



Trinity College Dublin

Coláiste na Tríonóide, Baile Átha Cliath

The University of Dublin

Genotypic and phenotypic insights into an Irish axial spondyloarthritis cohort

Marcus Kenyon

A thesis submitted to the University of Dublin for
the degree of Doctor of Philosophy

School of Medicine

Supervisor: Prof. Ross McManus

Trinity College Dublin

2023

Declarations

I hereby declare that this thesis has not been submitted as an exercise for a degree at this or any other university and it is entirely my own work. I agree to deposit this thesis in the University's open access institutional repository or allow the Library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement. I consent to the examiner retaining a copy of the thesis beyond the examining period, should they so wish (EU GDPR May 2018).

This work was undertaken by Marcus Kenyon within the Department of Clinical Medicine Genomics Research Group, Trinity Translational Medicine Institute, Trinity College Dublin. Any work performed by individuals other than the author is clearly credited throughout this thesis.

Full and informed consent was obtained from all study participants to include the processing of clinical and genetic information. The study adhered to good clinical practice guidelines at all participating centres. Ethical approval was obtained from each of the local hospital ethics committees at participating centres, as well as from the Joint Research Ethics Committee for St. James' and Tallaght University Hospital.

Signed:  Date: 27/09/2023
Marcus Kenyon

Summary

Axial spondyloarthritis (AxSpA) encompasses a group of closely related chronic inflammatory conditions which primarily affect the lumbar spine. The prototypic presentation of the disease is ankylosing spondylitis, which comprises more than half of AxSpA cases. AxSpA presentation is highly heterogeneous, with a wide range of axial disease severity and concomitant peripheral manifestations, such as uveitis, peripheral arthritis, psoriasis and IBD.

AxSpA is a highly heritable disease, with strong generational recurrence and robust evidence of genetic association. The strongest of these is the MHC class I type, HLA-B27, which was first reported in 1973 and is present in over 80% of AxSpA patients. 50 years since this discovery, over 100 independent genetic loci have been linked with the disease. Genetic risk points to the involvement of aberrant antigen presentation by HLA-B27 and dysregulation of the IL-17 pathway as key factors in AxSpA aetiology. However, only a small proportion of genetically susceptible individuals ever develop the disease. The specific event which triggers onset of the disease in genetically susceptible individuals remains elusive, but is generally considered to be a ubiquitous environmental influence. However, a significant proportion of AxSpA patients present with sub-clinical gut inflammation and gut dysbiosis has also been reported. This may point to an altered relationship between the immune system and gut microbiome in disease susceptible individuals.

Here, we provide a comprehensive account of the clinical presentation of AxSpA within a unique Irish registry cohort while investigating the role of genetic traits in shaping disease characteristics. Our study employed a range of genetic techniques, including genome-wide and gene-specific approaches, to identify key genetic variants associated with disease onset and severity. Additionally, we utilised polygenic risk scores to explore how composite genetic risk, both within and beyond the HLA complex, might influence disease presentation. Furthermore, I performed a meta-analysis of genetic risk, incorporating findings from our study cohort with a previous genome-wide association study to identify novel risk variants. Lastly, we investigated how the regulation of specific genes and bio-molecular pathways differ between AxSpA patients and healthy controls using single-cell RNA sequencing.

We found that some peripheral manifestations associated with AxSpA occur at a steady rate over time, while others appear to either precede or occur early in AxSpA disease. In particular, dactylitis seemed to occur early in AxSpA disease and was strongly associated with other peripheral manifestations, especially uveitis and was also inversely associated with HLA-B27 and spinal radiographic involvement. Our results suggest that patients presenting with dactylitis represent a distinct subset of AxSpA. Disease presentation was also found to be influenced by

genetic risk variants. We revealed that multiple genetic variants within the ERAP1 gene were associated with the age at which AxSpA disease developed and the likelihood of concomitant uveitis episodes within patients. This was supported by polygenic risk scores, which demonstrated that genetic variants, both within and outside of the HLA region are associated with earlier AxSpA disease onset. We found further evidence for lymphocyte involvement through the identification of several disease risk variants located near to genes relevant to the differentiation and function of these cells. This was supported using single cell RNA expression analysis, which revealed increased expression of IL-17 pathway genes in multiple cell subtypes and a strong signature of immune cell activation in patient relative to control samples. Several genes associated genetically with AxSpA were found to be differentially regulated in immune cells. Finally, molecular pathways differentially expressed between AxSpA patients and healthy controls support the hypothesis that gut dysbiosis is involved in AxSpA pathogenesis.

