Assessment of the complexities of genetic contributions to Ankylosing Spondylitis

by

©Alexandra Munn

A thesis submitted to the

School of Graduate Studies

in partial fulfilment of the requirements for the degree of

Master of Science in Medicine

Discipline of Clinical Epidemiology

Memorial University of Newfoundland

May 2017

St. John's, Newfoundland

Abstract

Objectives

This research assesses whether structural variations can predict disease expression and whether gene-environment interactions can predict radiographic severity and progression.

Methods

In a cohort of 896 AS patients the associations of *UGT2B17* CNV with *HLA B*27* and extra-articular features were compared using the Chi- Squared test.

A second cohort of 76 AS patients were assessed using multiple linear regression analysis of serial radiographic assessments and SNPs from three genes related to smoking metabolism.

Results

No association was found between having *UGT2B17* CNV 2 and IBD (p-value = 0.155), psoriasis (p-value = 0.481), uveitis (p-value = 0.466) or *HLA-B*27* positivity (p-value = 0.118).

Two SNPs from *CHRNA7* were associated with radiographic severity: rs11071593 and rs7178176 (p-value = 0.0005). Two SNPs from *CHRNA7* were associated with radiographic progression: rs2337506 and rs12910885 (p-value < 0.0001).

Conclusion

No association was found between *UGT2B17* CNVs and disease expression; however, we have demonstrated an association between *CHRNA7* SNPs and AS radiographic severity and progression.

Acknowledgements

First and foremost, I would like to extend my thanks to my supervisor, Dr. Proton Rahman, for his expert knowledge, guidance and patience throughout my research. Without your incredible support this thesis would have been too overwhelming a pursuit to complete in medical school. Dr. Darren O'Reilly: thank you so much for your continuous assistance and availability. Likewise, to Dr. Gerry Mugford: thank you for your input and revisions.

I would also like to thank Dianne Codner for her assistance in the lab.

To Rose Ardern: thank you for making meetings and scheduling run smoothly. You made everything that much easier.

To my friend, Claire Woodworth, for teaching me to edit; thank you for tolerating with my horrendous grammatical errors, thank you for teaching me the "oxford comma" and, most importantly, thank you for learning to say "UGT2B17" flawlessly.

Finally, I would like to thank my parents and my uncle for their unconditional support throughout these last four years. I would not have been able to complete this thesis without their encouragement.

Table of Contents

Abstract ii

Acknowledgements iii

Table of Contents iv

List of Tables vi

List of Figures viii

List of Symbols, Nomenclature or Abbreviations ix

List of Appendices xi

Chapter 1: Introduction 1
- 1.1: Background of Ankylosing Spondylitis 2
- 1.2: Genetic Analysis Studies 5
  - 1.2.1: Heritability of AS 5
  - 1.2.2: Genetic Studies 6
    - 1.2.2a: Linkage Studies 6
    - 1.2.2b: Genome-Wide Association Studies 7
- 1.3: Review of Major AS Genes 9
  - 1.3.1: Immunogenetics 9
    - 1.3.1a: *HLA-B*27* 9
  - 1.3.2: Gene Variants Identified via GWAS 10
    - 1.3.2a: *IL-23…* 11
    - 1.3.2b: *IL-12B…* 12
    - 1.3.2c: *ERAP1* 12
    - 1.3.2d: *CARD9* 13
    - 1.3.2e: *STAT3…* 13
    - 1.3.2f: *TYK2* 14
    - 1.3.2g: *RUNX3…* 14
    - 1.3.2h: *IL-6…* 15
    - 1.3.2i: *IL-27…* 15
  - 1.3.3: Missing Heritability 17
    - 1.3.3a: Copy Number Variants 17
    - 1.3.3b: Gene-Gene Interactions 20
    - 1.3.3c: Gene-Environment Interaction: Smoking and AS 21
    - 1.3.3d: Genetics of Disease Expression: Genetic Association with Extra Articular Features 22
    - 1.3.3e: Rare Variants 25
- 1.4: Relevance of Research 25
  - 1.4.1: Study Objectives 26

Chapter 2: Materials and Methods 28

2.1: Ethical/Administrative Approval 29
2.2: Study Design 29
2.3: AS Patient Ascertainment 29
2.3.1: Location 29
2.3.2: Diagnostic Criteria 30
2.4: Data Collection 31
2.4.1: Clinical Data 31
2.4.2: Genetic Data 34
2.4.2a: Part 1: CNV Information: *UGT2B17* Gene 34
2.4.2b: Part 2: Selection of Related Smoking SNPs 34
2.4.2c: Part 2: Genotyping of Related Smoking SNPs 36
2.5: Sample Size 37
2.6: Statistical Analysis 38
2.6.1: Data Organization and Coding 38
2.6.1a: Variables: Extra-Articular Features, *HLA-B*27* and *UGT2B17..38*
2.6.1b: Variables: Smoking SNPs and Attached Clinical Information...40
2.6.2: Variables: Data Filtering 40
2.6.3: Data Analysis 41
2.6.3a: Part 1: *UGT2B17* CNV and Extra-Articular Features in SpA 42
2.6.3b: Part 1: *UGT2B17* CNV and *HLA-B*27* 42
2.6.3c: Part 2: Disease Severity and Genotypes 42
2.6.3d: Part 2: Disease Progression and Genotypes 43

Chapter 3: Results 44
3.1: Part 1: UGT2B17 CNV and extra- articular features in SpA 45
3.1.1: Demographics 45
3.1.2: *UGT2B17* CNV and AS Extra-Articular Features 46
3.1.3: Gene-Gene Interaction: *UGT2B17* and *HLA-B*27* 50
3.2: Part 2: Genetics of Radiographic Severity and Progression 52
3.2.1: Demographics 52
3.2.2: Disease Severity and Smoking SNPs 53
3.2.3: Disease Progression and Smoking SNPs 56

Chapter 4: Discussion 61
4.1: Discussion 62
4.2: Conclusion 75

References 76

Appendix 1: Additional Gene Variants Identified via GWAS 92

Appendix 2: Results Displaying Inheritance Models (Dominant, Recessive & Additive)..... 100

List of Tables

Table 1-1 Genes found to be significantly associated with AS through SNPL based GWA studies……..……...15

Table 1-2 Genes found to be significantly associated with AS through CNV-based GWAS……..……20

Table 2-1 1984 Modified New York Criteria for diagnosis of AS……...……...31

Table 2-2 SNP IDs with associated genes 35

Table 2-3 Sample size calculations based on varying power indexes for Part 1 38

Table 3-1 Demographic information for AS patients in a cohort of 896 patients……..46

Table 3-2 CNV Frequency in the three populations analyzed……..…..…..46

Table 3-3 Contingency table of CNV categories and history of IBD used in the Chi- square Fisher's exact test analysis……..…….47

Table 3-4 Contingency table of CNV categories and history of psoriasis used in the Chi-square Fisher's exact test analysis……...……48

Table 3-5 Contingency table of CNV categories and history of uveitis used in the Chi- square Fisher's exact test analysis……..…….49

Table 3-6 Contingency table of CNV categories and *HLA-B*27* status used in the Chi- square Fisher's exact test analysis…….…..50

Table 3-7 Frequency of individuals having ever presented with each extra-articular feature and *HLA-B*27* in each CNV category with associated p-values from statistical analysis using the Chi-squared Fisher's exact test…….51

Table 3-8 Demographics of AS patients in a cohort of 75 individuals with serial mSASS scores……...……52

Table 3-9 Disease severity in association with SNPs for genotypes……...54

Table 3-10 Disease severity and interaction with smoking in association with SNPs for genotype…….54

Table 3-11 Disease severity and interaction with gender in association with SNPs for genotype……..54

Table 3-12 SNPs from the *CHRNA7* gene having significant association with disease severity……..56

Table 3-13 Disease progression in association with SNPs for genotypes..... 57

Table 3-14 Disease progression and interaction with smoking in association with SNPs for genotype..... 57

Table 3-15 Disease progression and interaction with smoking in association with SNPs for genotype..... 58

Table 3-16 SNPs from the *CHRNA7* gene having significant association with disease progression…..... 59

Table 3-17 Comparison of progression values between non-smokers and smokers with the addition of C allele in the genotype for the rs1800925 SNP..... 59

Table 3-18 Comparison of mSASSS1 values between non-smokers and smokers with GG, GT, and TT genotypes for the rs61750900 SNP……...60

Table A-1 Interaction between selected genetic variants, radiographic severity (mSASSS) and both smoking status and sex……101

Table A-2 Interaction between selected genetic variants, radiographic progression (change in mSASSS/time) and both smoking status and sex……...103

Table A-3 Three inheritance models (additive, dominant and recessive) examined to determinate the association between each SNP and the mean values of mSASSS1 and progression……106

List of Figures

Figure 3-1: Proportions of individuals with 2 copies of the *UGT2B17* gene having self-reported IBD and never having self-reported IBD 47

Figure 3-2: Proportions of individuals with 2 copies of the *UGT2B17* gene having self-reported psoriasis and never having self-reported psoriasis 48

Figure 3-3: Proportions of individuals with 2 copies of the *UGT2B17* gene having self-reported uveitis and never having self-reported uveitis 49

Figure 3-4: Proportions of individuals with 2 copies of the *UGT2B17* gene with a positive *HLA-B*27* status and a negative *HLA-B*27* status 51

List of Symbols, Nomenclatures or Abbreviations

AAU	Acute anterior uveitis
ACh	Acetylcholine
ANTXR2	Anthrax toxin receptor 2
AS	Ankylosing spondylitis
AxSpA	axial Spondyloarthritis
BACH2	Transcription regulator protein BACH2
BiP	Binding immunoglobulin protein
CARD9	Caspase recruitment domain-containing protein 9
CD4+ cells	Cluster of differentiation 4 cells
CD8+ cells	Cluster of differentiation 8 cells
CHRNA7	Cholinergic receptor, nicotinic, alpha 7
CNV	Copy number variation
DMARDS	Disease modifying anti-rheumatic drugs
DNA	Deoxyribonucleic acid
DRGLOBAL	Physician global health rating
EOMES	Eomesodermin
ER	Endoplasmic reticulum
ERAP1	Endoplasmic reticulum aminopeptidase 1
GPR35	G protein-coupled receptor 35
GPR37	G protein-coupled receptor 37
GPR65	G protein-coupled receptor 65
GWAS	Genome wide association studies
HLA-A*0201	Human leukocyte antigen A02
HLA-B*27	Human leukocyte antigen B27
HLA-C*06:02	Human leukocyte antigen C06:02
HWE	Hardy Weinberg equilibrium
IBD	Inflammatory bowel disease
IGAS	International Genetics of Ankylosing Spondylitis
IL1	Interleukin 1
IL1R1	Interleukin 1 receptor, type I
IL1R2	Interleukin 1 receptor, type II
IL6	Interleukin 6
IL6R	Interleukin 6 receptor
IL7R	Interleukin 7 receptor
IL17	Interleukin 17
IL12B	Interleukin 12 subunit beta
IL22	Interleukin 22
IL23	Interleukin 23
IL23R	Interleukin 23 receptor
IL27	Interleukin 27
KIF5A	Kinesin family member 5A
KIF21B	Kinesin family member 21B

LD	Linkage disequilibrium
LMP	Low molecular weight proteosome
LTBR	Lymphotoxin beta receptor
MALDI_TOF	Matrix assisted laser desorption/ionization time-of-flight
MHC	Major histocompatibility complex
miRNA	microRNA
MRI	Magnetic resonance imaging
mSASSS	Modified Stoke Ankylosing Spondylitis Spine Score
nAChR	Alpha 7 nicotinic acetylcholine receptor
NFkB	Nuclear factor kappa-light-chain-enhancer of activated B cells
NNAL	Glucuronidation of 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol
NNK	Nitrosamine, 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone
NOS2	Nitric oxide synthase 2
NSAIDS	Non-steroidal anti-inflammatory drugs
PCR	Polymerase chain reaction
PsA	Psoriatic arthritis
PTGER46	Prostaglandin E_2 (PGE_2) receptor EP4 subtype
ReA	Reactive Arthritis syndrome
RUNX3	Runt-related transcription factor 3
SH2B3	SH2B adapter protein 3
SNP	Single nucleotide polymorphism
SpA	Spondyloarthropathy
SPARCC	Spondyloarthritis Research Consortium of Canada
STAT3	Signal transducer and activator of transcription 3
TBKBP1	TBK binding protein 1
TCAG	The Center for Applied Genomics
Th2 cells	T helper 2 cells
Th17 cells	T helper 17 cells
TNF- α	Tumor necrosis factor alpha
TNFR1	Tumor necrosis factor receptor 1
TNFRSF1A	Tumor necrosis factor receptor superfamily member 1A
TYK2	Tyrosine kinase 2
UC	Ulcerative Colitis
UGT2B10	UDP glucuronosyltransferase 2 family, polypeptide B10
UGT2B17	UDP-glucuronosyltransferase 2 family, polypeptide B17
ZMIZ1	Zinc finger MIZ domain-containing protein 1
3HC	Trans-3′-hydroxycotinine

List of Appendices

Appendix 1: Additional Gene Variants Identified via GWAS 86

Appendix 2: Results Displaying Inheritance Models (Dominant, Recessive & Additive)… 93

1

Introduction

1.1 - Background of Ankylosing Spondylitis

Ankylosing Spondylitis (AS) is an inflammatory disease belonging to the disease group spondyloarthropathies (SpA), which primarily affects the axial skeleton and sacroiliac joints (Edavalath, 2010). Other disease entities that belong to the SpA group are psoriatic arthritis (PsA), enteropathic arthritis, inflammatory bowel disease (IBD) and Reactive Arthrtis (ReA) (Atzeni et al., 2014). AS occurs in approximately 0.1% to 1% of the population and onset of disease is generally in the early to late twenties. Men are two to three times more likely to be affected than women. Traditionally AS has been diagnosed by the modified New York criteria on the basis of its primary clinical and radiographic characteristics: low back pain and stiffness, limitation of lumbar spine, limitation of chest expansion and radiographic sacroiliitis (National Institute for Health and Clinical Excellence Technology Appraisal Guidance, 2008). With the most recent classification criteria, AS now falls into a broader spectrum of chronic back pain, which is now referred to as axial spondyloarthritis (Rudwaleit et al., 2009). For the purpose of this study, we investigate AS rather than axial SpA, as this represents a homogenous subset of patients with axial SpA. While the pain from inflammation is often dull and persistent, it has the potential to be severe and disabling, primarily due to a significant loss of range of movement in the axial spine.

Increased inflammation causes the ankylosis of individual vertebrae, leading to decreased spinal mobility and secondary diseases as syndesmophytes form in the joints between bordering vertebrae (Lian et al., 2013; Spondylitis Association of America, 2013). A common phenotypic expression of the disease is a bamboo spine, an exaggerated concave in the thoracic spine, which results from decreased spine mobility

due to spinal fusion (Boonen & van der Linden, 2006). With respect to spinal inflammation, individuals often experience other inflammatory musculoskeletal manifestations, such as enthesitis and dactylitis. Enthesitis occurs when the insertion point of tendons and ligaments to the bone become inflamed. Common sites of enthesitis include insertion of Achilles tendon, costochondral and costovertebral insertions (Salvarani & Fries, 2009). Dactylitis is characterized by diffuse swelling of the fingers and toes due to inflammation of the tendon sheath around the digits.

AS is associated with select extra-articular features. Common extra-articular features include acute anterior uveitis (AAU), psoriasis and IBD. AAU occurs in about 25 to 40% of AS patients and is characterized by unilateral eye pain, blurring of vision and photophobia. AAU is often recurrent and must be treated immediately or else it could lead to permanent ocular damage. Psoriasis occurs in up to 10% of AS patients and is characterized by hyperproliferative scaly eruptions, particularly in the scalp, nails and extensor surfaces of the joints. Clinical Crohn's disease or ulcerative colitis (UC) occurs in 10% of patients with AS, although sub-clinical evidence of bowel inflammation has been reported in up to 60% of AS patients (Salvarani & Fries, 2009). Recent studies have also reported an increased incidence of cardiovascular disease with AS (Haroon, Paterson, Li, Inman, & Haroon, 2015). Other extra-articular manifestations include apical pulmonary fibrosis, aortic insufficiency and osteopenia.

Radiographic progression is considered to be an important marker for damage in individuals with AS (Ramiro et al., 2013). It encompasses structural changes in the spine and sacroiliac joints. Measurement of the degree of severity of sacroiliitis in a radiograph is a requirement for the diagnosis of AS by the modified New York criteria (Poddubnyy

& Sieper, 2012). Radiographic progression is scored on the modified Stoke Ankylosing Spondylitis Spine Score (mSASSS; Poddubnyy & Sieper, 2012). The most substantial amount of radiographic progression is observed in the first 10 years after disease onset (Lian et al., 2013). mSASSS is the tool used for measuring radiographic progression, one of the core outcomes for AS. The interobserver reliability, truth, discrimination and feasibility of modified Stoke Ankylosing Spondylitis Spine Score (mSASSS) was assessed in AS over a 3 year period (Salaffi et al., 2007). This was compared to the Bath Ankylosing Spondylitis Radiology Index--spine (BASRI-spine). Interobserver agreement was analyzed by intraclass correlation coefficients (ICC). The BASRI-spine reached intra- and interobserver ICC of 0.755 and 0.831, respectively. The mSASSS scores were more reliable, with ICC of 0.874 and 0.941, respectively. With regards to sensitivity to change, it was found that mSASSS classified the highest percentage of patients with more changes than the BASRI-spine (mSASSS: 35.8% vs. BASRI-spine: 15.8%). Thus mSASSS offers advantages in measurement properties and is the most appropriate method by which to assess progression of structural damage in AS.

The condition is initially treated with anti-inflammatory medications, particularly non-steroidal anti-inflammatory drugs (NSAIDs), although some patients with AS are able to manage their condition without any pharmacological agents. For those AS patients receiving NSAIDs, about 70 to 80% report a significant benefit. There is no convincing evidence that the class of NSAID makes a significant difference, although the maximum dose of any particular NSAID is required. At present, there is controversy about the use of continuous versus on demand NSAID therapy. Earlier reports suggested that continuous

NSAIDs might delay radiographic progression; however, more recent studies have not reported this benefit (Sieper et al., 2015). Continuous NSAIDs are also associated with more adverse events, particularly gastroduodenal ulcers, ischemic heart disease and renal toxicity. Traditional disease-modifying antirheumatic drugs (DMARDs) are not used to treat axial symptoms, although selected DMARDs, such as sulfasalazine, may be used for peripheral arthritis. Biologic agents, particularly anti-tumor necrosis factor (TNF) agents, are increasingly being used in treatment regimes. Anti-TNF agents have demonstrated significant benefit in reducing pain and stiffness as well as in enhancing mobility and overall quality of life. These medications have also demonstrated reduction in spinal inflammation and possibly decreased radiographic progression with early and prolonged use. More recently, interleukin (IL)-17 inhibition has demonstrated efficacy in AS treatment (Baeten et al., 2015). Management throughout the course of AS can also include exercise, stretching and physical therapy (Spondylitis Association of America, 2013).

1.2 - Genetic Analysis Studies

The etiology of AS remains unknown, but it likely results from interplay between genetic, immunological and environmental factors. Multiple lines of evidence support a genetic basis of AS. Evidence for a genetic basis of AS is derived from family based investigations, association studies with human leukocyte antigens (HLA) and mounting evidence from genome-wide association studies (GWAS).

1.1.1 - Heritability of AS

Large epidemiological studies have estimated the sibling recurrence risk ratio in AS to be greater than 52, while the index of heritability for AS is greater than 90% (Brown et al., 1997; Brown, 2008; Brown, 2011). There is substantial evidence

through twin studies that have demonstrated that disease susceptibility is largely due to a genetic component (Brown et al., 1997; Reveille, 2011). Studies comparing monozygotic and dizygotic twins have confirmed that AS is a familial disease (Breban, 2006; Brown, 2008). AS twin studies have consistently reported significantly higher correlations of the disease between monozygotic twins compared with dizygotic twins, with coherence rates of 50 to 75% and 12.5 to 15%, respectively (Breban, 2006; Brown, Crane & Wordsworth, 2002; Jarvinen, 1995). Clinical manifestations of AS, in particular age of onset and disease outcomes (including disease activity and function), were also noted to be heritable.

