNPs from all 828 cases with MSSS scores.....141

4.41 Most significant SNPs from 451 cases with extreme MSSS scores.....142

4.42 Most Significant Armitage trend test results from all secondary analyses 145

5.1 Armitage Trend Test for all chromosome 16 SNPs from 60K BeadChip assay.....154

5.2 Summary of most significant SNPs from each Region that were used in logistic regression models.....163

5.3 Linkage disequilibrium values for these most significant SNPs 163

5.4 List of interaction p-values for each pairwise combination of SNPs in the overall dataset 164

5.5 List of interaction p-values for each pairwise combination of SNPs in the HLA+ stratum.....165

5.6 List of interaction p-values for each pairwise combination of SNPs in the HLA- stratum.....165

6.1 Genes shared by several autoimmune diseases.....177

LIST OF FIGURES

Figure

1.1 Types of MS 4

1.2 Lifetime risk of MS 13

2.1 IL7/IL7RA pathway genes 25

3.1 Flowchart of QC Process 42

5.1 Linkage Disequilibrium for chromosome 16 region (D' values) 161

5.2 Linkage Disequilibrium for chromosome 16 region (r^2 values) 162

5.3 The conserved synteny between these chromosome 16 genes 166

LIST OF ABBREVIATIONS

3UTR	3’ untranslated region
5UTR	5’ untranslated region
ACP	Accelerated Cure Project
AITD	autoimmune thyroid disease
AS	ankylosing spondylitis
BWH	Brigham and Women’s Hospital
CD	Crohn’s disease
CD4+	“helper” T cells
CD8+	cytotoxic T cells
CDCV	common disease/common variant hypothesis
CEU	US residents with northern and western European ancestry collected by the Centre d’Etude du Polymorphisme Humain (CEPH)
CHB	Chinese people from Beijing
CIS	clinically isolated syndrome
CMS	University of Cambridge, England
CNS	central nervous system
CSF	cerebrospinal fluid
dbSNP	database of single nucleotide polymorphisms
DNA	deoxyribonucleic acid
EDSS	Extended Disability Status Score
FS	functional systems
GD	Graves’ disease
GRRs	genotype relative risks
GWAS	genome-wide association studies

HLA	human leukocyte antigen
HWE	Hardy-Weinberg equilibrium
IgG	immunoglobulin G
IMSGC	International Multiple Sclerosis Genetics Consortium
JPT	Japanese people from Tokyo area
kb	kilobases
LD	linkage disequilibrium
MALDI TOF MS	Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry
MDR	Multi-factor dimensionality reduction
MHC	Major Histocompatibility Complex
MRI	magnetic resonance imaging
MRV	multiple rare variants hypothesis
MS	multiple sclerosis
MSSS	Multiple Sclerosis Severity Score
nsSNP	nonsynonymous single nucleotide polymorphism
OR	odds ratio
PCR	polymerase chain reaction
PPMS	primary progressive multiple sclerosis
PRMS	progressive relapsing multiple sclerosis
PSA	psoriatic arthritis
PSO	psoriasis
QC	quality control
RA	rheumatoid arthritis
RRMS	relapsing-remitting multiple sclerosis
rs	reference sequence (NCBI)
RUSH	Rush University of Chicago

SLE	systemic lupus erythematosus
SNP	single nucleotide polymorphism
SPMS	secondary progressive multiple sclerosis
T1D	type 1 diabetes
Th1	T helper cells type 1
Th2	T helper cells type 2
Th17	T helper cells type 17
UCSF	University of California San Francisco
UK	United Kingdom
US	United States
WashU	Washington University
WASP	Whole-genome Association Study Pipeline
YRI	Yoruba people of Ibadan, Nigeria

Nucleotide Abbreviations:

A	adenine
C	cytosine
G	guanine
T	thymine

Amino Acid Abbreviations:

A	alanine
G	glycine
I	isoleucine
P	proline
Q	glutamine

R	arginine
S	serine
T	threonine
Y	tyrosine

Gene Abbreviations:

AKT	v-akt murine thymoma viral oncogene homolog 1
AP-1	jun oncogene (JUN)
Bad	Bcl-2-associated agonist of cell death
Bax	Bcl-2-associated X protein
Bcl-2	B-cell CLL/lymphoma 2
Bcl-2L1	Bcl-2-like 1
Bcl-3	B-cell CLL/lymphoma 3
BID	BH3 interacting domain death agonist
Bim	Bcl-2-like 11 (apoptosis facilitator) (Bcl2l11)
c-myc	myelocytomatosis oncogene homolog
CIITA	major histocompatibility complex class II transactivator
C16orf75	chromosome 16 open reading frame 75
CD6	Cluster of Differentiation 6
CD58	Cluster of Differentiation 58
CD226	Cluster of Differentiation 226
Cdc25a	cell division cycle 25 homolog A
CLEC16A	C-type lectin domain family 16 member A
CPLA2	cytosolic phospholipase a2
CRLF2	cytokine receptor-like factor 2 (thymic stromal lymphopoietin receptor)
CSK	c-src tyrosine kinase

GCET2	germinal center expressed transcript 2
GSK-3	frequently rearranged in advanced T-cell lymphomas (FRAT1)
HLA	human leukocyte antigen
IL1R2	interleukin-1-receptor type II
IL2	interleukin-2
IL2RA	interleukin-2 receptor subunit alpha
IL2RG	interleukin-2 receptor subunit gamma (gamma c)
IL4	interleukin-4
IL4R	interleukin-4 receptor
IL5	interleukin-5
IL7	interleukin-7
IL7RA	interleukin-7 receptor subunit alpha
IL9	interleukin-9
IL12A	interleukin-12 subunit alpha
IL12B	interleukin-12 subunit beta
IL12RB1	interleukin-12 receptor, beta 1
IL12RB2	interleukin-12 receptor, beta 2
IL13	interleukin-13
IL13RA1	interleukin-13-receptor alpha 1
IL15	interleukin-15
IL17	interleukin-17
IL17A	interleukin-17 subunit alpha
IL17RA	interleukin-17 receptor subunit alpha
IL17RD	interleukin-17 receptor subunit delta
IL18	interleukin-18
IL21	interleukin-21
IL23	interleukin-23

IL27RA	interleukin-27-receptor subunit alpha
INFG	interferon gamma
IRF1	interferon regulatoyr factor 1
IRF8	interferon regulatory factor 8
IRS1	insulin receptor substrate 1
IRS2	insulin receptor substrate 2
Jak1	Janus kinase 1
Jak2	Janus kinase 2
Jak3	Janus kinase 3
LAG3	lymphocyte-activation gene 3
LOC400499	hypothetical protein 400499
LOC729954	hypothetical protein 729954
MCL-1	myeloid cell leukemia sequence 1
NFATC1	nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 1
NFATC2	nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 2
p27	Cyclin-dependent kinase inhibitor member 1B (*CDKN1B*)
p38	mitogen-activated protein kinase 14 (*MAPK14*)
PI3K	phosphoinositide-3-kinase (alpha, beta, gamma, delta subunits)
PLA2G4A	phospholipase A2, group IVA
PLA2G4B	phospholipase A2, group IVB
PLA2G4C	phospholipase A2, group IVc
PRKCA	protein kinase C subunit alpha
PRKCE	protein kinase C subunit epsilon
PRM1/2/3	protamine 1, 2, and 3
PTEN	phosphatase and tensin homolog

PTK2B	protein tyrosine kinase 2 beta
SHC1	SHC (Src homolog 2 domain containing) transforming protein 1
SHIP	inositol polyphosphate-5-phosphatase (*INPP5D*)
SKAP2	src kinase associated phosphoprotein 2
SOCS	suppressors of cytokine signaling (*SOCS1-7*)
Stat1	signal transducer and activator of transcription 1
Stat3	signal transducer and activator of transcription 3
Stat5a	signal transducer and activator of transcription 5A
Stat5b	signal transducer and activator of transcription 5B
Stat6	signal transducer and activator of transcription 6
TBET	Th-1-specific T box factor
TNFRSF1A	tumor necrosis factor receptor superfamily member 1A
TNP2	transition protein 2
TSLP	thymic stromal lymphopoietin
TYK2	tyrosine kinase 2

Statistic Abbreviations/Symbols:

$\lambda_{\square}$	recurrence risk ratio for first degree relatives

LIST OF WEBSITES

http://genome.ucsc.edu

http://www.ensembl.org/index.html

http://www.r-project.org/

www.genome.gov/19518660#3

www.mult-sclerosis.org

www.ncbi.nlm.nih.gov