Acknowledgements

I would like to express my sincere thanks to everyone who very generously gave their time to support me over the past number of years.

In particular, I would like to thank my supervisor, Professor Ross McManus. It is through your knowledge and enthusiasm that I have developed a deep interest in the field of genetics and genomics. You have pushed me far beyond my comfort zone into cutting edge techniques and helped me to understand how to tackle the most complex problems. Thanks to your wise insight and guidance I finally have the tools to emerge into the world of research on my own two feet. Thank you for your incredible patience over the past number of years, particularly through the most difficult times. I am very much looking forward to continuing our research work on the single cell sequencing analysis over the coming months.

I would also like to thank Dr. Deepak Bharti. Firstly, for all of your help in translating our single cell data, but most of all for what I have learned from you. If it weren't for your teaching, I would still be typing BASH scripts on the terminal interface. You have given me inspiration to expand my knowledge of data analytics and bioinformatics.

To Dr. Shigeki Nakagome. Thank you so much for your truly unique approaches and insights into some of the problems we have come across. I have also secretly been learning about population genomics whenever you have been discussing work with your PhD students in the office.

To Dr. Elizabeth Heron. Thank you for your wisdom in statistics. Thank you for debating techniques with me and offering guidance whenever I needed it. You have challenged my thinking and helped me to consider problems from different angles.

To Dr. Conor Finlay and the TCD Omics team. Thank you for your absolutely incredible work in the single cell pipeline. The work you have performed has opened a remarkable array of research opportunities for us.

To Ric Anney and Lesley Bates in the MRC Institute for Molecular Genetics and Genomics. Thank you for your steadfast support in scanning the seemingly endless supply of prepared gene chips we sent your way.

Thank you to Dr. Finbar O'Shea, Phil Gallagher and all of the ASRI team of nurses and doctors who have diligently recruited patients, and collected the data and samples which make this type of research possible.

Related publications

Kenyon, M., Maguire, S., Rueda Pujol, A., O'shea, F. & Mcmanus, R. 2022. The genetic backbone of ankylosing spondylitis: how knowledge of genetic susceptibility informs our understanding and management of disease. *Rheumatology International*.

Manuscripts in preparation for submission:

Kenyon, M. Gallagher, P., Dinneen, B., O'Shea, F., & McManus, R. Peripheral arthritis and dactylitis are closely related in AxSpA but are independently associated with worse patient outcomes.

Kenyon, M., Bharti, D., Gallagher, P., Dinneen, B., Bates, L., Anney, R., O'Shea, F. & McManus, R. Meta-analysis of GWAS reveals genetic association in the regions of CD80, EOMES and 6 additional loci in axial spondyloarthritis.