1.1.2 - Genetic Studies

1.2.2a Linkage Studies

Linkage Mapping of genetic variants within families occurs when a gene at a particular locus segregates dependent on another gene at a neighboring locus on the same chromosome (Sims, Wordsworth & Brown, 2004). Linkage studies are particularly helpful in identifying genetic variants with large effect sizes when a complex disease segregates in a Mendelian inheritance.

To date, three large genome-wide linkage studies have been implemented in AS populations. All three studies suggested strong linkage in the major histocompatibility complex (MHC) chromosomal regions. Additionally, these studies implicate regions outside of the MHC genes. An assessment of an AS population in the United Kingdom, defined by the modified New York criteria, found significant linkage in chromosome 16q and suggestive linkage in chromosomes 2q, 9q, 10q and 19q (Laval et al., 2001). A study of North American sibling pairs found suggestive

linkage in chromosomes 6q and 11q (Zhang et al., 2004). The patients in this population were also defined by the modified New York criteria for AS. The third study, a French AS population analysis defined by the European Spondyloarthropathy Study Group, found suggestive linkage in chromosomes 5q, 9q, 13q and 17q (Groupe Français d'Etude Génétique des Spondylarthropathies, 2004). When the data from these three linkage studies were combined the results attest to suggestive linkage in chromosomes 10q and 16q and nominal linkage in chromosomes 1q, 3q, 5q, 6q, 9q, 17q and 19q (Chandran & Rahman, 2010). Despite the initial enthusiasm related to genome-wide linkage studies, very few robust candidate genes have been identified using this approach. Thus genome-wide linkage studies have largely been abandoned.

1.2.2b Genome-Wide Association Studies

GWAS are used for identifying the risk alleles of small effect sizes. These studies identify genetic variants that occur in high frequencies in affected versus unaffected individuals across the human genome (Patnala, Clements & Batra, 2013). Linkage Disequilibrium (LD) mapping has greatly facilitated the GWAS approach (Morota, Valente, Rosa, Weigel & Gianola, 2012). LD blocks are regions of chromosomes containing genetic loci that segregate and are inherited together.

Four GWA studies have been performed in AS populations to date (Brown, 2011; IGAS, 2013; Pointon et al., 2010a; Pointon et al., 2010b; TASC, 2010; TASC & WTCCC2, 2007). The first three GWAS identified 12 loci relating to the disease beyond the well-established association of *HLA-B*27* to AS disease susceptibility (Brown, 2011). A recent GWA study in June 2013, by the International Genetics of

Ankylosing Spondylitis (IGAS) Consortium, identified 13 new genetic loci associated with disease in AS patients. To date, there are 37 genes that have reached genome-wide significance in AS (Table 1-1). With the addition of this latest GWAS, 24.4% of the heritability of AS is genetically described. Of the 24.4%, 20.1% is explained by *HLA-B*27* and 4.3% is explained by the additional 25 loci (IGAS, 2013). Even with the identification of 25 new loci associations, there is still a large portion of AS heritability that is undetermined.

The identification of 13 new gene loci and the verification of 12 previously identified gene loci in the GWA study by IGAS (2013) gives preference to four biological pathways in the explanation of AS pathogenesis. First, among the loci identified are the IL-23 pathway genes, suggesting that the IL-23 pro-inflammatory pathway is involved in disease presentation in AS patients. In the second case, some loci identified are aminopeptidase genes suggesting that pathways involved in gut immunity are possibly linked to pathways involved in AS pathogenesis. Thirdly, pathways involved in T cell lymphocyte activation and differentiation are implicated. IGAS (2013) found an association between the single nucleotide polymorphisms (SNPs) implicated in AS and the cell counts of both $CD8^+$ and $CD4^+$ T cell lymphocytes. This further suggests that lymphocyte activation pathways play a role in AS disease pathogenesis. Lastly, the pathways involved in peptide signaling before HLA class I presentation are implicated (IGAS, 2013). These hypotheses of AS pathogenesis are explained in greater detail below in reference to the specific genetic loci to which they pertain.

1.3 - Review of Major AS Genes

1.3.1 - Immunogenetics

*1.3.1a - HLA-B*27*

Of individuals affected by AS, approximately 80% are *HLA-B*27* positive. Only 1 to 5% of *HLA-B*27* positive individuals have AS, leaving a large amount of the heritability of AS unexplained by this gene (Reveille, 2011; Reveille, 2012; van der Linden, Valkenburg, de Jongh & Cats, 1984). *HLA-B*27*, a MHC class I molecule, accounts for approximately 20.1% of the heritability of AS (IGAS, 2013). This gene is located on the short arm of chromosome 6 and codes for four major proteins that are associated with AS, each varying in portion among different ethnicities. The most abundant protein in Western European Caucasian individuals is *HLA-B*27:*05*,* the parent allele, which comprises approximately 90% of the variants within this ethnicity. The remaining 10% of *HLA-B*27* variations in Western European Caucasian individuals is within *HLA-B*27:02*. The two other primary variants, *HLA-B*27:07* and *HLA-B*27:04,* are found mainly in Southern Asia and Eastern Asia, respectively. Of the 69 identified subtypes of *HLA-B*27,* these four are found most often in populations. The remaining subtypes are considered rare and are generally identified only in families (Reveille, 2011). Three hypotheses explain the role of *HLA-B*27* in mediating vertebral joint inflammation in AS. The first hypothesis suggests that $CD8^+$ T cell receptors recognize*HLA-B*27*. A cytotoxic T cell autoimmune response is initiated by a receptor on $CD8^+$ T cells that identifies its arthrotigenic peptide, *HLA-B*27* (Reveille,

2011; Salvarani & Fries**,** 2009). The second hypothesis suggests that the heavy chain of *HLA-B*27* misfolds, or homodimerizes, while in the endoplasmic reticulum (ER) causing stress by activating an unfolded protein response. The Cys67 residue on the α1 domain of the heavy chains causes *HLA-B*27* to homodimerize, or misfold, which inhibits transcription and translation. The homodimerized domain directs BiP, the ER chaperone, eliciting downstream inflammation. The last hypothesis is that tissue-specific homing cells of the inflammatory response in the gut epithelium are inappropriately rerouted to synovial compartments, therefore moving from the gut into the joints (Fantini, Pallone & Monteleone, 2009; Reveille, 2011; Salvarani & Fries**,** 2009). Interestingly, this particular hypothesis supports mechanisms that guide genetic overlap between AS and IBD.

MHC genes other than *HLA-B*27* have been implicated in AS susceptibility. The most notable non-B*27 genes identified have been *HLA-B*40* and *HLA-B*60* (Reveille, 2014). Inconsistent associations have been noted for non-MHC genes, including MIC-A genes as well as TAP, low molecular weight proteosome (LMP) and TNF alleles.

1.3.2 - Gene Variants Identified via GWAS

The candidate genes in this study were selected on the basis of the following rationale. Firstly, genes reaching genome wide significance in multiple GWAS studies were preferred as we were wondering whether established genes is related to radiographic disease severity and function. Secondly, genes reaching statistical significance on previous candidate gene studies for radiographic severity were ideal for inspection. Lastly, high priority functional candidates that were related to new

bone formation, but not necessarily shown to be associated with AS, were chosen. We were suggesting that these genes are only relevant for disease expression and not susceptibility.

GWA studies have been instrumental in identifying multiple candidate genes that have reached genome-wide significance. These genes include *IL-23R, RUNX3, IL-1* region and *IL-1R2, ANTXR2, IL-12, CARD9, PTGER4, KIF21B, STAT3*, intergenic regions, TNF pathway associated genes, *ERAP1, TYK2, IL-6, IL-27, EOMES, ZMIZ1, IL7R, BACH2, SH2B3, GPR35, GPR37, GPR65, GPR25* and *HLA-A*0201* (Appendix 1; Brown, 2011; IGAS, 2013; Reveille, 2012). *IL-23R* and intergenic regions were associated with AS in both *HLA-B*27* positive and negative patients, confirming that the association of these genes was not the result of gene-gene interaction but simply a genetic component independently associated with the disease.

1.3.2a - IL-23R

The IL-17/IL-23R axis has been identified as having a significant role in AS, IBD and psoriasis susceptibility (De Vos, Hindryckx & Laukens, 2011; Robinson & Brown, 2012). One SNP, rs11209026, is common to all three diseases. *IL-23R* accounts for approximately 0.31% of the heritability of AS (Brown, 2011; Reveille, 2012). IL-23R regulates a downstream inflammatory response by initiating the differentiation of $CD4^+$ T cells into Th-17 helper (Th17) cells. The Th-17 cells produce IL-17, a pro-inflammatory cytokine, which stimulates fibroblasts, epithelial cells and macrophages to induce

inflammation through the production of pro-inflammatory mediators: IL-1, IL-6, TNF-α, NOS-2, metalloproteases and chemokines (De Vos et al., 2011; Iwakura & Ishigame, 2006; Matzkies et al., 2012; Reveille, 2012; TASC, 2010).

1.3.2b - IL-12B

IL12B, at chromosome 5q33, accounts for approximately 0.11% of AS heritability (Brown, 2011; Reveille, 2012). The protein encoded by this gene, IL12B40, has a p40 subunit and is shared by both the IL-12 and IL-23 pathways (Wong et al., 2012). The product binds *IL-23R* (Apel et al., 2013). The *IL-12* product, in addition to binding *IL- 23R,* plays a role in the Th-17 signaling pathway by overseeing $CD4^+$ T cell differentiation into Th-17 cells. The presence of Th-17 cells initiates cytokine production leading to the assemblage of pro-inflammatory factors. This pathway is involved in IBD and PsA pathogenesis; an association between the diseases and *IL-12B* has been reported (Jadon et al., 2013; Matzkies et al., 2012; Reveille, 2012; Ruutu et al., 2012). Again, the genetic overlap between the three inflammatory diseases, AS, IBD and PsA, suggests a contribution to disease susceptibility.

1.3.2c - ERAP1

The endoplasmic reticulum aminopeptidase 1 (*ERAP1)* gene encodes an aminopeptidase linked to the ER and trims peptides in preparation for binding MHC class I molecules. It is responsible for approximately 0.34% of the heritability of AS (Brown, 2012; Reveille, 2012). *ERAP1* was only found to be associated with AS in *HLA-B*27* positive patients (Reveille, 2011; Robinson & Brown, 2012*).*

1.3.2d - CARD9

The product of *CARD9* expression is caspase recruitment domain-containing protein 9 (hCARD9), which is responsible for signaling production of TNF, IL-6 and IL-23. hCARD9 uses NFκB to induce production of TNF, IL-6 and IL-23 while also differentiating T cells into Th-17 cells, which produce IL-17 and IL-23. The SNPs of *CARD9* were found to be relatively strongly associated with AS while still only accounting for 0.034% of the heritability of the disease (Brown, 2011; Reveille, 2012). *CARD9* was also associated with PsA and IBD. The high frequency of common genetic factors between the three inflammatory diseases supports the role of *CARD9* in AS susceptibility (Ruutu et al., 2012; TASC & WTCCC2, 2011).

1.3.2e - STAT3

Signal transducer and activator of transcription 3 (*STAT3*) was found to be associated with AS susceptibility. It codes for a molecule involved in the IL-17/IL-23 signaling pathway, which is an inflammatory response (Reveille, 2011; Reveille, 2012). STAT3 mediates the signal of IL-22 and IL-23 to the target cells, lymphoid cells and epithelial cells (Ciccia et al., 2012). The activated gene product is phosphorylated due to the binding of IL-23 to IL-23R (Di Meglio et al., 2011). The activation of the *STAT3* product elicits IL-23 expression, rendering an autoimmune inflammatory response. In individuals with AS and Crohn's disease, IL-22 and IL-23 expression was upregulated, resulting in the subsequent over-expression of *STAT3* in the ileum of the small intestine. *STAT3* has a genetic overlap with other autoimmune diseases such as Crohn's disease and type 1 diabetes (TASC & SPARCC, 2010).

1.3.2f - TYK2

Tyrosine kinase 2 (*TYK2*) is involved in intercellular signaling in the IL-23 pathway through signal transduction from IL-23R and other cytokine receptors. A common genetic variant (rs35164067) has been associated with AS and Crohn's disease. Meanwhile another variant in *TYK2* (rs6511701) is associated only with AS and not Crohn's disease (IGAS, 2013).

1.3.2g - RUNX3

RUNX3 is located on chromosome 1p36 and is responsible for about 0.12% of AS heritability (Brown, 2011; Reveille, 2012). It encodes a transcription factor involved in differentiation of $CD8^+$ T cells, which suggests that *RUNX3* may affect *HLA-B*27*. HLA class I molecules, such as *HLA*B27*, present antigens to $CD8^+$ T lymphocytes (Reveille, 2012). $CD8^+$ T cell molecules recognize the HLA-B*27 receptors eliciting an immune response with cytotoxic T cells. *RUNX3* has also been demonstrated to be involved in PsA susceptibility, with an effect size similar to that found in AS. Considering the known genetic overlap between the two inflammatory diseases, the additional observation on the co-familial genetic nature of the *RUNX3* gene in both PsA and AS supports the involvement of the gene in AS heritability (Apel et al., 2013).

1.3.2h - IL-6

IL-6 promotes naïve T cells to differentiate to Th-17 cells. The exact mechanism in which IL-6R exerts an effect is unclear. IL-6R functions simultaneously with transforming growth factor-β. This combination of factors also mediates the inhibition of transforming growth factor-β directed differentiation to T_{reg} cells (IGAS, 2013).

1.3.2i - IL-27

IL-27 affects the IL-23 pathway by promoting the differentiation of $CD4^+$ T cells.

This gene suppresses the differentiation of Th-2 and Th-17 cells (IGAS, 2013).

Table 1-1: Genes found to be significantly associated with AS through SNP-based GWA

studies

Nearby Gene(s)	**SNP**	**Odds Ratio**
IL-23R	rs11209026	1.90
LTBR-TNFRSF1A	rs11616188	1.38
2p15	rs10865331	1.36
ERAP1	rs30187	1.35
KIF21B	rs2297909	1.25
21q22	rs378108	1.25
TBKBP1	rs8070463	1.24
ANTXR2	rs4389526	1.21

PTGER4	rs10440635	1.20
RUNX3	rs11249215	1.19
IL-12B	rs6556416	1.18
CARD9	rs10781500	1.18
IL-1R2	rs2310173	1.18
IL-6R	rs4129267	1.14
FCGR2A	rs1801274	1.11
UBE2E3	rs12615545	1.12
GPR35	rs4676410	1.13
BACH2	rs17765610	1.15
ZMIZ1	rs1250550	1.11
NKX2-3	rs11190133	1.15
SH2B3	rs11065898	1.11
GPR65	rs11624293	1.20
IL-27-SULT1A1	imm_16_ 28525386	1.11
NOS2	rs2531875	1.12
TYK2	rs35164067	1.14
ICOSLG	rs7282490	1.11

1.3.3 - Missing Heritability

The addition of 25 more genetic associations with the disease adds little to the heritability of AS, where only an extra 4.3% of the genetic variance is explained. The total heritability of AS identified to date is 24.4%, suggesting that other genetic variations such as structural variations, epigenetics and rare variants along with complexities of genetic interactions (gene-gene interactions and gene-environment interactions), could explain some of the missing heritability (IGAS, 2013; Reveille, 2012).

1.3.3a - Copy Number Variations

Copy number variations (CNV) are structural variants in the genes that involve duplications or deletions of large regions of deoxyribonucleic acid (DNA). The previously discussed 24.4% of genetic heritability of AS has been explained through SNP-based GWAS. There is a large amount of unexplained genetic heritability that could be elucidated through an alternative method of genetic analysis. There is evidence to support CNV analysis in AS as an effective means for determining genetic heritability. In 2014, Jung et al. were the first to identify an AS-associated CNV variant in a Korean population using a SurePrint G3 Human CNV 180K microarray. The study found that 227 CNV regions were associated with AS. Of the candidate genes that were examined through the CNV GWA study and independent replications, 9 were successfully replicated in the analysis: 1q32.2 (*HHAT*), 1p34.2 (*BMP8A*), 2q31.2 (*PRKRA*), 6p21.32 (*HLA-DPB1*), 11q22.1 (*CNTN5*), 13q13.1 (*EEF1DP3*), 14q24.2 (*RGS6*), 16p13.3 and

22q11.1 (*IL17RA*; Table 1-2). The 5 deletion-type CNV regions, in 1q32.2, 2q31.2, 6p21.32, 13q13.1 and 16p13.3, were associated with an increased risk of AS. The other 4 CNV regions were protective. Of the 9 genes that were found to be significant in the first replication set, 4 CNV regions in 1q32.2, 2q31.2, 6p21.32 and 16p13.3 were consistently significantly replicated (Jung et al., 2014). This study was the first to explore CNV based GWAS to explain the pathogenesis of AS.

A recent study of a multiplex family identified that the UDP-glucuronosyltransferase 2 polypeptide B17 (*UGT2B17*) gene is present in copy number 2 in individuals affected by AS (Uddin, Sturge, Peddle, O'Rielly & Rahman, 2013). *UGT2B17* is a functional gene associated with androgen metabolism. This gene is responsible for glucoronidation of testosterone (Ekstrom et al., 2013). Given that glucuronidation is an important step in the metabolism of nicotine, the *UGT2B17* gene likely plays a role in the pathway of nicotine metabolism (Chen et al., 2010). With the identified association between AS and *UGT2B17* in a multiplex family, and given the role of *UGT2B17* in nicotine metabolism, we are postulating that *UGT2B17* may influence radiographic progression or disease severity in AS through a gene-environment interaction.

UGT2B17 is a gene that codes for an enzyme belonging to the UDP-glucuronosyltransferase family. It is a part of the UGT2B subfamily, which is primarily found in the hepatocytes and extrahepatic tissues involved in androgen metabolism (Turgeon, Carrier, Levesque, Hum & Belanger, 2010). It metabolizes steroid hormones by catalyzing the binding of the 17β-hydroxyl position of dihydrotestosterone, testosterone and 3α-diol to the gene product, which

glucuronidates androsterone (Pâquet et al., 2012). *UGT2B17,* therefore, has a role in the formation of estrogen, a hormone affecting bone structure and strength. The *UGT2B17* gene dosage, corresponding to CNV, affects the concentration of androgen metabolites produced and results in varying amounts of bone formation (Angstadt et al., 2013; Conrad et al., 2010). The CNV of *UGT2B17* has been found to have an influence on bone mineral density, cortical thickness and buckling ratio (Giroux, Bussieres, Bureau & Rousseau, 2012; Yang et al., 2008).

Enhanced expression of the gene affects bone formation, increasing the risk of osteoporosis. AS is characterized by the formation of new bone in the vertebral column; yet, osteoporosis is a common pathogenic characteristic of the disease, which increases the risk of spinal fracture (Yang et al., 2008). The role of the *UGT2B17* gene CNV in bone formation and thickness provides the basis for further investigation of CNV analysis in individuals with AS in this study.

Expression of *UGT2B17* and *UGT2B10* produces enzymes that catalyze the N- glucouronidation pathways of nicotine metabolism. Both genes, belonging to the UDP- glucuronosyltransferase family, are polymorphic. *UGT2B17*, along with *UGT2B10*, is involved in the metabolism of nicotine by initiating the glucuronidation of 4- (methylnitrosamino)-1-(3-pyridyl)-1-butanol (NNAL), a major metabolite from the nicotine-derived tobacco-specific nitrosamine, 4-(methylnitrosamino)-1-(3-pyridyl)-1- butanone (NNK). *UGT2B17* deletion significantly lowers the glucoronidation activity in smokers. The gene catalyzes the phase II detoxification process for nicotine to cotinine to trans-3'-hydroxycotinine (3HC; Chen et al., 2010; Chen, Giambrone & Lazarus, 2012). Radiographic progression in individuals with

AS has been demonstrated to occur faster in patients who smoke as compared to non-smokers. The previously found relationship between the *UGT2B17* gene and AS, in combination with the role of UGT enzymes in nicotine metabolism, warrants further investigation of *UGT2B17* CNV, smoking status and AS disease outcomes.