List of abbreviations

AKT - Protein kinase B

Ag - Antigen

AS - Ankylosing spondylitis

ASAS - Assessment of SpondyloArthritis international Society

AAU - Acute anterior arthritis

ASRI – Ankylosing Spondylitis Registry of Ireland

AxSpA - Axial spondyloarthritis

CARD9 - Caspase associated recruitment domain 9

CD - Crohn's disease

CREB - cAMP-response element binding protein

D' - D-prime

DMARD - Disease modifying anti-rheumatic drug

EAM - Extra-articular manifestation

EP4 - Prostanoid EP4 receptor

ERAP1 - Endoplasmic reticulum aminopeptidase 1

ER - Endoplasmic reticulum

Fc - Fragment crystallisable

FDA - U.S. Food & Drug Administration

FDR - False discovery rate

FUMA GWAS - Functional Mapping and Annotation of Genome-Wide Association Studies

GWAS - Genome wide association study

GRCh37 (hg19) - Genome Reference Consortium build 37

HAQ-s - Health Assessment Questionnaire for the Spondyloarthropathies

HLA - Human leukocyte antigen

HWE - Hardy-Weinburg equilibrium

IFN- λ - Interferon gamma

IgG - Immunoglobulin G

IL-17 - Interleukin 17

IL-23 - Interleukin 23

IL23R - Interleukin 23 receptor gene

Indel - Insertion or deletion

IQR - Interquartile range

JAK2 - Janus kinase 2

JAKinib - Janus kinase inhibitor

KS - Kolmogorov-Smirnov

LD - Linkage disequilibrium

lncRNA – Long non-coding RNA

LPS - Lipopolysaccharide

MAF - Minor allele frequency

MAGMA - Multi-marker Analysis of GenoMic Annotation

MAMP - Microbe associated molecular pattern

MHC Class I - Major histocompatibility complex class I

mNY - Modified New York Criteria for ankylosing spondylitis, 1984

MS - Multiple sclerosis

OR - Odds ratio

PBWT - Positional Burrow-Wheeler transformation

ped - Pedigree file format

PI_HAT - Kinship coefficient (ϕ)

PI3K - Phosphoinositide 3-kinase

PKA - Protein kinase A

PRO – Patient reported outcome

Ps - Psoriasis

PsA - Psoriatic arthritis

QC - Quality control

R² - R-squared

rAxSpA - Radiographic axial spondyloarthritis

SLE - Systemic lupus erythematosus

STAT3 - Signal transduction and transcription 3

Th17 - T helper 17

TNF- α - Tumour necrosis factor alpha

TNFi - Tumour necrosis factor alpha inhibitor

TSS - Transcription start site

TYK2 - Tyrosine kinase 2

UC - Ulcerative colitis

UTR - Untranslated region

vcf - Variant call file format

wnt - Wingless pathway

Table of Contents

Declarations	ii
Summary	iii
Acknowledgements	v
Related publications	vi
List of abbreviations	vii
Index of tables	xiii
Index of figures	xvi
1.1 An introduction to axial spondyloarthritis	2
1.2 Diagnosis of AS	3
1.3 Epidemiology	6
1.4 Pathogenesis	7
1.5 Clinical Indicators of Disease Status	8
1.6 Serum Biomarkers in AS	10
1.6.1 Inflammatory Markers	10
1.6.2 The Adaptive Immune Response	10
1.6.3 Bone Metabolic Markers	11
1.7 The Genetic Nature of AS	11
1.7.1 The Role of HLA-B27 in AS Pathology	12
1.7.2 The Discovery of Further Genetic Associations through GWAS	13
1.7.2 The significance of the <i>ERAP1</i> locus	14
1.7.3 Genetic implication of T-cell function	15
1.7.4 Genetic associations found within ‘gene desert’ regions	20
1.7.5 The use of fine mapping techniques to establish causal variants	20
1.8 The relationship between genetics and clinical phenotype	22
1.8.1 Genetic commonality in disease comorbidity	22
1.8.2 How genetics impact on disease severity	24
1.8.3 The polygenic risk approach to predicting disease onset and severity	25
1.8.4 Genetics play a role in therapeutic selection in AS	26
1.9 Summary	27
1.10 Rationale	29
1.10.1 Objectives	30
2 A clinical appraisal of the ASRI cohort	32
2.1 Introduction	32
2.1.1 AS and AxSpA: increasingly blurred definitions	32
2.1.2 Measuring disease outcomes in AxSpA	33

2.1.2 Differences in presentation between males and females.....	34
2.1.2 The effects of smoking and obesity on AxSpA disease outcomes.....	34
2.2 Aims.....	35
2.3 Methods.....	36
2.3.1 Data collection	36
2.3.2 Statistical analysis	37
2.4 Results	38
2.4.1 Demographics of the ASRI cohort.....	38
2.4.2 The effect of disease duration on severity and extra-articular manifestations...39	
2.4.3 Gender differences in presentation	44
2.4.4 The effects of extra-axial manifestations on AxSpA disease severity	50
2.4.5 Investigating potential outcomes of obesity in past smokers.....	58
2.5 Discussion.....	62
2.5.1 The relationship between disease duration and patient outcomes	62
2.5.2 AxSpA disease in men and women	63
2.5.3 The importance of EAMs in patient outcomes	64
2.5.4 The effects of past smoking on obesity and comorbidity in AxSpA patients	65
3 The relationship between genetics and phenotypic characteristics in the ASRI AxSpA cohort.	68
3.1 Introduction.....	68
3.2 Aims.....	70
3.3 Methods.....	71
3.3.1 Data collection	71
3.3.2 Genotyping.....	71
3.3.3 Hardy-Weinberg disequilibrium	73
3.3.4 Calculation of threshold for GWAS significance	75
3.3.5 Imputation of variants from whole genome sequence data.....	76
3.3.6 Conversion of IGAS variant ID's to match ASRI.....	77
3.3.7 Statistical evaluation of genotyping data	77
3.3.8 HLA-B27 typing and Immunochip genotyping validation.....	78
3.3.9 HLA-typing of patients from genotypic data.....	82
3.3.10 Interrogation of <i>ERAP1</i> haplotypes	83
3.3.11 Calculation of polygenic risk scores (PRS)	85
3.4 Results: Genome wide risk in AxSpA phenotypes	87
3.4.1 Demographic characteristics of genotyped cohort.....	87
3.4.2 Genome wide association analysis of phenotypic presentation	87

3.4.3 Acute anterior uveitis	88
3.4.4 Psoriasis	90
3.4.5 Inflammatory bowel disease	95
3.4.6 Early disease onset	96
3.5 Results: The role of the ERAP1 gene in clinical phenotype	99
3.5.1 Phenotypic characteristics associated with HLA-B27 in AxSpA	100
3.5.2 Phenotypic characteristics associated with ERAP1 variant rs30187	102
3.6 Results: Investigation of causative haplotypes within the ERAP1 gene	106
3.6.1 Investigation of rs30187 as a causative SNP in AxSpA	108
3.6.2: Phenotypic characteristics associated with rs149544 and rs7063	116
3.6.3: Investigation of the <i>ERAP2</i> gene as a possible independent AxSpA association	119
3.7 Results: The relationship between polygenic risk and clinical phenotype	120
3.8 Discussion	130
3.8.1 Genome-wide association analysis of clinical phenotypes	130
3.8.2 The relationship between HLA-B27, ERAP1 and disease presentation	131
3.8.4 The contribution of polygenic risk to disease phenotype and severity	133
3.9 Limitations of this study	135
4. A meta-analysis of genetic risk in AxSpA	137
4.1 Introduction	137
4.2 Aims	138
4.3 Methods	139
4.3.1 Acquisition of control samples and quality control	139
4.3.2 Genome wide association analysis of ASRI data	139
4.3.3 Meta-analysis of GWAS studies	140
4.3.4 Quality control: assessment of genetic association directionality between IGAS and ASRI	141
4.3.5 Determination of P-value cut off for significance of markers	141
4.3.6 Evaluation of meta-analysis results	141
4.3.7 Calculation of LD between novel and previously reported variants	142
4.4 Results	143
4.4.1 Investigation of meta-analysis data at genome wide significance	143
4.4.2 Investigation of further variants of interest	150
4.5 Discussion	159
4.6 Limitations of this study	162
5. Analysis of peripheral blood mononuclear cell gene expression profiles using a single cell quantitative RNA sequencing approach.	164