Table 1-2: Genes found to be significantly associated with AS through CNV-based GWAS

Gene	**Chromosome**
HHAT	1q32.2*
BMP8A	1p34.2
PRKRA	2q31.2*
BMP6	6p24.3
HLA-DPB1	6p21.32*
CNTN5	11q22.1
EEF1DP3	13q13.1
RGS6	14q24.2
--------	16p13.3*
IL17RA	22q11.1

1.3.3b - Gene-Gene Interactions

The remaining genetic variability may be attributed to multiple gene-gene interactions. For example, the gene-gene interaction between *HLA-B*27* and *ERAP1* contributes to disease progression in AS patients. *ERAP1* was only associated with the disease in individuals who were *HLA-B*27* positive, thus confirming that it is this interaction that elicits disease susceptibility and not the genetic contribution of *ERAP1* alone (TASC & WTCCC2, 2011; Reveille, 2011; Reveille, 2012). That *ERAP1* is associated with AS susceptibility only in the presence of *HLA-B*27* suggests that AS predisposition does not always stem from the presence of *HLA-B*27* alone. A malfunctioning *ERAP1* leads to misfolding of the *HLA-B*27* heavy

chain, which causes stress by activating an unfolded protein response that inhibits appropriate transcription and translation. When compared with controls, individuals with AS have overexpressed *ERAP1* within dendritic cells, resulting in loss of optimal regulation of the gene product (TASC & WTCCC2, 2011). Gene-gene interaction has also been noted between *ERAP1* and *HLA-B*40* in AS, *HLA-B*51* in Bechet's disease and *HLA-Cw6* in psoriasis. This suggests that the interaction of *ERAP-1* and *HLA* genes is very important in the pathogenesis of these disorders, as there are very few instances where gene-gene interaction has been noted for a complex disease (Brown, Kenna, & Wordsworth, 2015).

1.3.3c - Gene-Environment Interaction: Smoking and AS

Associations between environmental factors and disease outcomes have been observed in PsA patients. Smoking has been linked to worsening disease outcomes in PsA individuals, making smoking a prognostic factor for PsA (Duffin et al., 2009; Tillett et al., 2013). The *IL-13* region has been identified as being associated with increased susceptibility of PsA in psoriasis patients. Interestingly, smoking is protective in psoriasis patients, delaying the onset of PsA, especially in those patients with *IL-13* polymorphisms. AS and PsA are co-familial diseases, being included in the spondyloarthropathy family and overlap genetically (Dougados & Baeten, 2011). The genetic similarities between AS and PsA warrant investigation of the effect of smoking on AS disease outcomes. An interaction between the environment and genetic factors of AS may contribute to the missing genetic heritability.

AS disease activity and radiographic progression have been observed to be

significantly worsened and occur more rapidly in affected individuals who smoke (Kaan & Ferda, 2005; Mattey, Dawson, Healey & Packham, 2011; Poddubnyy et al., 2012, Ward et al., 2009). A gene-environment interaction exists between smoking and bone metabolism in osteoblast cells (Rothem, Rothem, Soudry, Dahan & Eliakim, 2009). Spinal bone formation is a characteristic feature in AS. Median disease scores on standardized disease functioning, pain and quality of life questionnaires for AS patients have been found to be significantly higher in individuals who smoked as compared to those who never smoked (Averns et al., 1996; Chung, Machado, van der Heijde, D'Agostino & Dougados, 2012). An additional dose-dependent relationship between pack years and disease worsening was also established (Mattey et al., 2011). Smoking has been found to be associated with spinal radiographic progression in individuals with early stage SpA and, furthermore, smoking status was positively associated with mSASS scores (Chung et al., 2012; Poddubnyy et al., 2012). Evidence suggests that smoking status alters the disease outcomes of AS; however, no association has been made between smoking and *HLA-B*27* positivity. Another genetic component may thus be the factor that interacts with the nicotine metabolism pathways that result in poor disease outcomes.

1.3.3d - Genetics of Disease Expression: Genetic Association With Extra-Articular Features

Overlap in genes and inflammatory pathways occurs in comorbid diseases such that individuals affected by AS exhibit a number of extra-articular features like IBD, psoriasis and uveitis (Boone & van der Linden, 2006; Rosenbaum & Chandran,

2012). There are a number of genetic irregularities associated with AS, IBD, psoriasis and uveitis. An example of such an overlap is seen in the upregulation of the IL-17/IL-23R pathway. This suggests that the remaining genetic heritability in AS may be co-familial (Brown, 2009; Reveille, 2012). That IBD, psoriasis and uveitis are common comorbidities associated with AS suggests that the four diseases share common genetic components, often causing them to occur in succession.

IBD is an inflammation of the gastrointestinal tract. The autoimmune disorder encompasses two major disease classifications: Crohn's disease and UC (Lees, Barrett, Parkes & Satsangi, 2011). IBD occurs in approximately 5 to 10% of all AS patients (Matzkies et al., 2012; Rosenbaum & Chandran, 2012). The two diseases share common inflammatory pathways that cause IBD to occur in combination with AS (Fantini, Pallone, & Monteleone, 2009). One of these pathways involves the Th-17 cell inflammatory response, which was observed to be a common pathogenic mechanism in both IBD and AS (Salvarani & Fries, 2009). The genetic component *IL-23R* plays a role in the signaling pathway involving Th-17 cells (Brown, 2009). The *IL-23R* gene is commonly mutated in protective and risk variants in both IBD and AS cases, adding evidence that inflammatory mechanisms for both disorders are similar (Lees et al., 2011). Genetic overlap between Crohn's disease and AS has also been identified using GWAS at loci *IL-23R*, *STAT3*, *IL-12B* and *TNFSF15* as well as the intergenic region at chr1q32. In addition to *IL-23R*, *STAT3* and *IL-12p40* play a role in the Th-17 inflammatory pathway (Matzkies et al., 2012).

Uveitis, or iritis, is an inflammation of the uvea, the middle layer of the eye. It causes pain, watering and redness in the eye, leading to blurred vision and sensitivity

to light (Spondylitis Association of America, 2013). AAU occurs in 40% of patients with AS (Reveille, 2011). The genetic overlap between AS and uveitis is primarily from the strong association of both diseases with the *HLA-B*27* gene. Uveitis is significantly associated with cervical spinal pain and *HLA-B*27* positivity (De Vos et al., 2011). It has been observed that over 90% of individuals with AAU due to *HLA-B*27* positivity also have SpA (Reveille, 2011). It has been suggested that the partnership between AS and uveitis is heavily influenced by environmental factors (Wendling et al., 2012).

Psoriasis is an autoimmune disorder in which thick white or red patches form on the skin of affected individuals. Psoriasis occurs in 20% of all individuals with SpA (De Vos et al., 2011). The primary genetic marker indicative of psoriasis susceptibility has been previously identified as *HLA-C*06:02*. This genetic component of psoriasis does not overlap with those genes identified as being associated with AS susceptibility; however, similar to AS, the primary genetic component of psoriasis appears to have an interactive relationship with *ERAP1*. *HLA-C*06:02* appears to be involved in a gene-gene interaction with *ERAP1* in affected individuals in a similar fashion to the interaction between *HLA- B*27* and *ERAP1* in individuals with AS. *HLA-C*06:02* and *HLA-B*27* are both peptides coded from genes found in MHC regions. The genetic similarities between AS and psoriasis in this instance further support misfolding of MHC peptides as a likely explanation of SpA susceptibility. Furthermore, other genetic overlap exists between AS and psoriasis beyond the *ERAP1* and MHC peptide relationship. Both diseases express *IL-23R* induction of Th-17 cells. The production of differentiated Th-17 cells results

in the assemblage of IL-17, an inflammatory response regulator (Iwakura & Ishigame, 2006; Reveille, 2012). The *IL-23R* gene may be a common mediator in the induction of inflammatory responses in AS, IBD and psoriasis.

1.3.3e - Rare Variants

The common variants identified via GWA studies marginally contribute to AS susceptibility. The common variant hypothesis dictates that multiple common variants of small effect size account for the heritability of common disease. GWA studies overlook rare variants as the SNPs chosen usually have minor allele frequency of over 1%. Additionally, linkage studies are grossly underpowered to identify genetic variants with low relative risk. With the advent of next generation sequencing, it is now possible to interrogate the genome for rare variants. These variants occur infrequently; however, when present, there is a large effect size. To date, there has only been one study performed to identify rare variants in AS patients (O'Rielly et al., 2015). Two familial deletions in *SEC16A* and *MAMDC4* correspond with increased susceptibility to axial Spondyloarthritis (AxSpA) in *HLA-B*27* positive patients belonging to a multigenerational family.

1.4 - Relevance of Research

The course of AS is extremely variable: some patients only have spinal involvement while others have peripheral arthritis, peri-articular disease (enthestis and dactylitis) or extra-articular manifestations (uveitis, psoriasis, IBD). The range of spinal ankylosis is also quite large with some individuals having long standing disease with no radiographic progression while others demonstrate rapid onset of

spinal ankylosis over just a few years. Since these patients are young at disease onset, the lifelong impact of the progression is significant. Early identification of extra-articular features or disease

progression, coupled with appropriate management, may improve the quality of life and prevent disability.

At present, when a patient with SpA presents to the physician, it is not clear what path their disease will take. It is possible to have a mild disease, which does not result in deformity, or a rapidly progressive course that results in considerable damage. Our proposed research will identify markers for disease expression, severity and progression. Clinicians may be able to utilize this work to determine the future course of the disease at onset. That, in turn, will allow individualized treatment such that patients destined to have mild disease do not get aggressive treatment exposing them to untoward side effects, while patients with a severe pattern will be treated early with appropriate medications to prevent progression of damage. Modification of the natural history of disease is possible if an environmental factor is shown to be associated with worsening disease outcomes.

1.4.1 - Study Objectives

The proposed research will help assess whether structural variations (i.e. CNVs) can predict disease expression and whether gene-environment interactions can predict radiographic severity and progression of AS. This study has two parts:

Part 1:

To assess the genetics of **disease expression** using recently identified CNVs for extra- articular manifestations of AS;

Part 2:

A. To assess the genetics of **radiographic severity** in AS using selected smoking related genetic variants; and
B. To assess the genetics of **radiographic progression** in AS using selected smoking related genetic variants.

2

Materials and Method

2.1 - Ethical/Administrative Approval

This project fell under the broader study of Genetics of AS and was approved by the Health Research Ethics Authority (Ref. #99.172). Ethics approval was renewed annually. Permission was given by the University Health Network of the Toronto Western Hospital and the University of Alberta for use of the clinical information and laboratory results of their AS patients. Each of these centers also has approval from their respective ethics committees. A material transfer agreement has been in place between these universities and Memorial University for over a decade. The clinical information was extracted from The Spondyloarthritis Research Consortium of Canada (SPARCC) database. In each center, informed consent was obtained from all patients used in the study prior to their information being recorded in the database.

2.2 - Study Design

This is a retrospective case study. Patient data was previously collected by the SPARCC consortium. For the purpose of this study, AS cases were stratified based on extra-articular manifestations, radiographic severity or radiographic progression.

2.3 - AS Patient Ascertainment

2.3.1 - Location

The patients used in this study were recruited from three large AS cohorts located in Newfoundland, Toronto and Alberta. All genetic and statistical analysis was performed at Memorial University.

2.3.2 - Diagnostic Criteria

All AS patients satisfied the modified New York criteria (Table 2-1). *One of three clinical criteria must be satisfied in combination with one radiological criteria for full diagnosis of AS*. The first clinical requirement is that the individual must have experienced low back pain and stiffness that improves with exercise but is not relieved by rest for more than three months. The second clinical qualification states that the individual has limitation of the lumbar spine in both the sagittal and frontal planes. The third criterion requires that the individual have limitation of chest expansion relative to normal values correlated for age and sex. The individual must present with sacroiliitis with a grade of at least 2 bilaterally or a grade of 3 or 4 unilaterally.

There now is a new classification criterion for AxSpA, which includes AS. Patients satisfy the criterion for AxSpA if they have active inflammation on magnetic resonance imaging (MRI), highly suggestive of sacroiliitis, or definite radiographic sacroiliitis according to the modified New York criteria. Furthermore, individuals must have one feature of AxSpA or *HLA-B*27* positivity and two features of AxSpA. Our study focuses on AS patients rather than AxSpA, as those with AxSpA may have no radiographic changes. Thus, it would have been difficult to follow radiographic progression using radiographs.

Table 2-1: 1984 Modified New York Criteria for diagnosis of Ankylosing Spondylitis. Patients must have one clinical criteria and one radiological criteria for diagnosis

Clinical Criteria – One of the following:
▪ Low back pain ≥ 3 months, improved by exercise and not relieved by rest
▪ Limitation of lumbar spine in sagittal and frontal planes (relative to normal values corrected for age and sex)
▪ Limitation of chest expansion (relative to normal values corrected for age and sex)
Radiological Criteria – One of the following:
▪ Bilateral grade 2-4 sacroiliitis
▪ Unilateral 3-4 sacroiliitis
Grading system for sacroillitis 1. Grade 0: normal 2. Grade I: some blurring of the joint margins - suspicious 3. Grade II: minimal sclerosis with some erosion 4. Grade III: definite sclerosis on both sides of joint or severe erosions with widening of joint space with or without ankylosis 5. Grade IV: complete ankylosis

2.4 - Data Collection

2.4.1 - Clinical Data

The clinical information used in this study was collected from the SPARCC database. The Spondylitis Consortium of Canada is a large national network that prospectively evaluates AS and PsA patients with a detailed, standardized protocol.

The data is captured prospectively as patients are seen in this clinic every 6 to 12 months after entry to the clinic. Three major sites, Newfoundland, Ontario and Alberta account for over 90% of the patients. This study used data that was collected from this consortium. As the events captured were already manifested from this consortium, it is felt that this is retrospective case only study, as all patients had ankylosing spondylitis. The case was then stratified based on presence or absence of extra-articular features. Only patients of Northern European Ancestry were collected in these cohorts in an attempt to reduce population stratification with candidate gene studies. Since all patients were self-reported to be Caucasians, ethnicity was not considered. The cohort consists of patients with AS referred to the clinic by family physicians and other medical specialists. All information is collected on a protocol, which includes clinical, laboratory and radiographic features updated on 12-month intervals. The radiologic assessments were performed at the University of Toronto on selected patients at 2- to 3-year intervals. There were two radiological assessments for a subset of our AS patients. The first radiological scoring that was available was referred to as mSASSS1. This score was used to assess disease severity. For patients that had a second mSASSS score available at a later point were referred to as mSASSS2. The mSASSS1 and mSASSS2 values were used to determine the rate of change of X-ray progression. Only 76 patients have serial radiographs as collection of the X-rays only began to answer a clinical question. After routine use of TNF inhibitors for the treatment of AS, a study of AS patients on biologics was compared to a historical control of about 100 patients demonstrated that these biologic agents may not be inhibiting the radiographic progression of this entity. In order to assess this

question, an investigator from the SPARCC network started doing serial spinal X-rays. This is not necessarily within the standard of care for most physicians. To the best of our knowledge the samples were consecutive samples that attend the U of T clinic, where the patients consented to repeat radiographs. It is from this cohort that the 76 patients were obtained. Patients were not selected for any specific trait, and also did not appear to have severe radiographic disease. This is because the median radiographic progression for this group is one mSASSS unit per year, which is very similar to what is reported in the literature.

With respect to the specific variables in our study, the patients completed questionnaires regarding their demographic information. Physicians collected histories at onset of AS and presence of any extra-articular manifestations, particularly IBD, psoriasis and uveitis at any point in their disease course. Comorbidities were recorded if individuals had a history of the disease, not if the disease was present at the time of examination.

In all three AS cohorts, smoking history was collected by physicians and coded on the basis of whether the individual had ever smoked or never smoked. The occiput to wall measurement was collected for the Newfoundland and Toronto cohorts while the University of Alberta cohort collected occiput to wall and tragus to wall measurements.

The occiput to wall value used in this report was collected upon the patient's most recent visit with the physician. The total number of enthesitis was collected for the patients from all three cohorts. The enthesitis value collected at the individual's first physician visit was used for this report. All information was previously recorded in the SPARCC database and was extracted for all locations commencing September 12, 2012.

2.4.2 - Genetic Data

Part 1:

2.4.2a - CNV Information: UGT2B17 Gene

The initial detection of a significant association between *UGT2B17* in copy number 2 and AS was in a large multiplex family. In summary, a custom microarray was used to genotype CNVs. Validation of the CNV estimation of the *UGT2B17* gene was performed using the Taqman copy number assay Hs03185327_cn (Life Technologies) using the manufacturer's recommended protocol.

The custom genome-wide microarray used to identify the *UGT2B17* CNV in affected individuals has been previously described (Uddin et al., 2013; Uddin et al., 2011). The microarray was covered with 2 X 1 million probes that were spaced with a mean of 280bp (Uddin et al., 2013). For this study, this unique assay was used to analyze CNV in 896 patient DNA samples. The samples were extracted from a comprehensive biobank located in the Health Sciences Center at Memorial University. The bank contains the genetic information of the SpA patients belonging to the cohorts from Newfoundland, Toronto and Alberta. DNA samples were pipetted into 96 well trays at 10 ng/µg and analyzed for CNV count at The Center for Applied Genomics (TCAG) in Toronto using the Taqman copy number assay Hs03185327_cn (Life Technologies).

Part 2:

2.4.2b - Selection of Related Smoking SNPs

The genetic information from the biobank was used to analyze related smoking SNPs. In total, 36 SNPs were selected. The DNA of individuals with available smoking history information was pipetted into 96 well trays at 10 ng/µg. The smoking SNPs were

then genotyped using the Sequenom MassARRAY system (iPLEX GOLD). Smoking SNPs were chosen from five genes that have been known to be associated with either AS or the nicotine metabolism pathway (Table 2-2).

Table 2-2: SNP IDs with associated genes

SNP ID	Gene
rs62003625	CHRNA7
rs4779565	CHRNA7
rs28531779	CHRNA7
rs261160	CHRNA7
rs6494223	CHRNA7
rs76336281	CHRNA7
rs12916879	CHRNA7
rs6494216	CHRNA7
rs12908877	CHRNA7
rs78383742	CHRNA7
rs75599783	CHRNA7
rs79184402	CHRNA7
rs885071	CHRNA7
rs13329490	CHRNA7
rs11636680	CHRNA7
rs11636680	CHRNA7
rs113629315	CHRNA7
rs1355920	CHRNA7
rs4779978	CHRNA7
rs10438340	CHRNA7
rs10438341	CHRNA7
rs2133965	CHRNA7
rs11852727	CHRNA7
rs7167551	CHRNA7
rs141124772	CHRNA7
rs7178176	CHRNA7
rs60109258	CHRNA7
rs202192566	CHRNA7
rs374853	CHRNA7
rs60056915	CHRNA7
rs115291806	CHRNA7
rs2337503	CHRNA7
rs111910242	CHRNA7
rs2337506	CHRNA7

rs115112505	CHRNA7
rs11071593	CHRNA7
rs188889623	CHRNA7
rs7172276	CHRNA7
rs12910885	CHRNA7
rs4349859	HLA-B*27
rs1800925	IL-13
rs61750900	UGT2B10
rs72551387	UGT2B17

First, in *HLA-B*27*, the gene most strongly associated with AS heritability, a single tag-SNP (rs4349859) was located and genotyped. We next identified a single tag-SNP (rs1800925) within the *IL-13* gene to investigate whether the same protective effect of *IL-13* polymorphisms in psoriasis progression is present in AS patients (Duffin et al., 2009; Eder, et al., 2010). For the purpose of this study, 36 SNPs located within the cholinergic receptor, nicotinic, alpha 7 (*CHRNA7)* gene, which is involved in receptor presentation in the nicotine pathway, were genotyped. A single tag-SNP located in every LD block within the *CHRNA7* gene was selected for genotyping (Rothem et al., 2009). Additionally, the SNP (rs61750900) in the Asp67Tyr allele located in the *UGT2B10* gene and the SNP (rs72551387) in the Asp85Tyr allele located in the *UGT2B17* gene were genotyped (Berg et al., 2010; Chen et al., 2010). Given their roles in the nicotine pathway, it is appropriate to investigate these genes for the effects of smoking on AS susceptibility and pathogenesis.

2.4.2c - Genotyping of Related Smoking SNPs

Genotyping of the relevant SNPs was done using the Sequenom MassARRAY system (Danoy et al., 2010). The system was operated as per the protocol stated by the manufacturer, Agena. The polymerase chain reaction (PCR) amplification used a unique

primer constructed by the MassARRAY custom design software. Two multiplex assays were designed for each AS sample. Originally, 43 SNPs were submitted for design, however only 36 SNPS had data relevant to the analysis. Each custom assay consisted of 36 SNPs and was constructed using Sequenom's Assay Design Suite 1.0. The products of the Sequenom reactions were dispensed onto a 96 element Spectrochip bioarray using the MassARRAY nanodispenser. The SNPs were assayed on the OpenARRAY genotyping platform. The products were detected and genotyped using iPLEX chemistry on a matrix assisted laser desorption/ionization time-of-flight (MALDI_TOF) mass spectrometer. MassARRAY typer software was then used to identify the genotypes present in the assays.