5.1 Introduction.....	164
5.2 Aims.....	166
5.3 Methods.....	166
5.3.1 Patient recruitment.....	166
5.3.2 Patient recruitment and data collection.....	166
5.3.3 PBMC isolation and lymphocyte preparation.....	167
5.3.4 Single cell isolation and sequencing.....	168
5.3.4 Single cell sequencing data processing.....	169
5.3.5 Differential expression and geneset enrichment analysis.....	169
5.4 Results.....	171
5.4.1 Annotation of <i>ex-vivo</i> cell subsets.....	175
5.4.2 Differential expression analysis in <i>ex-vivo</i> cells of AxSpA patients and healthy controls.....	180
5.4.3 Molecular pathway enrichment analysis by cell type.....	186
5.4.4 Investigation of cells stimulated with CD3 & CD28.....	189
5.5 Discussion.....	190
5.6 Limitations of this study.....	194
6. Conclusion.....	196
6.1 The ASRI AxSpA Cohort is similar to, but not the same as other AxSpA cohorts.....	196
6.2 Dactylitis is an outlier amongst peripheral manifestations.....	197
6.3 HLA-B27 and <i>ERAP1</i> contribute to disease phenotype.....	198
6.4 HLA-B27 negative AxSpA remains unresolved.....	199
6.5 Possible involvement of infectious disease and the gut microbiome in AxSpA aetiology.....	199
6.6 Evidence for Th17 cells in AxSpA continues to mount, but the relationship is complicated.....	200
6.7 Future directions.....	202
Supplementary material.....	207
Chapter 1 supplementary.....	208
Chapter 3 supplementary.....	209
Chapter 4 supplementary.....	214
Chapter 5 supplementary.....	216
Bibliography.....	221

Index of tables

Table 1.1	Rome Diagnostic Criteria for Ankylosing Spondylitis.	4
Table 1.2	Modified New York Criteria for diagnosis of ankylosing spondylitis.	5
Table 1.3	ASAS criteria for axial SpA.	6
Table 1.4	Functional overview of susceptibility genes shared between AS, psoriasis, ulcerative colitis and Crohn's.	24
Table 2.1	General characteristics of the ASRI cohort.	39
Table 2.2	Univariate and multivariate regression of clinical outcomes against AxSpA disease duration.	42
Table 2.3	Univariate and multivariate regression of extra-spinal manifestations against AxSpA disease duration.	44
Table 2.4	and female subgroups.	46
Table 2.5	A comparison of women with and without history of acute anterior uveitis.	49
Table 2.6	Characteristics of AxSpA patients with and without concomitant uveitis.	51
Table 2.7	Characteristics of AxSpA patients with and without concomitant psoriasis.	53
Table 2.8	Characteristics of AxSpA patients with and without concomitant arthritis.	55
Table 2.9	Characteristics of AxSpA patients with and without concomitant dactylitis.	57
Table 2.10	Demographics and clinical presentation of patients based on smoking status.	60
Table 2.11	Multivariate regression of clinical presentation by smoking status.	61
Table 3.1	Sample QC summary.	73
Table 3.2	Counts of HWE divergence by chromosome.	74
Table 3.3	Statistical power to detect associations by allele frequency.	76
Table 3.4	Cross-reference table of the concordance (%) of HLA-B27 tagging variants and methodologies.	81
Table 3.5	HIBAG Report Summary.	82
Table 3.6	Demographic characteristics of genotyped patients.	87
Table 3.7	A brief explanation of phenotypes used in GWAS analysis.	88
Table 3.8	A comparison of rs10484554-T positivity between ASRI patients with and without current psoriasis and cases of the Strange et al psoriasis GWAS.	92
Table 3.9	Psoriasis HLA-B27 and HLA-Cw6 summary.	93
Table 3.10	HLA-Cw6 carriage in AxSpA patients and Irish controls.	94

Table 3.11 General characteristics of participants genotyped for <i>ERAP1</i> study.	99
Table 3.12 AxSpA patients.	101
Univariate and multivariate analysis of rs30187 allelotypes in HLA-	
Table 3.13 B27+ patients.	104
Association of haplotypes containing rs30187 to AxSpA, based on	
Table 3.14 641 patients and 1,911 healthy controls.	109
Comparison of haplotype frequencies between 641 AxSpA patients	
Table 3.15 and 1,911 healthy controls, based on rs30187 allelotype.	109
Table 3.16 Novel and previously published tag SNPs associated with AxSpA.	112
Statistical significance of AxSpA disease associated <i>ERAP1</i> tag SNPs	
Table 3.17 after conditioning.	114
Residual statistical significance of <i>ERAP1</i> variants rs26653 and	
Table 3.18 rs149544 after conditioning.	115
Statistical significance of AxSpA disease protective <i>ERAP1</i> tag SNPs	
Table 3.19 after conditioning.	115
Univariate and multivariate analysis of patient characteristics	
Table 3.20 based on rs149544 allelotypes in HLA-B27+ patients.	118
Univariate and multivariate analysis between rs7063 allelotypes in	
Table 3.21 HLA-B27+ patients.	118
Table 3.22 SNP associations in the <i>ERAP2</i> coding region (Haplotype Block 11)	119
Table 3.23 Polygenic risk associated to patient demographics.	121
Polygenic risk excluding the HLA region and <i>ERAP1</i> associated with	
Table 3.24 HLA-B27.	122
Polygenic risk associated with disease activity, disability and gainful	
Table 3.25 employment.	125
Table 3.26 Polygenic risk associated to extra-axial manifestations.	127
Frequency (%) of high impact SNPs in the HLA region stratified by	
Table 3.27 HLA-B27 status and history of uveitis.	128
A post-hoc analysis of polygenic risk associated with gender in	
Table 3.28 AxSpA patients.	129
Table 4.1 Concordance of directionality of association stratified by P-value.	144
Table 4.2 Annotation of novel peak lead SNPs.	145
Association analysis of rs57271503 and rs28653390 in 641 AxSpA	
Table 4.3 patients and 1,911 Irish controls.	149
AxSpA associations of expanded haplotypes at the CD80 locus	
Table 4.4 between rs57271503 and rs28653390.	150
Additional novel variants of significance after Bonferroni	
Table 4.5 correction for multiple testing.	152
AxSpA association of simple haplotypes of rs1992661 and	
Table 4.6 rs73090828.	157
AxSpA association of derived expanded haplotypes between	
Table 4.7 rs1992661 and rs73090828.	157

Table 5.1	Patients included in the study.	171
Table 5.2	PBMC viability <i>ex-vivo</i> and cultured cell viability following 36 hour stimulation with CD3 and CD28	172
Table 5.3	Sequence read quality summary.	173
Table 5.4	Putative cells called by sample tag.	175
Table 5.5	Cell types determined by PCA of UMAP by sample <i>ex-vivo</i>	177
Table 5.6	Top up-regulated genes by statistical significance and cell type.	182
Table 5.7	Top down-regulated genes by statistical significance and cell type.	183
Table 5.8	Up-regulated and down-regulated genes in LD ($r^2 \geq 0.4$) of GWAS SNPs.	186
Table S1.1	Genetic concordance between associated diseases	205
Table S3.1	Haplotype associations within ERAP1 500Kb region, based on 1,000 permutations.	206
Table S3.2	Univariate and multivariate analysis between rs7063 allelotypes in HLA-B27- patients.	207
Table S4.1	LD between rs28653390 and both rs572715043 and rs6804441	212
Table S4.2	Investigation of further <i>IL1R2</i> SNPs.	212
Table S5.1	List of antibodies sourced from BD Biosciences used in single cell identification.	213
Table S5.2	Differential genes in monocytes.	215
Table S5.3	Differential genes in T cells.	216
Table S5.4	Differential genes in B cells and NK cells.	217