2.5 - Sample Size

The minimum sample size calculation for part 1 of the objectives was based upon previous research of the *UGT2B17* copy number in AS patients by Uddin et al. (2013) found in the Newfoundland and Alberta populations. Considering that the primary outcome of interest for this study is dichotomous and that the results are being assessed in a one-sided testing fashion, the results of the previous paper were used to estimate a larger sample size.

Based on the known frequency of the CNV from previous studies, a frequency of 40% was used for two or more copies of *UGT2B17*. With the assumption of a 10% difference in the allele frequency, a total of 200 patients were needed for an 80% power (i.e. $\beta = 0.20$; $\alpha = 0.05$). A breakdown of additional sample sizes is also presented (Table 2-3).

Table 2-3: Sample size calculations based on varying power indexes for Part 1

	N	
Index	**Total**	**Power**
1	100	0.535
2	200	0.839
3	300	0.954
4	400	0.988
5	500	0.997
6	600	>0.999
7	700	>0.999
8	800	>0.999

Power for the regression analysis performed to evaluate Part 2 of the objectives was calculated based on a 50% change of the mean value for either severity or progression. The calculation was completed assuming 38 patients per group. These assumptions considered will give 29.8% power to detect a difference of 4.35 in mSASSS1, which is 50% of the mean, and 20.6% power to detect a difference of 0.525 in progression, which is 50% of the mean.

2.6 - Statistical Analysis

SPSS Statistics v20 was used for data analysis.

2.6.1 - Data Organization and Coding

*2.6.1a - Variables: Extra-Articular Features, HLA-B*27 and UGT2B17*

The data from the three sites was initially entered and organized into a Microsoft

Excel spreadsheet after being extracted from the SPARCC database. The dichotomous and categorical variables were entered with a numerical code.

Dichotomous data were available for the following clinical variables: gender, *UGT2B17* CNV, IBD history, psoriasis history, uveitis history, *HLA-B*27* status and smoking history. Gender was coded as "1" being female and "2" being male. IBD, psoriasis and uveitis were analyzed as dichotomous variables by coding "0" as "no, never presented with the feature," and "1" as "yes, presently has the feature or once had the feature." For *HLA-B*27* status, positive was coded as "1" and negative was coded as "0." Smoking history was coded as "0" for "no, never smoked," and "1" as "yes, currently smoking or was a past smoker." For the primary analysis, AS individuals were divided into two groups based upon *UGT2B17* CNV. One group was "CNV 2" consisting of individuals with 2 or more copies of the *UGT2B17* gene while the second group was "CNV ≤ 1" consisting of individuals with CNV 0 or 1. CNV was coded as CNV 0 or 1 and CNV 2 because a previous study by Uddin et al 2013 found that CNV 2 was significantly associated with AS in a multiplex family. CNV 2 represents the wildtype gene and CNV 0 and 1 represent polymorphisms. Our analysis aims to determine whether the wildtype allele, that was significantly associated with AS in a multiplex family, is also associated with the extra-articular features of AS. We group CNVs in this way because we are using previous significant findings and expanding on them in a larger cohort to determine whether disease expression can be explained by the same allele frequency.

Physician Global Health Rating, written as DRGLOBAL, is an ordinal variable and was coded as "1" for "very good," "2" for "good," "3" for "fair," "4" for "poor" and

"5" for "very poor." Occiput to wall and age of onset of AS were measured on a continuous scale. The organized and coded data was then moved into SPSS Statistics v20 for analysis.

2.6.1b - Variables: Smoking SNPs and Attached Clinical Information

The results from the Sequenom MassARRAY system were organized into Microsoft Excel 2010 by SNP rs numbers. Genotype call information was divided into 36 SNP rs numbers. Each patient was matched to his or her corresponding genotype call and clinical information by SNP rs numbers. The organized and coded data was then moved into SPSS Statistics v20 for analysis.

2.6.2 - Variables: Data Filtering

Any AS patients with missing phenotype information for the variable under examination was removed from the analysis using a specially designed algorithm. Any smoking SNP with a single genotype call was removed from the analysis using the same algorithm. Also, any SNP failing the Hardy Weinberg equilibrium (HWE) was excluded from the analysis. Originally 43 SNPs were submitted for design; however, 7 SNPs produced no genotype calls due to poor multiplex design. This left 36 SNP's for analysis. In Sept. 2015, the rs 188889623 SNP was excluded from analysis due to poor performance (i.e., high UEP, low peaks). This brought the list down to 35 SNPs. From this list of 35, 11 SNPs failed to produce a genotype call for at least 1 sample, leaving 24 SNPs for subsequent analysis. From this list of 24 SNPs, 7

SNPs were uninformative meaning that all samples produced the same genotype. Therefore, 17 SNPs were included in the final analysis.

2.6.3 - Data Analysis

Comparisons involving the two *UGT2B17* CNV groups with another dichotomous variables were done using Chi- square analysis. A Chi-squared test is a statistical analysis that is used to evaluate the likelihood that an observed difference between two sets of categorical data occurred by chance. It tests a null hypothesis that no difference is observed between outcomes and each outcome is equally likely to occur. When the null hypothesis is true, the data follows a chi- squared distribution. Rejection of the null hypothesis indicates that the data sets are not independent. A Fisher's exact test is another statistical analysis that it used to evaluate contingency tables of categorical data. It also tests a null hypothesis that each outcome is equally likely to occur; however, a Fisher's exact test is used for data with small sample size and is only applicable to data when at least one cell of the contingency table is less than or equal to 5. For this reason we use the Chi-squared test in our analysis as all cells observed in our contingency tables are greater than 5.

Two-tailed tests were used instead of one-tailed tests in the assessment of whether CNV 2 was found to be more often associated with the extra-articular features of AS and *HLA- B*27*. It is possible that two copies of the *UGT2B17* gene may be more or less frequent in a subset of AS patients with extra-articular manifestations when compared to AS patients exhibiting no extra-articular features. Our null hypothesis states that the number of individuals with CNV of 2 is the same in both the patient group with extra-articular features or *HLA-B*27* positivity and the

group without extra-articular features or *HLA-B*27* positivity. Our alternative hypothesis is that the CNV groups are different between the two patient groups.

Associations between known smoking SNPs and disease severity and progression were evaluated using a multiple linear regression. The linear regression was controlled for age, gender and smoking habit.

2.6.3a -Part 1: UGT2B17 CNV and Extra-Articular Features

The associations of *UGT2B17* CNV with the extra-articular features of AS, IBD, psoriasis and uveitis were compared using the Chi-squared test. The comparison was done at a level of significance of $p \leq 0.05$.

*2.6.3b – Part 1: UGT2B17 CNV and HLA-B*27*

The gene-gene interaction of *UGT2B27* with *HLA-B*27* was tested using the Chi-squared test for association. This comparison was done at a level of significance of $p \leq 0.05$.

2.6.3c – Part 2: Disease Severity and Genotypes

A multiple linear regression was used to determine the association between each SNP and the mean values of mSASSS1. For the purposes of the analyses in the study, univariate analysis with each SNP was not performed. Instead, to evaluate the mSASSS data there were four variables and one interaction term in the regression models. The multiple linear regressions were controlled for age, gender and smoking habit. In this analysis three inheritance models were examined: additive, dominant and recessive. For each genotype, the HWE was tested, and only those SNPs which follow the HWE were included. An interaction term between smoking habit ("yes" versus "no") and each SNP

was introduced into the multiple linear regression models to examine the association among smokers and non-smokers, respectively. Dependent variables used in the multivariate regression analysis were both mSASSS1 and progression.

2.6.3d – Part 2: Disease Progression and Genotypes

A multiple linear regression was used to determine the association between each SNP and the mean values of disease progression. The multiple linear regressions were controlled for age, gender and smoking habit. In this analysis the same three inheritance models were examined. For each genotype, the HWE was tested, and only those SNPs which follow the HWE were included. An interaction term between smoking habit ("yes" versus "no") and each SNP was introduced into the multiple linear regression models to examine the association among smokers and non-smokers, respectively. A common element of genetic studies is to assess multiple genetic variants with a very small sample set. Multiple testing must be taken into account when performing candidate genetic testing. As long as the SNPs are not in linkage disequilibrium, the most straightforward way to account for multiple testing is the use the Bonferroni's correction. For larger sample sets such as a genome wide association study, the calculation of a false discovery rate can be used, as many SNPs are in linkage disequilibrium. Thus Bonferroni's correction is too conservative in this case. For each given SNP, they were assessed using additive and multiplicative models. Genetic interactions are often difficult to measure and interpret, as statistical interactions do not necessarily lead to biological interactions. Gene-gene interaction was assessed for *HLA-B*27* and *UGBT17,* as stated in the above section.

3

Results

3.1 -Part 1: *UGT2B17* Copy Number Variants and Extra-articular features in SpA

3.1.1 - Demographics

Of the 896 AS patients from Newfoundland, Toronto and Alberta, 555 (61.9%) were male while 341 (38.1%) of the patients were female. The patients were all analyzed for CNV of the *UGT2B17* gene. Of the 896 patients, 625 (69.8%) had available family history information. Of these, 448 (78.1%) did not have a family history of AS and 137 (21.9%) individuals had an affected family member. Of the 896 patients, 587 (65.5%) individuals had available smoking history information. Of these cases, 312 (52.3%) individuals had never smoked and 275 (46.8%) individuals had smoked, past or present (Table 3-1).

Additional frequency analysis was completed for *UGT2B17* alleles in each of the three populations analyzed. Of the 166 patients in the Newfoundland population 71 (43%) patients had CNV 2 and 95 (57%) had CNV 0 or 1. Of the 503 patients in the Toronto population, 183 (36%) had CNV 2 while 320 (64%) had CNV 0 or 1. Of the 227 patients in the Alberta population, 92 (41%) had CNV 2 and 135 (59%) had CNV 0 or 1 (Table 3-2)

Table 3-1: Demographic information for AS patients in a cohort of 896 patients

Clinical Information	Values
Male	61.9%
Family History of AS	21.9%
Smoking	46.8%

Table 3-2: CNV Frequency in the three populations analyzed

Population	CNV 2 Frequency	CNV ≤ 1 Frequency
Newfoundland	71 (43%)	95 (57%)
Toronto	183 (36%)	320 (64%)
Alberta	92 (41%)	135 (59%)

3.1.2 -*UGT2B17* CNV and AS Extra-Articular Features

Of the 896 patients identified by the *UGT2B17* CNV, 548 (61.2%) individuals expressed 0 or 1 copy of the gene and 348 (38.8%) individuals expressed 2 or more copies of the gene. The two CNV groups were then compared for frequencies of IBD, psoriasis and uveitis. Of the 896 patients identified by the *UGT2B17* CNV, 602 had available IBD history information. Of these patients, 486 individuals (80.7%) self- reported having never had symptoms of IBD and 116 patients (19.3%) reported a history of IBD symptoms. Between IBD and *UGT2B17* CNV, 308 patients had CNV 0 or 1 and never had IBD; 67 patients had CNV 0 or 1 and an IBD history; 178 patients had CNV 2 and never had IBD; and 49 patients had CNV 2 and an IBD history (Table 3-3; Figure 3- 1). Using the two-sided Chi-Squared test, no association was found between having

UGT2B17 CNV 2 and ever having IBD (p-value = 0.262).

Table 3-3: Contingency table of CNV categories and history of IBD used in the Chi-square test analysis

CNV CALLS	IBD	NO IBD
≤1	67 (0.58)	308 (0.63)
2	49 (0.42)	178 (0.37)

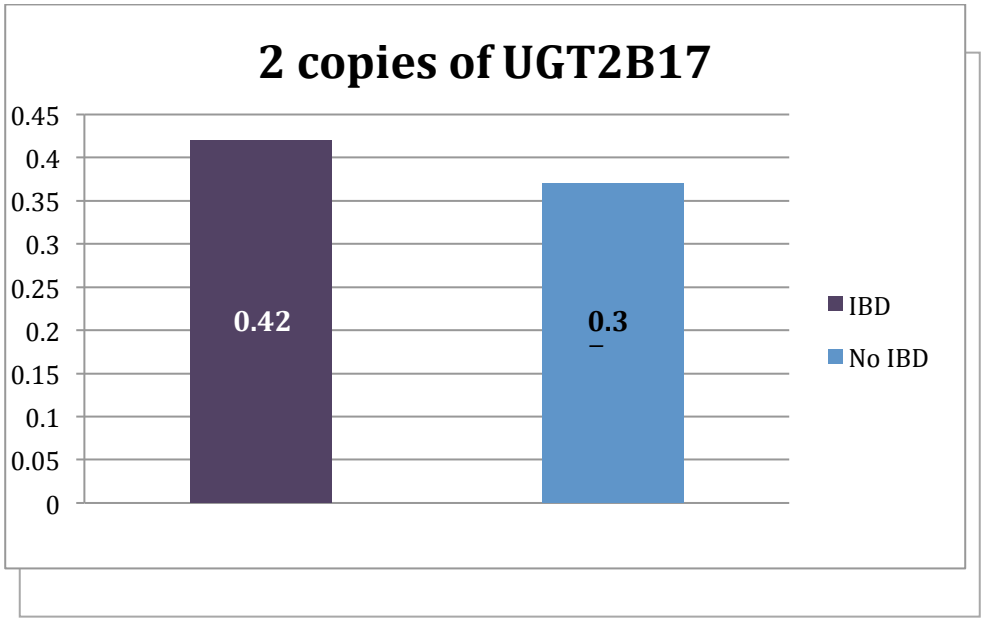

Fig 3-1: Proportions of individuals with 2 copies of the *UGT2B17* gene having self-reported IBD and never having self-reported IBD

Psoriasis history information was available for only 385 of the patients with CNV information. Of the total patients with psoriasis, 325 individuals (84.4%) reported never experiencing symptoms of psoriasis while 60 patients (15.6%) reported a history of psoriasis. Between psoriasis and *UGT2B17* CNV: 196 patients had CNV 0 or 1 and never

had psoriasis; 37 patients had CNV 0 or 1 and a psoriasis history; 129 patients had CNV 2 and never had psoriasis; and 23 patients had CNV 2 and a psoriasis history. Using the two-sided Chi-Squared test, no association was found between having *UGT2B17* CNV 2 and ever having psoriasis (p-value = 0.843; Table 3-4; Figure 3-2).

Table 3-4: Contingency table of CNV categories and history of psoriasis used in the Chi-square test analysis

CNV CALLS	PSORIASIS	NO PSORIASIS
≤1	37 (0.62)	196 (0.60)
2	23 (0.38)	129 (0.40)

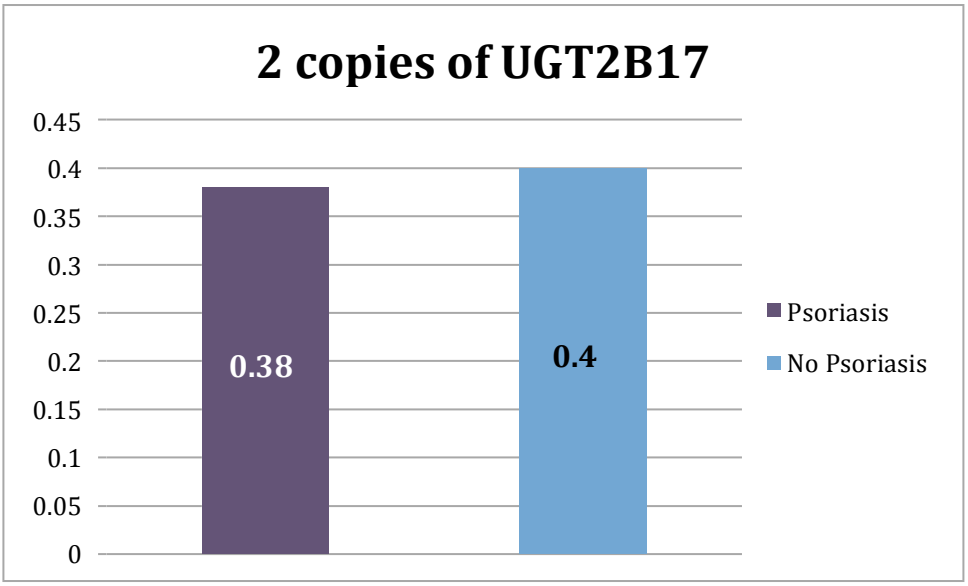

Fig 3-2: Proportions of individuals with 2 copies of the *UGT2B17* gene having self-reported psoriasis and never having self-reported psoriasis

Of the total patient group, 771 individuals had available uveitis information. Of these, 522 patients (67.7%) reported never having uveitis and 249 patients (32.3%)

reported a history of uveitis. Chi- squared test for association was preformed between CNV groups and each extra-articular feature separately. Between uveitis and *UGT2B17* CNV: 326 patients had CNV 0 or 1 and never had uveitis; 154 patients had CNV 0 or 1 and a uveitis history; 196 patients had CNV 2 and never had uveitis; and 95 patients had CNV 2 and a uveitis history. Using a two-sided Chi-squared test, no association was found between having *UGT2B17* CNV 2 and ever having uveitis (p-value = 0.871; Table 3-5; Figure 3-3).

Table 3-5: Contingency table of CNV categories and history of uveitis used in the Chi-square test analysis

CNV CALLS	UVEITIS	NO UVEITIS
≤1	154 (0.62)	326 (0.62)
2	95 (0.38)	196 (0.38)

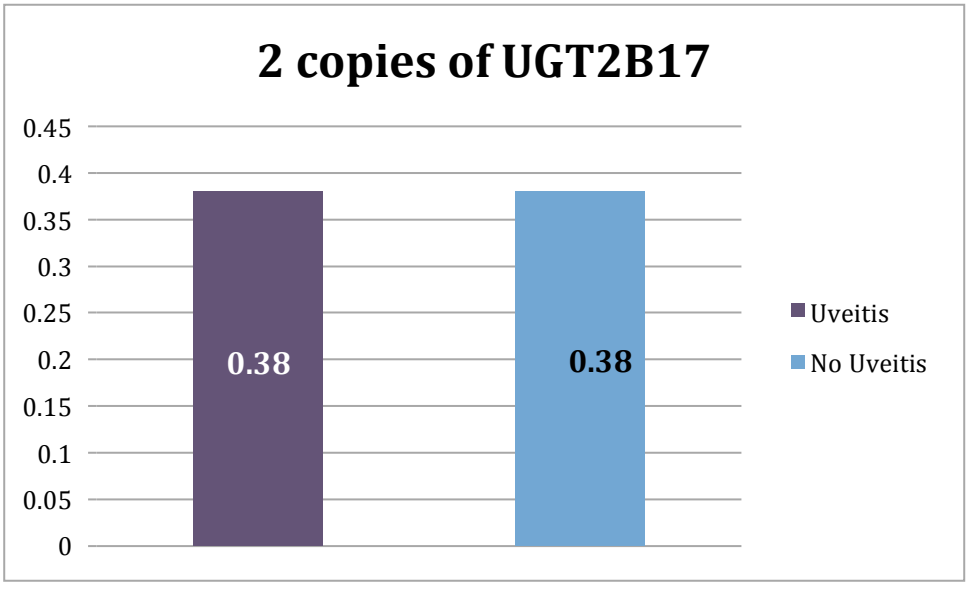

Fig 3-3: Proportions of individuals with 2 copies of the *UGT2B17* gene having self-reported uveitis and never having self-reported uveitis

3.1.3 - Gene-Gene Interaction: *UGT2B17* CNV and *HLA-B*27* Status

Of the 896 patients identified with the *UGT2B17* CNV, 754 of the patients had been tested for *HLA-B*27* positivity. Of these 754 patients, 601 patients (79.7%) tested positive for the *HLA-B*27* gene and 153 (20.3%) patients tested negative for the *HLA-B*27* gene. Between *UGT2B17* CNV status and *HLA-B*27* status: 102 individuals had CNV 0 or 1 and tested negative for *HLA-B*27*; 367 had CNV 0 or 1 and tested positive for *HLA-B*27*; 51 patients had CNV 2 and tested negative for *HLA-B*27*; and 234 patients had CNV 2 and tested positive for *HLA-B*27*. Using a two-sided Chi-squared test, no association was found between *UGT2B17* CNV 2 and *HLA-B*27* positivity (p-value = 0.202; Table 3-7; Figure 3-4).

Table 3-7 summarizes CNV frequencies in the four extra-articular feature groups.

Table 3-6: Contingency table of CNV categories and *HLA-B*27* status used in the Chi-square Fisher's exact test analysis

CNV CALLS	***HLA-B*27***	**NO *HLA-B*27***
≤1	367 (0.61)	102 (0.67)
2	234 (0.39)	51 (0.33)

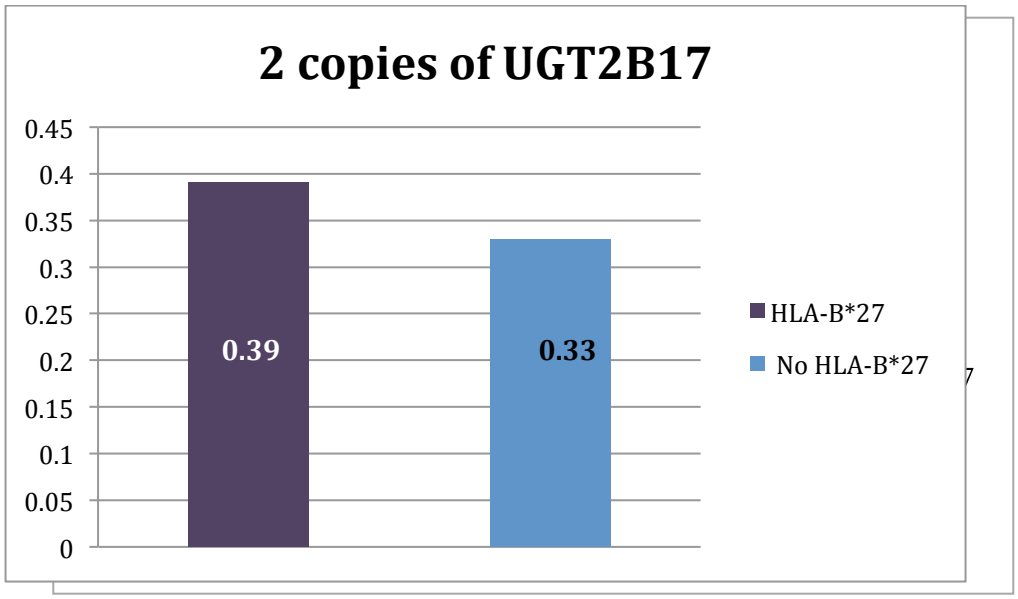

Fig 3-4: Proportions of individuals with 2 copies of the *UGT2B17* gene with a positive *HLA-B*27* status and a negative *HLA-B*27* status

Table 3-7: Frequency of individuals having ever presented with each extra-articular feature and *HLA-B*27* in each CNV category with associated p-values from statistical analysis using the Chi-squared test

Clinical Features	*UGT2B17* CNV 0 or 1	*UGT2B17* CNV 2	p-value
History of IBD (n = 602)	67 (17.9%)	49 (21.6%)	0.262
History of Uveitis (n = 771)	154 (32.1)	95 (32.6)	0.843
History of Psoriasis (n = 385)	37 (15.9%)	23 (15.1%)	0.871
*HLA-B*27* Positivity	367 (78.2%)	234 (82.1%)	0.202

3.2 -Part 2: Genetics of Radiographic Severity and Progression

3.2.1 - Demographics

Of the 75 patients used in the multiple linear regression analysis, 78% were male, 40% had a smoking history and 76% were *HLA-B*27* positive. The average onset of disease presentation in the 75 patients was 24.8 years; however, on average, the disease was not formally diagnosed until 31.2 years. An analysis was preformed to find an association with disease severity and disease progression with smoking SNPs. The radiographs used in the analysis were taken at an average of 3 years apart. Of the 75 patients, 35 patients demonstrated a change in serial radiograph values above 0 (i.e. radiographic progression). The mean disease severity (i.e. mSASSS1) for the entire 75 patients was 8.65, with a standard deviation of 13.22, and the mean mSASSS1 for the 35 patients with radiographic progression was 14.65, with a standard deviation of 15.28. The mean disease progression value for the entire 75 patient cohort was 0.51, with a standard deviation of 0.9, and the mean disease progression value for the 35 patients demonstrating disease progression was 1.09, with a standard deviation of 1.14 (Table 3-8).

Table 3-8: Demographics of AS patients in a cohort of 75 individuals with serial mSASS scores

Clinical Information	Values
Male	78%
Smoking	40%
*HLA-B*27* positive status	76%
Average age of disease onset	24.8 years

Average age of diagnosis	31.2 years
Average timeframe between serial radiographs	3 years
Number of patients with radiographic progression (change in mSASSS score > 0)	35
Mean disease severity (mSASSS1) for 75 patients	8.65 ± 13.22
Mean disease severity (mSASSS1) for 35 patients with progression	14.65 ± 15.28
Mean disease progression per year for entire 75 patients	0.51 ± 0.9
Mean disease progression per year for 35 patients with progression	1.09 ± 1.14

3.2.2 - Disease Severity and Smoking SNPs

Analysis of the interaction between selected genetic variants, radiographic severity and both smoking status and sex was performed. Three inheritance models are described as dominant (CC or CT), recessive (TT), and additive (Appendix 2). Additive represents an allele dose effect by comparing 2 C alleles (CC), 1 C allele (CT) and 0 C alleles (TT). Four tables (Table 3-9, Table 3-10, Table 3-11 and Table 3-12) display the results of the analyses. Two SNPs were found to be significantly associated with disease severity, rs7178176 and rs11071593, both with a significant of $p = 0.0005$ for the dominant model and $p = 0.001$ for the genotype. Table 3-12 displays only SNPs with statistical significance. Both of these SNPs are located in the deep intronic region of *CHRNA7* and play a role in the smoking and nicotine dependence pathway.

Table 3-9: Disease severity p-values in association with SNPs for genotypes

SNPs for Genotype	Disease Severity (mSASSS1)
rs885071	0.6801
rs1800925	0.2936
rs2337506	0.2068
rs4779565	0.8455
rs4779978	0.441
rs6494223	0.5751
rs7167551	0.8284
rs7178176	0.001
rs11071593	0.0087
rs11636680	0.8427
rs12910885	0.2068
rs12916879	0.8982
rs13329490	0.8343
rs60109258	0.4716
rs61750900	0.2351
rs62003625	0.4859
rs113629315	0.2834

Table 3-10: Disease severity and interaction with smoking p- values in association with SNPs for genotype

SNPs for Genotype	Disease Severity (mSASSS1) and Interaction with Smoking
rs885071*smoking	0.7068
rs1800925*smoking	0.7014
rs2337506*smoking	0.686
rs4779565*smoking	0.6555
rs4779978*smoking	0.0127
rs6494223*smoking	0.7401
rs7167551*smoking	0.9237
rs7178176*smoking	0.5443
rs11071593*smoking	0.8752
rs11636680*smoking	0.592
rs12910885*smoking	0.686
rs12916879*smoking	0.2175
rs13329490*smoking	0.2391

rs60109258*smoking	0.3761
rs61750900*smoking	0.0435
rs62003625*smoking	0.5468
rs113629315*smoking	0.6482

* Indicates a variable interaction term

Table 3-11: Disease severity and interaction with gender p-values in association with SNPs for genotype

SNPs for Genotype	Disease Severity (mSASSS1) and Interaction with Gender
rs885071*gender	0.893
rs1800925*gender	0.2991
rs2337506*gender	0.8864
rs4779565*gender	0.9838
rs4779978*gender	0.5371
rs6494223*gender	0.7823
rs7167551*gender	NA
rs7178176*gender	0.4092
rs11071593*gender	0.9793
rs11636680*gender	0.9437
rs12910885*gender	0.8864
rs12916879*gender	0.2319
rs13329490*gender	0.4007
rs60109258*gender	0.7092
rs61750900*gender	0.6313
rs62003625*gender	0.9363
rs113629315*gender	0.6234

*Indicates a variable interaction term

Table 3-12: SNPs from the *CHRNA7* gene having significant association with disease severity

SNPs located in *CHRNA7*	Disease Severity (mSASSS1)
rs7178176	Dominant model p = 0.0005 Genotype p = 0.001
rs11071593	Dominant model p = 0.0005 Genotype p = 0.001

3.2.3 - Disease Progression and Smoking SNPs

An analysis of the interaction between selected genetic variants, radiographic progression (change in mSASSS > 0) and both smoking status and sex was performed. Three inheritance models are described as dominant (CC or CT), recessive (TT), and additive (Appendix 2). Additive represents an allele dose effect by comparing 2 C alleles (CC), 1 C allele (CT) and 0 C allele (TT). Four tables (Table 3-13, Table 3-14, Table 3- 15 and Table 3-16) display the results of the analysis. Two SNPs were found to be significantly associated with disease progression, rs2337506 and rs12910885. The SNP rs2337506 was associated with disease progression with a significance of $p < 0.001$ for the recessive model and $p = 0.003$ for the genotype. The SNP rs12910885 was associated with disease progression with a significance of $p < 0.0001$ for the recessive model and $p = 0.0001$ for the genotype. Table 3-16 displays only SNPs with statistical significance. All SNPs were run in a likelihood ratio test with radiographic progression. Each SNPs was run separately in its own model with radiographic progression. The two SNPs listed in Table 3-16 were the SNPs that were significant.

Both of these SNPs are located in the deep intronic region of *CHRNA7* and play a role in the smoking and nicotine dependence pathway. The final analysis evaluates the genetics of radiographic progression in AS using selected smoking related genetic variants.

Table 3-13: Disease progression p-values in association with SNPs for genotypes

SNPs for Genotype	Disease Progression
rs885071	0.071
rs1800925	0.1384
rs2337506	0.0003
rs4779565	0.7416
rs4779978	0.1038
rs6494223	0.2132
rs7167551	0.3206
rs7178176	0.1913
rs11071593	0.1155
rs11636680	0.0771
rs12910885	0.0003
rs12916879	0.0648
rs13329490	0.6645
rs60109258	0.2028
rs61750900	0.5182
rs62003625	0.6831
rs113629315	0.7398

Table 3-14: Disease progression and interaction with smoking p-values in association with SNPs for genotype

SNPs for Genotype	Disease Progression and Interaction with Smoking
rs885071*smoking	0.5936
rs1800925*smoking	0.1493
rs2337506*smoking	0.5652
rs4779565*smoking	0.9462
rs4779978*smoking	0.3521
rs6494223*smoking	0.8224
rs7167551*smoking	0.4022
rs7178176*smoking	0.3499
rs11071593*smoking	0.1144

rs11636680*smoking	0.184
rs12910885*smoking	0.5652
rs12916879*smoking	0.2824
rs13329490*smoking	0.5582
rs60109258*smoking	0.6555
rs61750900*smoking	0.662
rs62003625*smoking	0.7252
rs113629315*smoking	0.1655

*Indicates a variable interaction term

Table 3-15: Disease progression and interaction with gender p-values in association with SNPs for genotype

SNPs for Genotype	Disease Progression and Interaction with gender
rs885071*gender	0.3719
rs1800925*gender	0.9572
rs2337506*gender	0.2218
rs4779565*gender	0.9991
rs4779978*gender	0.7992
rs6494223*gender	0.9382
rs7167551*gender	NA
rs7178176*gender	0.33
rs11071593*gender	0.8648
rs11636680*gender	0.5848
rs12910885*gender	0.2218
rs12916879*gender	0.8
rs13329490*gender	0.9888
rs60109258*gender	0.9042
rs61750900*gender	0.7772
rs62003625*gender	0.7536
rs113629315*gender	0.8443

*Indicates a variable interaction term

Table 3-16: SNPs from the *CHRNA7* gene having significant association with disease progression

SNPs located in CHRNA7	Disease Progression
rs2337506	Recessive model $p < 0.001$ Genotype $p = 0.0003$
rs12910885	Recessive model $p < 0.0001$ Genotype $p = 0.0001$

Table 3-17 displays the significant results of comparing progression values between non- smokers and smokers with the addition of C allele in the genotype for the rs1800925 SNP. The addition of a C allele in the genotype was not associated with progression in non-smokers. The addition of a C allele in the genotype was significantly associated with progression in smoking patients. The R squared value in this multivariate regression analysis of progression with rs1800925 was 0.22.

Table 3-17: Comparison of progression values between non-smokers and smokers with the addition of C allele in the genotype for the rs1800925 SNP

Smoking Status	B-value of progression	P-value
Non-Smoker	0.64	0.2348
Smoker	-1.34	0.0497

Table 3-18 shows significant comparisons of mSASSS1 values between non-smokers and smokers with GG, GT, and TT genotypes for the rs61750900 SNP. There was no significant association between the SNP rs61750900 and mSASSS1 in non-

smokers. Smokers with GT genotype had significantly lower mSASSS1 than those with GG. The R squared value in this multivariate regression analysis of mSASSS1 with rs61750900 was 0.28.

The SNP rs61750900 is a regulatory region variant found in the *UGT2B10* gene, which produces one of three enzymes that primarily metabolize nicotine. *UGT2B10* is the primary, if not the only, catalyst of nicotine and cotinine N-glucuronidation (Berg et al., 2010; Chen et al., 2007). The SNP rs61750900 in particular codes the Asp67Tyr UT2B10 variant enzyme, a missense mutation. This enzyme significantly reduces nicotine and cotinine glucuronidation while also potentially influencing cigarette consumption (Berg et al., 2010; Chen et al., 2010).

Table 3-18: Comparison of mSASSS1 values between non-smokers and smokers with GG, GT, and TT genotypes for the rs61750900 SNP

Smoking Status	Genotype	Patient Count	mSASSS1	P-value
Non-smoker	GG	40	0	–––
	GT	3	3.96	0.557
	TT	1	1.41	0.903
Smoker	GG	26	0	–––
	GT	6	-13.23	0.0084
	TT	0	–––	–––

4

Discussion

4.1 - Discussion

AS is an inflammatory immune mediated disease that is characterized by neo-ossification of the sacroiliac joint and axial spine. Although the etiology of AS is not yet fully elucidated, there is substantive evidence for the role of genetic, immunological and environmental factors in AS pathogenesis. AS exhibits the strongest genetic contribution of any complex rheumatic disease. It is associated with the highest relative sibling risk (λ= 50-58) and the strongest genotype relative risk (*HLA-B*27*). Over the last five years, there has been a plethora of genes identified in AS arising from GWA studies. Consequently, GWA studies have greatly enhanced our understanding of AS pathogenesis and have identified SNPs within genes that suggest specific immunomodulatory pathways affecting innate and acquired immunity. One such genetic association is the IL-23/IL-17 pathway, which contributes to AS pathogenesis. Despite the large number of genes identified, only one quarter of the genetic contribution to AS has been explained. Various explanations are postulated to explain the missing heritability: (1) insufficient samples being investigated; (2) rare genetic variants implicated in the pathogenesis, especially for familial disease; (3) presence of gene/gene or gene/environment interactions; (4) genetic variants other than SNPs contributing significantly to the genetics of AS. Other genetic variants that are presently being investigated in AS include CNVs, epigenetics, and gene expression profiling, including microRNA (miRNA).

The first research question of this thesis was formulated following a genome-wide microarray study of a large multiplex AS family which revealed association of a *UGT2B17* CNV in affected family members. The association of two copies of *UGT2B17*

with AS may play a significant role in disease establishment. The deletion-type variant of *UGT2B17* has previously been associated with osteoporosis (Yang et al., 2008). In addition to inflammatory axial pain, several family members exhibited psoriasis, Crohn's disease and anterior uveitis. The *UGT2B17* variant association identified in the multiplex family was not associated with AS in a larger independent cohort of unrelated AS cases; however, the possibility of a genetic association between *UGT2B17* and extra-articular features was not examined. This gene is thought to encode a key enzyme in the inhibition of androgen activity, making it functionally relevant in AS pathogenesis (Uddin et al., 2013). Given that the functional consequences of a CNV alteration may be greater than many single-base pair sequence changes and that CNVs have the potential to significantly alter gene expression, a systematic evaluation is warranted.

In this thesis, a concerted effort was undertaken to interrogate the genetics of disease expression rather than disease susceptibility. Much of the effort to date has been focused on genetic susceptibility. The genetics of disease susceptibility can often be challenging, as numerous genes interacting with unknown environmental factors are likely to increase the risk of developing disease. Therefore, it may be more strategic to investigate a specific feature of AS, particularly if this is a heritable disease manifestation. This endophenotype-based approach has the potential to enhance the genetic dissection of complex diseases. Endophenotypes may result from a subset of the genes necessary to develop AS because only a portion of the disease is being evaluated. An endophenotype, by definition, should be heritable and ideally not associated with disease activity. In other words, it is primarily state-independent and co-segregates with disease within families. The endophenotype should be found in higher frequency in first

and second degree relatives that do not have AS, as compared with the general population (Lee, Iafrate & Brothman, 2007). All of these attributes exist for psoriasis, IBD and uveitis, which are extra-articular features of AS. As noted above, about 30 to 40% of AS patients will develop anterior uveitis, 20% will develop psoriasis and 5 to 10% will develop inflammatory colitis. GWA studies have demonstrated a link between AS and each of these diseases, suggesting that there is likely an underlying factor in AS that predisposes the patient to develop related autoimmune diseases.

CNVs are reported to occur in 12% of the human genome and to contribute 0.12% to 7.3% of the genomic variability observed in humans (Scherer et al., 2007). It is intriguing that over 41% of all CNVs identified overlap with known genes. By extension, CNVs may play a prominent role in modulating genetic expression (Alkan, Coe & Eichler, 2011; Jung et al., 2014; Tsui et al., 2014; Uddin, et al., 2014; Wang et al., 2015). CNVs contribute to phenotypic expression by several molecular mechanisms. One of the most commonly recognized mechanisms for altering phenotype is attributed to dosage-sensitive genes, known as the "dosage effect" (Lupski et al., 1992). CNVs also have a "position effect," wherein gene expression can be altered depending on CNV chromosomal location, which can often be translocated (Kleinjan & van Hayningen, 2005). Another molecular mechanism by which rearrangements of the genome may alter disease phenotype is the unmasking of recessive mutations or functional polymorphisms of the remaining allele when a deletion occurs (Kurotaki et al., 2005).

GWA studies have been successful on several occasions in exploring AS associated CNVs. One study, in particular, by Jung et al. reported a significant association between 8 CNVs regions (i.e., 1p34.2, 1q32.2, 2q31.2, 6p21.32, 11q22.1, 13q13.1,

16p13.3, and 22q11.1) and risk of AS. This association remained significant even after accommodating for the contribution to AS risk from the well known AS susceptibility gene, *HLA-B*27* (Jung et al., 2014). The 8 CNV regions identified in that study were near genes that have already been established in AS pathogenesis (e.g., *HHAT, IL-17RA, BMP6*, and *BMP8A*). There was a dose-dependent relationship observed in CNVs of these genetic components and an effect on AS presentation. Additional copy numbers in *IL-17RA* and reduction in the copy number in *BMP8A* were significantly associated with diminished disease expression in individuals. CNV deletion in the *HHAT* gene has been demonstrated in the promotion of osteoblastic activity in arthritis leading to joint remodelling. Furthermore, a CNV dose-dependent relationship was demonstrated in the *FCGR3A* and *FCGR3B* genes, specifically suggesting that copy number of 3 or less was significantly associated with AS disease presentation (Wang et al., 2015).

A relatively large cohort of AS patients (n = 896) was investigated to determine whether the *UGT2B17* CNV was associated with the extra-articular features of AS disease expression. We also set out to determine if gene-gene interactions exist between *UGT2B17* and *HLA-B*27. UGT2B17* in copy number 2 was not associated with any of the extra-articular features, whether it be uveitis, psoriasis or IBD. No gene-gene interaction was noted between *UGT2B17* and *HLA-B*27.* There are multiple reasons why an association was not detected. First, it is entirely conceivable and likely that a true association does not exist between *UGT2B17* and extra-articular manifestations. Although our study was not powered to confidently rule out such an association, this represents the most likely scenario. Likewise, the sample size is too small to appreciate any contribution that *UGT2B17* CNV had on expression of extra-articular features or interaction with the

MHC gene in disease expression, despite a true association existing. A larger cohort of individuals affected with AS may show significant association of the gene with these features. Secondly, the *UGT2B17* gene in copy number 2 was associated with AS susceptibility in a multiplex family, but no association appears in the general population. Therefore, *UGT2B17* may be a rare or intermediate variant contributing to the disease only in rare multiplex families and not in patients that predominantly have sporadic disease. It is also conceivable that *UGT2B17* is a population-specific variant in the Newfoundland population, but not in outbred Caucasian populations. The Newfoundland population is significantly different from populations in other areas of Canada and the world due to a founder effect that has resulted in increased signal-to-noise ratio as well as greater number and length of runs of homozygosity. The association of this gene with AS and extra-articular features may thus be more applicable to the Newfoundland population. The second research question of this thesis was formulated following the observation that there appeared to be great variability in spinal damage in AS, both in severity and disease progression. Radiographic disease severity and progression of the axial spine is not linear. Even though the mean radiographic progression was 1.3 mSASSS, some patients exhibit a much higher rate of progression. Clinical factors that have previously been associated with rapid radiographic progression include baseline radiographic damage, elevated acute phase reactants and cigarette smoking (Poddubnyy et al., 2012). Smoking has been associated with worse radiographic outcomes in AxSpA – a relationship that may be dose-dependent. Smoking has been associated with worse inflammation and radiographic damage in early AxSpA. Worsening radiographic severity even in longstanding disease has also been observed (Chung, Machado, van der Heijde,

D'Agostino & Dougados, 2012). As a result, we set out to investigate whether there is an association between genetic variants related to smoking metabolism and radiographic progression in AxSpA.

A multiple linear regression was used to determine the association between each SNP and the mean values of mSASSS. The multiple linear regression models were controlled for age, gender and smoking habit. In this analysis three inheritance models were examined: additive, dominant and recessive. For each genotype, the HWE was tested, and only those SNPs that follow the HWE were included. An interaction term between smoking habit ("yes" versus "no") and each SNP was introduced into the multiple linear regression models to examine the association among smokers and non- smokers, respectively. AS cohorts with repeated radiographic assessments that also have DNA collected is quite rare. From our knowledge to date, the U of T AS cohort has the largest collection of patients with repeated radiographic measurements. Our plan is to replicate our study, once another independent AS cohort has been established at University of Toronto. Although collection for additional samples have started, this will take some time, as the median intervals between radiographics is 3 years. Thus, our study was not validated as part of this thesis.

Disease severity scores (mSASSS) were significantly associated with two SNPs from the *CHRNA7* gene: rs7178176 (dominant model p-value = 0.0005; genotype p-value = 0.001) and rs11071593 (dominant model p-value = 0.0005; genotype p-value = 0.001). Radiographic progression was determined to be associated with two distinctly different SNPs from *CHRNA7*: rs2337506 (recessive model p-value < 0.001; genotype p-value =

0.0003) and rs12910885 (recessive model p-value < 0.0001; genotype p-value = 0.0001). When smoking was introduced as an interaction term, there was no association with *CHRNA7* SNPs with disease severity or progression; however, there was a significant association between the SNP rs61750900 and mSASSS in smokers. Among smokers, those with GT genotype in this SNP had significantly lower mSASSS1 than those with GG. The SNP rs61750900 is a regulatory region variant found in the *UGT2B10* gene, which produces one of three enzymes that primarily metabolize nicotine. *UGT2B10* is the primary, if not the only, catalyst of nicotine and cotinine N-glucuronidation (Berg et al., 2010; Chen et al., 2007). In particular, it codes the Asp67Tyr UT2B10 variant enzyme, a missense mutation. This enzyme significantly reduces nicotine and cotinine glucuronidation while also potentially influencing cigarette consumption (Berg et al., 2010; Chen et al., 2010). Thus, we are not able to postulate that it is smoking that led to the worsening radiographic outcomes, as it is conceivable that there is a genetic association between *CHRNA7* and radiographic severity and progression. *CHRNA7* transcribes the alpha 7 nicotinic acetylcholine receptor (nAChR), which mediates immune system responses through its interaction with acetylcholine (ACh; Jonge & Ulloa, 2007). ACh activates the cholinergic anti-inflammatory pathway of the innate immune system and decreases production of endotoxin-inducible pro-inflammatory cytokines such as TNF, IL-1β, IL-6 and IL-18 (Borovikova et al., 2000). *CHRNA7* produces the receptor that mediates the anti-inflammatory effects of ACh. Given that nicotine is another ligand that binds to nAChR, nicotine agonists inhibit the production of cytokines and thus inhibit a pro-inflammatory response.

CHRNA7 plays a pivotal role in the mediation of inflammatory pathways,

especially those producing TNF, a prominent cytokine implicated in AS pathogenesis. The nAChR has been shown to be represented on the surface of macrophages and to play a role in efferent vagal anti-inflammatory activity (Pavlov, Wang, Czura, Friedman & Tracey, 2003). A recent study to investigate α7nAChR expression on macrophages suggests that the α7 subunit on the macrophage is necessary for the suppression of TNF-involvement in the cholinergic anti-inflammatory pathway (Wang et al., 2003). The same study determined that the inhibition of the α7nAChR in the presence of nicotine was successful in restoring TNF production. The significance of the α7 subunit in mediating cholinergic anti-inflammatory activity is reinforced through the investigation of α7 subunit-deficient mice. By releasing increased amounts of TNF, IL-1 and IL-6 into the serum from macrophages, gene knockout mice have an increased sensitivity to pro-inflammatory stimuli as compared to wild-type mice (Pavlov et al., 2003; Wang et al., 2003). Additionally, the macrophages of these gene knockout mice did not cease TNF production in the presence of nicotine and ACh, further suggesting that α7nAChR presentation on macrophages is critical in the inflammatory pathway.

Contradictory responses have been observed in inflammatory disease progression with the introduction of nicotine through smoking. The suggestion that nicotine could be an appropriate treatment for some inflammatory diseases was made through a number of clinical trials that implicate transdermal nicotine as a potentially successful treatment for active UC, an inflammatory disease that has genetic overlap with AS (Jonge & Ulloa, 2007). While nicotine appears to silence the inflammatory effects of acute UC, it has no therapeutic effect in the inflammatory process of Crohn's disease. The difference between the two IBD responses to nicotine may be attributed to different pathogenic effects,

different macrophage responses or different response mechanisms to anticholinergic affects associated with the nAChR. Regardless of the exact mechanism, these two co-familial diseases appear to have different responses to nicotine agonists via the nAChR. Other inflammatory diseases, namely arthritis, appear to exhibit variable responses to nicotine as well. The introduction of nicotine seems to enhance an individual's susceptibility to rheumatoid arthritis while decreasing their risk of osteoarthritis (Papadopoulos et al., 2005). The nAChR may be susceptible to a variety of different mutations or deletions across inflammatory diseases that alter the binding capacity for ACh or nicotine. Variability in disease response to nicotine is apparent when comparing smoking versus non-smoking cases. Our study revealed an association between AS mSASSS score, disease progression and the *CHRNA7* gene that is responsible for the nAChR production; however, unlike UC, which seems to express improved disease outcomes with the introduction of nicotine, smoking has been observed to worsen disease outcomes in AS patients. As was previously discussed, more rapid progression and worsened outcomes have been observed in AS disease activity and radiographic progression in affected individuals who smoke (Kaan & Ferda, 2005; Mattey et al., 2011; Poddubnyy et al., 2012).

An alternative proposal is that the disease progression of AS is not strictly the result of an inflammatory response. AS is an inflammatory disease that is primarily characterized by the formation of new bone. Radiographic progression and the formation of new spinal syndesmophytes is, in part, an inflammatory process. Thus it's expected that a gene, such as *CHRNA7,* primarily involved in inflammation and cytokine release would be chiefly responsible for new bone formation. The initial inflammatory

presentation peripherally at the site of entheses and axially at the site of new bone proliferation is essential for triggering the progression of the disease; however, the rate at which new bone forms to produce radiographic changes and sydesmophytes does not appear to be simply a product of the inflammatory response (Slobodin, Rosner & Odeh, 2013). The exact mechanism behind the radiographic progression of AS thus remains unknown. This suggests that there may be an alternative mechanism of disease progression beyond the effects of the initial points of inflammation. An unknown facet may participate in the interaction with nicotine to hasten radiographic progression; however, we did not find any credible link between *CHRNA7* and new bone formation.

There are several limitations in this study that should be considered when interpreting the results. First, some of the clinical variables, such as IBD, psoriasis, uveitis and smoking history, were self-reported by the patient. A rheumatologist did not always diagnose the presence of extra-articular features at the time of clinical interview (e.g., a history of ever having the condition was recorded only if the patient reported it). A rheumatologist was thus unable to independently diagnose the comorbidities in all cases when a disease was treated and not present at the time of data collection. However we confident in our reporting as in the SPAARCC protocol, the presence of extra-articular features was determined by asking that patient about symptoms of IBD, uveitis or psoriasis. Once the patient stated they had symptoms compatible with these symptoms, they were asked if the diagnosis was made by a specialist (i.e. gastroenterologist for IBD, ophthalmologist for uveitis and dermatologist for psoriasis). Only if the diagnosis of IBD was made by a gastroenterologist , would this be counted in the database. As a result we are fairly

confident about our diagnosis. Also our prevalence of extra-articular features for our cohort is in keeping with previously published rates.

Treatment history could have also confounded the presence or absence of comorbidities. How treatment might have altered disease progression was not taken into account for this study, so we can neither confirm nor rule out that medications played a role in pathogenesis. This is especially true for systemic medications. Treatments such as biological agents for AS could conceal the presence of other inflammatory diseases thus decreasing the possibility of IBD, psoriasis, or uveitis presentation. Medications could also reduce the possibility of developing one of the AS comorbidities despite a genetic and/or environmental predisposition to the condition. When our study was conducted there were two classes of treatment: anti-inflammatory agents and TNF inhibitors. The TNF inhibitors were used by approximately 25% of the patients. The remaining majority of patients used anut-inflammatory agents. The use of anti-inflammatory agents will not alter the development of extra-articular features but the biologic agents may affect the presentation of the features. This is because TNF monoclonal antibodies are effective in treating psoriasis, IBD and uveitis. Meanwhile TNF receptor antagonists, such as Etanercept, are effective in treating psoriasis. As TNF agents were being used many years after the onset of the disease, a large number of patients had already developed their extra-articular features. The frequency of extra-articular features in our population is within the range of what has been reported in the literature. Thus it is possible that we may have masked extra-articular features in a subset of patients receiving TNF- inhibitors, but the effect of this was likely minimal. Regarding the

second part of the study, continuous NSAID use only marginally affects disease progression and despite the effectiveness of TNF inhibitors in treating extra-articular symptoms, the pivotal clinical trails have failed to show a delaying of radiographic progression. Many long term studies, such as the recent study from the OSKAR database, have also failed to show alternation of the mSASS scores with as long as five years of TNF inhibitor therapy (Kim et al., 2016). Thus we do not think treatment of TNF- inhibitors for a median of three years – as in our study – greatly influenced the radiographic progression in the majority of the patients.

There may have been recall bias with smoking history that resulted in an exposure misclassification. The patient reported smoking history if they had ever smoked, not if they were a smoker at the time of clinical examination. It is possible that a patient could have misreported smoking history for any number of reasons. For the purposes of this study, we used a positive or negative history of smoking in the analysis, not a detailed pack-year history. This limited our ability to determine if there was a dose-dependent relationship between mSASS scores and smoking history.

Limitations in study design may have also affected results. A blind assessment was not performed: rheumatologists ascertained patient information in a clinical setting and not in a controlled trial setting. As a result, there may be researcher bias, especially in the reporting of the radiographs for mSASS scores.

Another limitation regarding the analysis portion of this study stems from inconsistencies in recording clinical patient information. Phenotypic information was not collected for all 896 patients despite complete genotyping. Varying amounts of

information were missing from IBD, psoriasis, uveitis and *HLA-B*27* status causing there to be smaller sample sizes for some Fisher's exact tests – psoriasis in particular. SPARRC was designed as a real world prospective cohort. Relevant clinical and disease information was collected by participating sites during routine clinical visits. It was obtained over many years by rheumatology fellows and rheumatologists. Most clinicians received a few training sessions regarding the variables being collected, and this helped standardize the process. At times during clinic visit, physicians were short of time, and under these circumstances, the data was not systematically collected. Also it has been our experience, that some physicians appear to transcribe only positive attributes, and would leave the assessment blank if the event did not occur. This created a difficult situation – as we were not sure if these patients were negative for the extra-articular feature or if the assessment was missed. It was for this reason these patients were not included in the assessment. In addition, although a central protocol was used, each site had it own area of interest and this resulted in various parts of the protocol being more adhered to by certain sites. The clinical data in Alberta was collected by a large network of physicians who were capturing the data and then sending the patient to University of Alberta. The primary interest for the university was the development of a MRI and radiographic scoring system. University of Toronto had the most through clinical assessments and data collection. They were able to bring their patients back in a more regular fashion and had data entry personal working closely with the clinicians. The Newfoundland site was the smallest of the three core sites and focused exclusively on genetics. Follow up was sporadic due to the volume of urgent patients that need to be seen by a limited number of

rheumatologists. It would have been ideal to have more detailed data about smoking. Categorizing smokers as never, past, and current with the addition of pack year history would have beneficial for our analysis. Unfortunately, the importance of capturing detailed smoking information was not appreciated when the SPARCC protocol was initially rolled out and pack year history was not consistently available for each patient. This missing data may have had some bearing on the level of significance. However, to date the clinical data of those with missing data does not appear to be different from those with a complete set of clinical data.

Additionally, this study represented an association-based analysis. Though association studies are the analysis of choice for a cohort of unrelated individuals, like the one used in this thesis, there are characteristics that cause restriction in the understanding of our findings. Specifically, significant results only represent an observation of co- occurrence between alleles and phenotypes, not a causal relationship. High rates of false positive associations, false negative associations and a true association that is population- specific (i.e., relatively rare) are contributing factors in impeding the replication of the results from association analysis studies. Also, multiple testing should be accounted for in the case of an allele not being independent of LD. It is difficult to discern if an association study is successfully replicated because the SNP under study is a causative agent to the disease, or if the SNP is in a LD with the authentic disease mutation. As a result, there is a high frequency of false positive results in association analyses. The power of association studies is another important factor as underpowered studies may result in false negative findings. Our multiple linear regression analyses contained 76 patients, only half of which were smokers. This sample size may not be large enough to appreciate the true effects of

environmental factors with genetic interactions and smoking. With such a small sample size it is difficult to distinguish whether the genetic factors alone impacted disease severity and progression or if the interaction of smoking with the genes played a role in worsening severity and more rapid progression.

4.2 - Conclusion

The results of this study suggested that the *UGT2B17* gene found to be associated with AS in copy number 2 in a multiplex family was not significantly associated with the extra-articular features of AS, nor was it significantly associated with *HLA-B*27*. It appears unlikely that AS heritability stems from a gene-gene interaction between *UGT2B17* and *HLA-B*27*; however, there is the potential that a portion of the missing heritability of AS can be explained through *UGT2B17* as a rare variant. The recent identification of *SEC16A* and *MAMDC4* suggests that this is a plausible interpretation. Future studies could investigate *UGT2B17* as a rare variant using exome sequencing.

This study also identified that mSASSS score and disease progression were each significantly associated with two SNPs located within *CHRNA7,* a gene implicated in the mediation of inflammatory responses and the production of TNF. This finding suggests that smoking may interact with the receptor product of this gene to exacerbate the inflammatory response; however, interaction with smoking, genetic variants related to nicotine metabolism and ankylosis is not clearly demonstrated in this study. Collectively, the results of this study make a convincing argument to continue this research.

References

Alkan, C., Coe, B. P., & Eichler, E. E. (2011). Genome structural variation discovery and genotyping. *Nature Reviews Genetics, 12*(5), 363-376.

Angstadt, A., Berg, A., Zhu, J., Miller, P., Hartman, T. J. Lesko, S., . . . Gallagher, C. J. (2013).The Effect of Copy Number Variation in the Phase II Detoxification Genes UGT2B17 and UGT2B28 on Colorectal Cancer Risk. *Cancer, 119*(13), 2477-2485.

Apel, M., Uebe, S., Bowes, J., Giardina, E., Korendowych, E., Juneblad, K., . . . Huffmeier, U. (2013). Variants in RUNX3 contribute to susceptibility to psoriatic arthritis, exhibiting further common ground with ankylosing spondylitis. *Arthritis and Rheumatism, 65*(5), 1224-1231.

Atzeni, F., Defendenti, C., Ditto, M. C., Alberto Batticciotto, A., Ventura, D., Antivalle, M., . . . Sarzi-Puttini, P. (2014). Rheumatic Manifestations of Inflammatory Bowel Disease. *Autoimmunity Reviews, 13*(1), 20-23.

Australo-Anglo-American Spondyloarthritis Consortium (TASC). (2012). Evidence of genetic association between TNFRSF1A encoding the p55 tumour necrosis factor receptor, and ankylosing spondylitis in UK caucasians. *Clinical and Experimental Rheumatology, 30*(1), 110-113.

Australo-Anglo-American Spondyloarthritis Consortium (TASC). (2010). Genome-wide association study of ankylosing spondylitis identifies non-MHC susceptibility loci. *Nature Genetics, 42*(2), 123-127.

Australo-Anglo-American Spondyloarthritis Consortium (TASC) & Spondyloarthritis

Research Consortium of Canada (SPARCC). (2010). Association of variants at 1q32 and STAT3 with ankylosing spondylitis suggests genetic overlap with Crohn's disease. *PLoS Genetics, 6*(12), e1001195.

Australo-Anglo-American Spondyloarthritis Consortium (TASC) & Wellcome Trust Case Control Consortium 2 (WTCCC2). (2007). Association scan of 14,500 nonsynonymous SNPs in four diseases identifies autoimmunity variants. *Nature Genetics, 39*(11), 1329-1337.

Australo-Anglo-American Spondyloarthritis Consortium (TASC) & the Wellcome Trust Case Control Consorium 2 (WTCCC2). (2011). Interaction between ERAP1 and HLA-B27 in ankylosing spondylitis implicates peptide handling in the mechanism for HLA-B27 in disease susceptibility. *Nature Genetics, 43*(8), 761-767.

Averns, H. L., Oxtoby, J., Taylor, H. G., Jones, P. W., Dziedzic, K., & Dawes, P. T. (1996).Smoking and outcome in ankylosing spondylitis. *Scandinavian Journal of Rheumatology, 25*(3), 138-142.

Baeten, D., Sieper, J., Braun, J., Baraliakos, X., Dougados, M., Emery, P., . . . Richards, H. B. (2015). Secukinumab, an Interleukin-17A Inhibitor, in Ankylosing Spondylitis. *The New England Journal of Medicine*, *373*(26), 2534-2548.

Berg, J. Z., von Weymarn, L. B., Thompson, E. A., Wickham, K. M., Weisensel, N. A., Hatsukami, D. K., & Murphy, S. E. (2010). UGT2B10 genotype influences nicotine glucuronidation, oxidation, and consumption. *Cancer Epidemiology, Biomarkers & Prevention: A Publication of the American Association for Cancer*

Research, Cosponsored by the American Society of Preventive Oncology, 19(6), 1423-1431.

Boonen, A. & van der Linden, S. M. (2006). The Burden of Ankylosing Spondylitis. *The Journal of Rheumatology, 33*(78), 4-11.

Borovikova, L.V., Ivanova, S., Yang, H., Botchakina, G., Watkins, L., Abumrad, N., . . . Tracey, K. (2000). Vagus nerve stimulation attenuates the systemic inflammatory response to endotoxin. *Nature*, *405*(6785), 458-461.

Breban, M. (2006). Genetics of spondyloarthritis. *Best Practice & Research. Clinical Rheumatology, 20*(3), 593-599.

Brown, M. A. (2008). Breakthroughs in genetic studies of ankylosing spondylitis. *Rheumatology, 47*(2), 132-137.

Brown, M.A. (2009). Genetics and the pathogenesis of ankylosing spondylitis. *Current Option in Rheumatology, 21*(4), 318-323.

Brown, M. A. (2011). Progress in the genetics of ankylosing spondylitis. *Briefings in Functional Genomics, 10*(5), 249-257.

Brown, M. A., Crane, A. M., & Wordsworth, B. P. (2002). Genetic aspects of susceptibility, severity, and clinical expression in ankylosing spondylitis. *Current Opinion in Rheumatology, 14*(4), 354-360.

Brown, M. A., Kenna, T., & Wordsworth, B. P. (2015). Genetics of Ankylosing spondylitis insights into pathogenesis. *Nature Reviews Rheumatology*. doi: 10.1038/nrrheum.2015.133

Brown, M. A., Kennedy, L. G., MacGregor, A. J., Darke, C., Duncan, E., Shatford, J., . . . Wordsworth, P. (1997). Susceptibility to ankylosing spondylitis in twins: the role

of genes, HLA, and the environment. *Arthritis and Rheumatism*, *40*(10), 1823 1828.

Chandran, V. & Rahman, P. (2010). Update on the genetics of spondyloarthritis ankylosing spondylitis and psoriatic arthritis. *Best Practice & Research Clinical Rheumatology, 24(*2010), 579-588.

Chen, G., Giambrone, N. E., Jr., Dluzen, D. F., Muscat, J. E., Berg, A., Gallagher, C. J., & Lazarus, P. (2010). Glucuronidation genotypes and nicotine metabolic phenotypes: Importance of functional UGT2B10 and UGT2B17 polymorphisms. *Cancer Research, 70*(19), 7543-7552.

Chen, G., Giambrone, N. E., & Lazarus, P. (2012). Glucuronidation of trans-3' hydroxycotinine by UGT2B17 and UGT2B10. *Pharmacogenetics and Genomics, 22*(3), 183-190.

Chung, H. Y., Machado, P., van der Heijde, D., D'Agostino, M. A., & Dougados, M. (2012). Smokers in early axial spondyloarthritis have earlier disease onset, more disease activity, inflammation and damage, and poorer function and health-related quality of life: Resultsfrom the DESIR cohort. *Annals of the Rheumatic Diseases, 71*(6), 809-816.

Ciccia, F., Accardo-Palumbo, A., Alessandro, R., Rizzo, A., Principe, S., Peralta, S., . . . Triolo, G. (2012). Interleukin-22 and interleukin-22-producing NKp44+ natural killer cells in subclinical gut inflammation in ankylosing spondylitis. *Arthritis and Rheumatism, 64*(6), 1869-1878.

Conrad, D. F., Pinto, D., Redon, R., Feuk, L., Gokcumen, O., Zhang, Y., . . . Hurles, M. E. (2010). Origins and functional impact of copy number variation in the human genome. *Nature, 464*(7289), 704-712.

Danoy, P., Pryce, K., Hadler, J., Bradbury, L. A., Farrar, C., Pointon, J., . . . Brown, M. A. (2010). Association of variants at 1q32 and STAT3 with ankylosing spondylitis suggests genetic overlap with crohn's disease. *PLoS Genetics, 6*(12), e1001195.

Davidson, S. I., Liu, Y., Danoy, P. A., Wu, X., Thomas, G. P., Jiang, L., . . . Xu, H. (2011).Association of STAT3 and TNFRSF1A with ankylosing spondylitis in han chinese. *Annals of the Rheumatic Diseases, 70*(2), 289-292.

De Vos, M., Hindryckx, P., & Laukens, D. (2011). Novel development in extraintestinal manifestations and spondylarthropathy. *Best Practice & Research Clinical Gastroenterology, 25*(Suppl 1), S19-26.

Di Meglio, P., Di Cesare, A., Laggner, U., Chu, C., Napolitano, L., Villanova, F., . . . Nestle, F. O. (2011). The IL23R R381Q gene variant protects against immune mediated diseases by impairing IL-23-induced Th17 effector response in humans. *PloS One, 6*(2), e17160.

Dougados, M. & Baeten, D. (2011). Spondyloarthritis. *Lancet, 377*(9783), 2127-2137.

Duffin, K. C., Freeny, I. C., Schrodi, S. J., Wong, B., Feng, B. J., Soltani-Arabshahi, R., . . . Krueger, G. G. (2009). Association between IL13 polymorphisms and psoriatic arthritis is modified by smoking. *The Journal of Investigative Dermatology, 129*(12), 2777-2783.

Edavalath, M. (2010). Ankylosing spondylitis. *Journal of Ayurveda and Integrative Medicine, 1*(3), 211-214.

Eder, L., Chandran, V., Pellett, F., Pollock, R., Shanmugarajah, S., Rosen, C. F., . . . Gladman, D. D. (2011). IL13 gene polymorphism is a marker for psoriatic arthritis among psoriasis patients. *Annals of the Rheumatic Diseases, 70*(9), 1594-1598.

Ekstrom, L., Cevenini, L., Michelini, E., Schulze, J., Thorngren, J. O., Belanger, A., . . . Rane, A. (2013). Testosterone challenge and androgen receptor activity in relation to UGT2B17 genotypes. *European Journal of Clinical Investigation, 43*(3), 248-255.

Fantini, M. C., Pallone, F., & Monteleone, G. (2009). Common immunologic mechanisms ininflammatory bowel disease and spondylarthropathies. *World Journal of Gastroenterology, 15*(20), 2472-2478.

Franke, L., Bannoudi, H., Jansen, D. T., Kok, K., Trynka, G., Diogo, D., . . . Zhernakov, A. (2015). Association analysis of copy numbers of FC-gamma receptor genes for rheumatoid arthritis and other immune-mediated phenotypes. *European Journal of Human Genetics*. Advance online publication. Retrieved from: http://www.nature.com/ejhg/journal/vaop/ncurrent/full/ejhg201595a.html

Giroux, S., Bussieres, J., Bureau, A., & Rousseau, F. (2012). UGT2B17 gene deletion associated with an increase in bone mineral density similar to the effect of hormone replacement in postmenopausal women. *Osteoporosis International, 23*(3), 1163-1170.

Groupe Français d'Etude Génétique des Spondylarthropathies. (2004). Significant linkage to spondyloarthropathy on 9q31-34. *Human Molecular Genetics*, *13*(15), 1641-1648.

Haroon, N. N., Paterson, J. M., Li, P., Inman, R. D. & Haroon, N. (2015). Patients with Ankylosing Spondylitis Have Increased Cardiovascular and Cerebrovascular Mortality: A Population Based Study. *Annals of Internal Medicine, 163*(6), 409-416.

Huang, X., Li Y., Tanaka K., Moore K. G., & Hayashi, J. I. (1995). Cloning and characterization of Lnk, a signal transduction protein that links T-cell receptor activation signal to phospholipase Cγ1, Grb2, and phosphatidylinositol 3-kinase. *Proceedings of the National Academy of Science of the United States of America*, *92*, 11618-11622.

International Genetics of Ankylosing Spondylitis Consortium (IGAS). (2013). Identification of multiple risk variants for ankylosing spondylitis through high density genotyping of immune related loci. *Nature Genetics, 45*(7)*,* 730-738.

Iwakura, Y. & Ishigame, H. (2006). The IL-23/IL-17 axis in inflammation. *The Journal of Clinical Investigation, 116*(5), 1218-1222.

Jadon, D., Tillett, W., Wallis, D., Cavill, C., Bowes, J., Waldron, N., . . . McHugh, N. J. (2013).Exploring ankylosing spondylitis-associated ERAP1, IL23R and IL12B gene polymorphisms in subphenotypes of psoriatic arthritis. *Rheumatology, 52*(2), 261-266.

Jarvinen, P. (1995). Occurrence of ankylosing spondylitis in a nationwide series of twins. *Arthritis and Rheumatism, 38*(3), 381-383.

Jonge, W. J. & Ulloa, L. (2007). The alpha7 nicotinic acetylcholine receptor as a pharmacological target for inflammation. *British Journal of Pharmacology, 151*(7), 915-929.

Jung, S. H., Yim, S. H., Hu, H. J., Lee, K. H., Lee, J. H., Sheen, D. H., . . . Chung, Y. (2014).Genome-wide copy number variation analysis identifies deletion variants associated with ankylosing spondylitis. *Arthritis and Rheumatology, 66*(8), 2103-2112.

Kaan, U. & Ferda, O. (2005). Evaluation of clinical activity and functional impairment in smokers with ankylosing spondylitis. *Rheumatology International, 25*(5), 357-360.

Kenna, T., Davidson, S., & Thomas, G. (2011). The genomics and genetics of ankylosing spondylitis. *Advances in Genomics and Genetics, 1*, 9-25.

Khan, M.A., Kushner, I., & Braun, W.E. (1981). Association of HLA-A2 with uveitis in HLA-B27 positive patients with ankylosing spondylitis. *The Journal of Rheumatology*, *8*(2), 295-298.

Kim T.J., Shin J.H., Kim S., Sung I.H., Lee S., Song Y. & Kim T.H. (2016). Radiographic progression in patients with ankylosing spondylitis according to tumor necrosis factor blocker exposure: Observation Study of Korean Spondyloarthropathy Registry (OSKAR) data. *Joint, Bone, Spine*. [Epub ahead of print].

Kleinjan, D.A. & van Hayningen, V. (2005). Long-range control of gene expression: emerging mechanisms and disruption in disease. *American Journal of Human Genetics, 76*(1), 8-32.

Kugathasan, S., Baldassano, R. N., Bradfield, J. P., Sleiman, P. M., Imielinski, M., Guthery, S. L., . . . Hakonarson, H. (2008). Loci on 20q13 and 21q22 are associated with pediatric-onset inflammatory bowel disease. *Nature Genetics, 40*(10), 1211-1215.

Kurotaki, N., Shen, J.J., Touyama, M., Kondoh, T., Visser, R., Ozaki, T., . . . Lupski, J.R. (2005).Phenotypic consequences of genetic variation at hemizygous alleles: Sotos syndrome is a contiguous gene syndrome incorporating coagulation factor twelve (FXII) deficiency. *Genetics Medicine: official journal of the American College of Medical Genetics, 7*(7), 479-483.

Laval, S. H., Timms, A., Edwards, S., Bradbury, L., Brophy, S., & Milicic, A. (2001). Whole-genome screening in ankylosing spondylitis: evidence of non-MHC genetic-susceptibility loci. *American Journal of Human Genetics, 68*(4), 918-926.

Lee, C., Iafrate, A. J., & Brothman, A. R. (2007). Copy nuber variations and clinical cytogenetic diagnosis of constitutional disorders. *Nature Genetics, 39*(7 Suppl), S48-54.

Lees, C. W., Barrett, J. C., Parkes, M., & Satsangi, J. (2011). New IBD genetics: common pathways with other diseases. *Gut, 60*(12), 1739-1753.

Lian, Z., Chai, W., Shi, L. L., Chen, C., Liu, J., & Wang, Y. (2013). Analysis of PPARGC1B, RUNX3 and TBKBP1 polymorphisms in chinese han patients with ankylosing spondylitis: A case-control study. *PloS One, 8*(4), e61527.

Lupski, J. R., Wise, C.A., Kuwano, A., Pentao, L., Parke, J.T., Glaze, D. G., Ledbetter, D. H., . . . Patel, P. I. (1992). Gene dosage is a mechanism for Charcot Marie-Tooth disease type 1A. *Nature Genetics, 1*(1), 29-33.

Mattey, D. L., Dawson, S. R., Healey, E. L., & Packham, J. C. (2011). Relationship between smoking and patient-reported measures of disease outcome in ankylosing spondylitis. *The Journal of Rheumatology, 38*(12), 2608-2615.

Matzkies, F. G., Targan, S. R., Berel, D., Landers, C. J., Reveille, J. D., McGovern, D. P., & Weisman, M. H. (2012). Markers of intestinal inflammation in patients with ankylosing spondylitis: A pilot study. *Arthritis Research & Therapy, 14*(6), R261, 1-8.

Morota, G., Valente, B. D., Rosa, G. J., Weigel, K. A., & Gianola, D. (2012). An assessment of linkage disequilibrium in holstein cattle using a bayesian network. *Journal of Animal Breeding and Genetics, 129*(6), 474-487.

Muto, A., Ochiai, K., Kimura, Y., Itoh-Nakadai, A., Calame, K., Ikebe, D., . . . Igarashi, K. (2010). Bach2 represses plasma cell gene regulatory network in B cells to promote antibody class switch. *The EMBO Journal, 29*(23), 4048-4061.

National Center for Biotechnology Information. (2015a). NKX2-3 NK2 homeobox 3 [*Homo sapiens* (human)] Gene ID: 159296. Retrieved from: http://www.ncbi.nlm.nih.gov/gene/159296

National Center for Biotechnology Information. (2015b). NOS2 nitric oxide synthase 2, inducible [*Homo sapiens* (human)]. Gene ID: 4843. Retrieved from: http://www.ncbi.nlm.nih.gov/gene/4843

National Institute for Health and Clinical Excellence Technology Appraisal Guidance. (2008).Appendix C: Modified New York Criteria for diagnosis of Ankylsoing Spondylitis. In Adalimumab, etanercept and infliximab for ankylosing spondylitis (pp. 48).

O'Rielly, D. D., Uddin, M., Codner, D., Hayley, M., Zhou, J., Pena-Castillo, L., . . . Rahman, P. (2015). Private rare deletions in SEC16A and MAMDC4 may represent novel pathogenic variants in familial axial spondyloarthritis. *Annals of*

the Rheumatic Disease. Advance online publication. Retrieved from: http://ard.bmj.com/content/early/2015/06/16/annrheumdis-2014-206484.long

Papadopoulos, N.G., Alamanos, Y., Voulgari, P.V., Epagelis, E.K., Tsifetaki, N., & Drosos, A.A. (2005). Does cigarette smoking influence disease expression, activity and severity in early rheumatoid arthritis patients?. *Clinical and Expermimental Rheumatology, 23*(6), 861-866.

Pâquet, S., Fazli, L., Grosse, L., Verreault, M., Têtu, B., Rennie, P., . . . Barbier, O. (2012).Differential Expression of the Androgen-Conjugating UGT2B15 and UGT2B17 Enzymes in Prostate Tumor Cells during Cancer Progression**.** *The Journal of Clinical Endocrinology and Metabolism, 97*(3), e428-432.

Patnala, R., Clements, J., & Batra, J. (2013). Candidate gene association studies: A comprehensive guide to useful in silico tools. *BMC Genetics, 14*(39), 1-11.

Pavlov, V. A., Wang, H., Czura, C., Friedman, S. G. & Tracey, K. J. (2003). The Cholinergic Anti-inflammatory Pathway: A Missing Link in Neuroimmunomodulation. *Molecular Medicine, 9*(5-8), 125-134.

Pearce, E.L., Mullen, A.C., Martins, G.A., Krawczyk, C.M., Hutchins, A.S., Zediak, V.P., . . . Reiner, S.L. (2003). Control of effector CD8+ T cell function by the transcription factor Eomesodermin. *Science, 302*(5647), 1041-1043.

Plafker, K. S. & Plafker, S. M. (2015). The ubiquitin- conjugating enzyme UBE2E3 and its import receptor importin-11 regulate the localization and activity of the antioxidant transciption factor NRF2. *Molecular Biology of the Cell, 26*(2), 327-338.

Poddubnyy, D., Haibel, H., Listing, J., Marker-Hermann, E., Zeidler, H., Braun, J., . . .

Rudwaleit, M. (2012). Baseline radiographic damage, elevated acute-phase reactant levels, and cigarette smoking status predict spinal radiographic progression in early axial spondylarthritis. *Arthritis and Rheumatism, 64*(5), 1388-1398.

Poddubnyy, D. & Sieper, J. (2012). Radiographic progression in ankylosing spondylitis/axial spondyloarthritis: How fast and how clinically meaningful?. *Current Opinion in Rheumatology, 24*(4), 363-369.

Pointon, J. J., Harvey, D., Karaderi, T., Appleton, L. H., Farrar, C., Stone, M. A., . . . Wordsworth, B. P. (2010a). Elucidating the chromosome 9 association with AS; CARD9 is a candidate gene. *Genes and Immunity, 11*(6), 490-496.

Pointon, J. J., Harvey, D., Karaderi, T., Appleton, L. H., Farrar, C., . . . Wordsworth, B .P. (2010b). The chromosome 16q region associated with ankylosing spondylitis includes the candidate gene tumour necrosis factor receptor type 1-associated death domain (TRADD). *Annals of Rheumatic Disease*, 69(6), 1243-1246.

Ramiro, S., van Tubergen, A., Stolwijk, C., Landewé, R., van de Bosch, F., Dougados, M., & van der Heijde, D. (2013). Scoring radiographic progression in ankylosing spondylitis: Should we use the modified stoke ankylosing spondylitis spine score (mSASSS) or the radiographic ankylosing spondylitis spinal score (RASSS)?. *Arthritis Research & Therapy, 15*(1), 1-9.

Reveille, J. D. (2011). The genetic basis of spondyloarthritis. *Annals of the Rheumatic Diseases, 70*(Suppl 1), i44-i50.

Reveille, J. D. (2012). Genetics of spondyloarthritis – beyond the MHC. *Nature Reviews Rheumatology, 8*(5), 296-304.

Reveille, J.D. (2014). An update on the contribution of the MHC to AS susceptibility. *Clinical Rheumatology, 33*(6), 749-757.

Robinson, P. C. & Brown, M. A. (2012). The Genetics of Ankylosing Spondylitis and Axial Spondyloarthritis. *Rhuematic Disease Clinics of North America, 38*(3), 539- 553.

Rosenbaum, J. & Chandran, V. (2012). Management of comorbidities in ankylosing spondylitis. *The American Journal of the Medical Sciences, 343*(5), 364-366.

Rothem, D. E., Rothem, L., Soudry, M., Dahan, A., & Eliakim, R. (2009). Nicotine modulates bone metabolism-associated gene expression in osteoblast cells. *Journal of Bone and Mineral Metabolism, 27*(5), 555-561.

Rudwaleit, M., van der Heijde, D., Landewé, R., Listing, J., Akkoc, N., Brandt, J., . . . Sieper, J. (2009). The development of Assessment of SpondyloArthritis international Society classification criteria for axial spondyloarthritis (part II): validation and final selection. *Annals of the Rheumatic Diseases*, *68*, 777-783.

Ruutu, M., Thomas, G., Steck, R., Degli-Esposti, M. A., Zinkernagel, M. S., Alexander, K., . . . Thomas, R. (2012). Beta-glucan triggers spondylarthritis and crohn's disease-like ileitis in SKG mice. *Arthritis and Rheumatism, 64*(7), 2211-2222.

Salaffi F., Carotti M., Garofalo G., Giuseppetti G.M. & Grassi W.(2007).Radiological scoring methods for ankylosing spondylitis: a comparison between the Bath Ankylosing Spondylitis Radiology Index and the modified Stoke Ankylosing Spondylitis Spine Score. *Clin Exp Rheumatol. 2*5(1): 67-74.

Salvarani, C. & Fries, W. (2009). Clinical features and epidemiology of spondyloarthritides associated with inflammatory bowel disease. *World Journal*

of Gastroenterology, 15(20), 2449-2455.

Scherer, S. W., Lee, C., Birney, E., Altshuler, D. M., Eichler, E. E., Carter, N.P., Hurles, M. E., & Feuk, L. (2007). Challenges and Standards in integrating surveys of structural variation. *Nature Genetics, 39*(7 Suppl), S7-15.

Sieper, J., Listing, J., Poddubnyy, D., Song, I., Hermann, K., Callhoff, J., . . . Rudwaleit, M. (2015). Effect of continuous versus on-demand treatment of ankylosing spondylitis with diclofenac over 2 years on radiographic progression of the spine: results from a randomised multicenter trial (ENRADAS). *Annals of the Rheumatic Diseases*. doi: 10.1136/annrheumdis-2015-207897

Sims, A. M., Wordsworth, B. P., & Brown, M. A. (2004). Genetic susceptibility to ankylosing spondylitis. *Current Molecular Medicine, 4*(1), 13-20.

Slobodin, G., Rosner, I., & Odeh, M. (2013). Bone Formation in Ankylosing Spondylitis. *OA Arthritis, 1*(1), 1-7.

Spondylitis Association of America. (2013). Ankylosing Spondylitis: Most Common Symptoms. Retrieved from: http://www.spondylitis.org/about/as_sym.aspx

Tillett, W., Jadon, D., Shaddick, G., Cavill, C., Korendowych, E., de Vries, C. S., & McHugh, N. (2013). Smoking and delay to diagnosis are associated with poorer functional outcome in psoriatic arthritis. *Annals of the Rheumatic Diseases, 72*(8), 1358-1361.

Tsui, F. W., Tsui, H. W., Akram, A., Harron, N., & Inman, R. D. (2014). The genetic basis of ankylosing spondylitis: new insights into disease pathogenesis. *The application of clinical genetics, 7*, 105-115.

Turgeon, D., Carrier, J. S., Levesque, E., Hum, D. W., & Belanger, A. (2001). Relative

enzymatic activity, protein stability, and tissue distribution of human steroid-metabolizing UGT2B subfamily members. *Endocrinology, 142*(2), 778-787.

Uddin, M., Maksymowych, W. P., Inman, R., Gladman, D., Munn, A., Yazdani, R., . . . Rahman, P. (2013). UGT2B17 copy number gain in a large ankylosing spondylitis multiplex family. *BMC Genetics, 14*(67), 1-6.

Uddin, M., Sturge, M., Peddle, L., O'Rielly, D. D. & Rahman, P. (2011). Genome-Wide Signatures of 'Rearrangement Hotspots' within Segmental Duplications in Humans. *PLoS ONE*, *6*(12), e28853.

van der Linden, S., Valkenburg, H., de Jongh, B., & Cats, A. (1984). The risk of developing ankylosing spondylitis in HLA-B27 positive individuals. A comparison of relatives of spondylitis patients with the general population. *Arthritis and Rhuematism, 27*(3), 241-249.

Wang, H., Yu, M., Ochani, M., Amella, C. A., Tanovic, M., Susarla, S., . . . Tracey, K. J. (2003).Nicotinic acetylcholine receptor alpha7 subunit is an essential regulator of inflammation. *Nature*, *421*(6921), 384-388.

Wang, L., Yang, X., Cai, G., Xin, L., Xia, Q., Zhang, X., Li, X., . . . Pan, F. (2015). Association study of copy number variants in FCGR3A and FCGR3B gene with risk of ankylosing spondylitis in a Chinese population. *Rheumatol Int*, *36*(3), 437-442.

Ward, M. M., Hendrey, M. R., Malley, J. D., Learch, T. J., Davis, J. C., Jr., Reveille, J. D., & Weisman, M. H. (2009). Clinical and immunogenetic prognostic factors

for radiographic severity in ankylosing spondylitis. *Arthritis and Rheumatism, 61*(7), 859-866.

Wendling, D., Prati, C., Demattei, C., Miceli-Richard, C., Daures, J. P., & Dougados, M. (2012).Impact of uveitis on the phenotype of patients with recent inflammatory back pain: Data from a prospective multicenter french cohort. *Arthritis Care & Research, 64*(7), 1089-1093.

Wong, R. H., Wei, J. C., Huang, C. H., Lee, H. S., Chiou, S. Y., Lin, S. H., . . . Yang, S. F. (2012). Association of IL-12B genetic polymorphism with the susceptibility and disease severity of ankylosing spondylitis. *The Journal of Rheumatology, 39*(1), 135-140.

Yang, T. L., Chen, X. D., Guo, Y., Lei, S. F., Wang, J. T., Zhou, Q., . . . Deng, H. W. (2008).Genome-wide copy-number-variation study identified a susceptibility gene, UGT2B17, for osteoporosis. *American Journal of Human Genetics, 83*(6), 663-674.

Zhang, G., Luo, J., Bruckel, J., Weisman, M. A., Schumacher, H. R., Khan, M. A, . . . Reveille, J. D. (2004). Genetic studies in familial ankylosing spondylitis susceptibility. *Arthritis and Rheumatism, 50*(7), 2246-2254.

Appendix 1:

Additional Gene Variants Identified via

GWAS

Intergenic Regions

The intergenic regions at chromosomes 2p15 and 21q22 have been found to be susceptible loci for the disease in multiple different AS cohorts. The chromosomal region 21q22 has also been previously identified as containing loci eliciting pediatric IBD susceptibility. The occurrence of IBD as a comorbidity of AS indicates that the inflammatory mechanisms in both diseases might be similar. The previously identified common genetic factors in both diseases provide convincing evidence that the intergenic regions found have potential action in disease susceptibility (Kugathasan et al., 2008; Reveille, 2012). The combined influence of these regions explains 0.575% of the heritability of AS (Brown, 2011; Reveille, 2012).

IL-1 Region Genes and IL1R2

The locus *IL-1* is found at chromosome 2p15 and within this region lies *IL1R2* at 2p12 (Reveille, 2012). The gene product is the IL-1 receptor type 2 (IL1R2), which acts to bind IL-1, preventing its binding to IL1R1. When IL-1 binds, its activity is repressed. IL-1 cytokines have action in producing inflammatory responses (Reveille, 2011). IL1R2 may have a relationship with *ERAP1*, a gene previously mentioned to be related to AS susceptibility. ERAP1 has a potential role in the cleavage of the *IL1R2* product from the cell membrane. It regulates the function of *IL1R2* by decreasing its expression. ERAP1 thus inhibits the binding of the IL1R2 product to IL-1, allowing for IL-1 to produce an inflammatory effect (Reveille, 2011). *IL1R2* has very little influence in accounting for AS heritability as it is only responsible for 0.12% of the variation (Brown, 2011; Reveille, 2012).

TNF Pathway Associated Genes

AS is associated with three genes involved in the TNF pathway. *LTBR* and *TNFRSF1A* are found on the same chromosome, 12p13, and account for 0.075% of AS heritability (Brown, 2011; Reveille, 2012). *TNFRSF1A* has previously been associated with IBD (Reveille, 2012). This overlap between comorbid diseases supports the conjecture that *TNFRSF1A* is associated with inflammation pathways common to both conditions. Additionally, mice studies have indicated a potential link between *TNFRSF1A* and the development of sacroiliitis, a feature of AS pathogenesis (Kenna, Davidson & Thomas, 2011). A relationship between *TNFRSF1A* and *ERAP1* has been proposed as *ERAP1* is responsible for cleavage of *TNFR1* from the cell surface, which aids in regulating the optimal function of *TNFRSF1A* (Davidson et al., 2011; TASC, 2012).

TBKBP1, a gene located on chromosome 17q21, is a susceptible loci for AS pathogenesis, accounting for 0.054% of AS heritability (Reveille, 2012; TASC & WTCCC2, 2011). *TBKBP1* differs in genotype frequencies between AS cases and healthy controls (Lian et al., 2013).

ANTXR2

ANTXR2 is located on chromosome 4q21 and its expression codes for the protein capillary morphogenesis protein-2. It accounts for 0.054% of AS heritability (Brown, 2011; Reveille, 2012). The gene is suspected to be involved in extracellular matrix adhesion as its protein has been observed to bind to type IV collagen and laminin.

Although *ANTXR2* has been associated with AS in GWAS studies, its role in disease pathogenesis remains largely unknown (Reveille, 2011).

PTGER46

PTGER46 has been associated with AS disease pathogenesis (IGAS, 2013). It is responsible for approximately 0.052% of AS heritability (Brown, 2011; Reveille, 2012). *PTGER46* is expressed in the presence of mechanical stress and results in a protein product called prostaglandin E_2. This protein initiates the production of IL-23 by the dendritic cells, which subsequently commences the IL17-IL23 pathway that results in an inflammatory response. Additionally, prostaglandin E_2 causes bone formation, a primary characteristic of AS pathogenesis (Reveille, 2012).

KIF21B

KIF21B, located on chromosome 1q23, was identified in a study of Crohn's disease-associated genetic regions in individuals with AS (Reveille, 2012). The gene belongs to the family encoding kinesin motor proteins, which are responsible for the transport of components along microtubules. *KIF21B* is not the only gene from this family encoding kinesin motor proteins to be associated with arthritis or autoimmune disorders. *KIF5A* has been associated with rheumatoid arthritis and type 1 diabetes susceptibility (TASC & SPARCC, 2010). The loci associated with the autoimmune disorder multiple sclerosis is also located near *KIF5A*. This association of functional genes with autoimmune disorders supports the role of *KIF21B* in AS susceptibility. The

gene was found to be associated with the disease, contributing about 0.25% to the overall heritability of AS (Brown, 2011; Reveille, 2012).

EOMES

EOMES encodes a transcription factor called eomesodermin, which is expressed in the presence of RUNX3, another gene protein associated with AS. Together, EOMES and RUNX3 mediate cell differentiation of $CD8^+$ T cells (IGAS, 2013; Pearce et al., 2003).

ZMIZ1

Zinc finger MIZ domain-containing protein 1 (*ZMIZ1*) is a coactivator of an inhibitory protein that affects the $CD8^+$ T cell count through STAT-mediated cytokine signaling. It is not certain whether AS pathogenesis is directly affected by the alteration of $CD8^+$ T cell count by *ZMIZ1* (IGAS, 2013).

IL7R

IL7R is a gene associated with AS; its protein product alters $CD8^+$ T cell counts.

The gene was found to be a risk variant for AS in combination with low $CD8^+$ T cell counts (IGAS, 2013).

BACH2

Transcription regulator protein BACH2 has a complex effect on B cell differentiation and function, which elicits a response in the $CD4^+$ T cell counts. Theeffects of BACH2 transcription factor thus indirectly alter $CD4^+$ T lymphocyte count (IGAS, 2013; Muto et al., 2010).

SH2B3

SH2B3 is a leukocyte adaptor protein that has an established effect on $CD4^+$ T cell counts in AS patients. The gene product of SH2B3 is involved in T cell receptor signaling, which facilitates the conversion of $CD8^+$ T cells to cytotoxic or Th cells (Huang, Li, Tanaka, Moore & Hayashi, 1995; IGAS, 2013).

GPR35, GPR37* and *GPR65

The functions of the genes *GPR35, GPR37*, and *GPR65* are not very well understood. The study by IGAS (2013) found these loci to be associated with risk of AS. *GPR35, GPR37* and *GPR65* are G protein-coupled receptor coding loci. *GPR65* is also associated with Crohn's disease. Very little is known about the function of these loci in the pro-inflammatory process of either disease (IGAS, 2013).

HLA-A*0201

*HLA-A*02* is a MHC gene found to be associated with AS. Previously, this locus had been associated with anterior uveitis, a common extra-articular feature of AS (IGAS, 2013; Khan, Kushner & Braun, 1981).

FCGR2A

Fc fragment of IgG, low affinity IIa, receptor gene (*FCGR2A)* was found to be significantly associated with AS disease risk (IGAS, 2013). This gene has known function in coding a protein that is a cell surface receptor found on phagocytic cells. It is involved in immunological processes, mainly the clearance of immune complexes (Franke et al., 2015).

UBE2E3

Ubiquitin-conjugating enzyme E2E 3 (*UBE2E3)* is significantly associated with AS pathogenesis (IGAS, 2013). Ubiquitination is an important mechanism for targeting protein degradation. This gene is highly conserved in eukaryotes, encoding a member of the E2 ubiquitin-conjugating enzyme family. It has known function in conjugating proteins using ubiquitin (Plafker & Plafker, 2015).

NOS2

Nitric oxide synthase 2 *(NOS2)* is a gene found on chromosome 17. It is inducible by lipopolysaccharides and certain cytokines (NCBI, 2015b). This gene, which encodes

nitric oxide synthase, is expressed in the liver and contributes to AS heritability (IGAS, 2013).

NKX2-3

NK2 homeobox 3 (*NKX2-3)* encodes a protein belonging to the NKX family of homeo-domain transcription (IGAS, 2013; NCBI, 2015a).

Appendix 2:

Results Displaying Inheritance Models

(Dominant, Recessive & Additive)

Table A-1: Interaction between selected genetic variants, radiographic severity (mSASSS) and both smoking status and sex. Three inheritance models are described as Dominant (CC or CT), Recessive (TT), and Additive

SNP		**Interaction between selected genetic variants, radiographic severity (mSASSS) and smoking status**	**Interaction between selected genetic variants, radiographic severity (mSASSS) and sex**
rs885071			
	Additive	0.5209	0.776
	Recessive	0.6174	0.8496
	Dominant	NA	NA
	Genotype	0.7068	0.893
rs1800925			
	Additive	0.6468	0.1621
	Recessive	0.938	0.2271
	Dominant	NA	NA
	Genotype	0.7014	0.2991
rs2337506			
	Additive	0.5814	0.6825
	Recessive	0.8124	0.6267
	Dominant	0.448	0.8346
	Genotype	0.686	0.8864
rs4779565			
	Additive	0.3959	0.8642
	Recessive	0.4022	0.9115
	Dominant	0.5928	0.8824
	Genotype	0.6555	0.9838
rs4779978			
	Additive	0.0314	0.3307
	Recessive	0.0037	0.2689
	Dominant	0.4865	0.6001
	Genotype	0.0127	0.5371
rs6494223			
	Additive	0.6342	0.5938
	Recessive	0.7717	0.5131

	Dominant	0.4526	0.7031
	Genotype	0.7401	0.7823
rs7167551			
	Additive	0.8761	0.8983
	Recessive	NA	NA
	Dominant	0.9237	0.9605
	Genotype	0.9237	NA
rs7178176			
	Additive	0.315	0.5947
	Recessive	0.2194	0.3625
	Dominant	0.6909	0.6909
	Genotype	0.5443	0.4092
rs11071593			
	Additive	0.1642	0.3701
	Recessive	0.279	0.5007
	Dominant	NA	NA
	Genotype	0.8752	0.9793
rs11636680			
	Additive	0.4023	0.7161
	Recessive	0.5847	0.7221
	Dominant	0.3282	0.9383
	Genotype	0.592	0.9437
rs12910885			
	Additive	0.5814	0.6825
	Recessive	0.8124	0.6267
	Dominant	0.448	0.8346
	Genotype	0.686	0.8864
rs12916879			
	Additive	0.1374	0.325
	Recessive	0.4427	0.107
	Dominant	0.0846	0.8222
	Genotype	0.2175	0.2319
rs13329490			
	Additive	0.5061	0.5103
	Recessive	0.3321	0.4394
	Dominant	NA	NA
	Genotype	0.2391	0.4007
rs60109258			
	Additive	0.4789	0.5439
	Recessive	0.7532	0.5223
	Dominant	0.166	0.5428

	Genotype	0.3761	0.7092
rs61750900			
	Additive	0.0251	0.2589
	Recessive	0.0351	0.4407
	Dominant	NA	NA
	Genotype	0.0435	0.6313
rs62003625			
	Additive	0.3928	0.8655
	Recessive	0.4316	0.8978
	Dominant	NA	NA
	Genotype	0.5468	0.9363
rs113629315			
	Genotype	0.6482	0.6234

Table A-2: Interaction between selected genetic variants, radiographic progression (change in mSASSS/time) and both smoking status and sex. Three inheritance models are described as Dominant (CC or CT), Recessive (TT), and Additive

SNP		**Interaction between selected genetic variants, radiographic progression (change in mSASSS/time) and smoking status**	**Interaction between selected genetic variants, radiographic progression (Change in mSASSS/time) and sex**
rs885071			
	Additive	0.7068	0.6175
	Recessive	0.3977	0.4735
	Dominant	NA	NA
	Genotype	0.5936	0.3719
rs1800925			
	Additive	0.0261	0.6428
	Recessive	0.0607	0.8631
	Dominant	NA	NA
	Genotype	0.1493	0.9572
rs2337506			

	Additive	0.215	0.4385
	Recessive	0.4656	0.7092
	Dominant	0.2896	0.22
	Genotype	0.5652	0.2218
rs4779565			
	Additive	0.6439	0.964
	Recessive	0.6749	0.9882
	Dominant	0.8315	0.9854
	Genotype	0.9462	0.9991
rs4779978			
	Additive	0.1182	0.4517
	Recessive	0.1831	0.6298
	Dominant	0.2218	0.5369
	Genotype	0.3521	0.7992
rs6494223			
	Additive	0.4074	0.7013
	Recessive	0.6028	0.9734
	Dominant	0.5781	0.7205
	Genotype	0.8224	0.9382
rs7167551			
	Additive	0.5866	0.6192
	Recessive	NA	NA
	Dominant	0.5633	0.376
	Genotype	0.4022	NA
rs7178176			
	Additive	0.1812	0.6501
	Recessive	0.2541	0.9267
	Dominant	0.1918	0.1918
	Genotype	0.3499	0.33
rs11071593			
	Additive	0.1028	0.9347
	Recessive	0.0591	0.9042
	Dominant	NA	NA
	Genotype	0.1144	0.8648
rs11636680			
	Additive	0.247	0.7253
	Recessive	0.597	0.9399
	Dominant	0.0672	0.3007
	Genotype	0.184	0.5848
rs12910885			
	Additive	0.215	0.4385

	Recessive	0.4656	0.7092
	Dominant	0.2896	0.22
	Genotype	0.5652	0.2218
rs12916879			
	Additive	0.1258	0.7632
	Recessive	0.2246	0.6736
	Dominant	0.1673	0.7419
	Genotype	0.2824	0.8
rs13329490			
	Additive	0.3508	0.8099
	Recessive	0.448	0.9149
	Dominant	NA	NA
	Genotype	0.5582	0.9888
rs60109258			
	Additive	0.3083	0.6357
	Recessive	0.63	0.962
	Dominant	0.3775	0.6365
	Genotype	0.6555	0.9042
rs61750900			
	Additive	0.44	0.4582
	Recessive	0.5685	0.6227
	Dominant	NA	NA
	Genotype	0.662	0.7772
rs62003625			
	Additive	0.5899	0.8054
	Recessive	0.6792	0.7672
	Dominant	NA	NA
	Genotype	0.7252	0.7536
rs113629315			
	Genotype	0.1655	0.8443

Table A-3: Three inheritance models (additive, dominant and recessive) examined to determinate the association between each SNP and the mean values of mSASS1 and progression, respectively, in multiple linear regression analysis controlling for age, gender and smoking habit.

SNP		mSASSS1	progression
rs885071			
	Additive	0.9861	0.3039
	Recessive	0.8341	0.5773
	Dominant	0.4204	0.0219
	Genotype	0.6801	0.071
rs1800925			
	Additive	0.8464	0.7776
	Recessive	0.7993	0.4084
	Dominant	0.1621	0.142
	Genotype	0.2936	0.1384
rs2337506			
	Additive	0.4204	0.0205
	Recessive	0.0868	<0.0001
	Dominant	0.9146	0.3932
	Genotype	0.2068	0.0003
rs4779565			
	Additive	0.5638	0.9576
	Recessive	0.6333	0.7391
	Dominant	0.6226	0.6138
	Genotype	0.8455	0.7416
rs4779978			
	Additive	0.2668	0.074
	Recessive	0.2008	0.3457
	Dominant	0.6232	0.034
	Genotype	0.441	0.1038
rs6494223			
	Additive	0.4618	0.2239
	Recessive	0.3078	0.089
	Dominant	0.9161	0.8714
	Genotype	0.5751	0.2132
rs7167551			
	Additive	0.5484	0.1544

	Recessive	0.8022	0.2164
	Dominant	0.5408	0.222
	Genotype	0.8284	0.3206
rs7178176			
	Additive	0.0013	0.0738
	Recessive	0.0212	0.125
	Dominant	0.0005	0.1527
	Genotype	0.001	0.1913
rs11071593			
	Additive	0.0319	0.0465
	Recessive	0.1717	0.0878
	Dominant	0.0023	0.099
	Genotype	0.0087	0.1155
rs11636680			
	Additive	0.7277	0.2094
	Recessive	0.6034	0.7556
	Dominant	0.9245	0.0258
	Genotype	0.8427	0.0771
rs12910885			
	Additive	0.4204	0.0205
	Recessive	0.0868	<0.001
	Dominant	0.9146	0.3932
	Genotype	0.2068	0.0003
rs12916879			
	Additive	0.9904	0.2643
	Recessive	0.7956	0.0407
	Dominant	0.7829	0.717
	Genotype	0.8982	0.0648
rs13329490			
	Additive	0.7335	0.6568
	Recessive	0.6272	0.4966
	Dominant	0.8646	0.7415
	Genotype	0.8343	0.6645
rs60109258			
	Additive	0.545	0.1735
	Recessive	0.2875	0.0782
	Dominant	0.8309	0.7554
	Genotype	0.4716	0.2028
rs61750900			
	Additive	0.1779	0.3658
	Recessive	0.1142	0.2883
	Dominant	0.9384	0.9514

	Genotype	0.2351	0.5182
rs62003625			
	Additive	0.3441	0.4872
	Recessive	0.4674	0.4199
	Dominant	0.2718	0.8933
	Genotype	0.4859	0.6831
rs113629315			
	Genotype	0.2834	0.7398