Index of figures

Fig. 1.1	Diagrammatic of the findings made by Bernard Connor in a French cemetery in 1691.	3
Fig. 1.2	The central role of STAT3 in the IL-17 producing T-cell pathway.	17
Fig. 1.3	PTGER4 receptor activation promotes the TH17 phenotype through at least two distinct pathways.	19
Fig. 1.4	An example of a haplotype block around the SNP rs9901869 on chromosome 17 (chr17:47497740, GRCh38).	21
Fig. 2.1	Univariate regression analysis of patient reported outcomes by AxSpA disease duration.	41
Fig. 2.2	Univariate binary logistic regression of extra-axial manifestations by AxSpA disease duration.	43
Fig. 2.3	A model of extra-musculoskeletal manifestations with AxSpA disease duration in males and females.	48
Fig. 3.1	Cluster plots of rs30187 and rs2726864 generated by Illumina GenomeStudio V2011.1.	73
Fig. 3.2	Cluster plots of variants with acceptable genotype clustering failing HWE generated by Illumina GenomeStudio V2011.1.	75
Fig. 3.3	Manhattan plot of GWAS of 480 individuals (442 B27+, 38 B27-).	78
Fig. 3.4	Linkage disequilibrium within 50Kb radius of rs7771432.	79
Fig. 3.5	The relationship between AAU and AxSpA disease duration.	88
Fig. 3.6	Genome wide significance of AAU in the ASRI patient cohort, based on 225 patients with and 405 without AAU.	89
Fig. 3.7	QQ plot of genetic significance associated with AAU in AxSpA patients.	90
Fig. 3.8	The relationship between psoriasis and AxSpA disease duration.	91
Fig. 3.9	Genome wide significance of psoriasis in the ASRI patient cohort, based on 107 patients with and 524 patients without psoriasis.	91
Fig. 3.10	QQ plot of the genetic significance associated with psoriasis in AxSpA patients.	92
Fig. 3.11	The relationship between IBD and AxSpA disease duration.	95
Fig. 3.12	Genome wide significance of IBD in the ASRI patient cohort, based on 66 patients with and 566 without IBD.	96
Fig. 3.13	QQ plot of the genetic significance associated with psoriasis in AxSpA patients.	96
Fig. 3.14	Genome wide significance of early disease onset (<21 years) in the ASRI patient cohort, based on 238 patients with early onset and 403 with later onset.	97
Fig. 3.15	QQ plot of the genetic significance associated with early disease onset (<21 years) in AxSpA patients.	98
Fig. 3.16	Distribution of patient age and AxSpA disease duration in AxSpA patients with a history of AAU.	102

Fig. 3.17	Distribution of age and AxSpA disease duration subdivided by ERAP1 rs30187 genotype.	105
Fig. 3.18	LD structure within the ERAP1-LNPEP region extending from Chr5: 96,000,000 – 96,500,000, (GRCh37, hg19).	107
Fig. 3.19	Haplotype blocks within ERAP1 500KB region.	107
Fig. 3.20	Tag variants (6) within the ERAP1 coding region generated by clumping methodology with $r^2 < 0.8$.	111
Fig. 3.21	Locus Zoom of r149544 region indicates AxSpA association and partial genetic correlation with rs30187.	113
Fig. 3.22	Statistical significance of PRS in relation to age at first symptom onset and diagnosis at various PRS thresholds.	122
Fig. 3.23	Relationship between PRS taken at a threshold of $P < 0.0001$ in relation disease onset, diagnosis and HLA-B27 status.	123
Fig. 4.1	Manhattan plot of genetic association to AS in the ASRI cohort, based on 69,636 SNP's in 641 patients and 1,911 controls.	143
Fig. 4.2	QQ plot of statistical significance of genetic risk associated with AxSpA, based on 641 patients and 1,911 controls.	144
Fig. 4.3	Manhattan plot of genome wide significance, based on 69,636 SNPs in 11,260 patients and 17,056 controls.	145
Fig. 4.4	eQTL reports for rs11129295	147
Fig. 4.5	Tissue expression of the EOMES gene.	147
Fig. 4.6	Tissue expression of the LINC02084 gene.	148
Fig. 4.7	LD block analysis of 641 AxSpA patients and 1,911 Irish controls between rs57271503 and rs28653390.	150
Fig. 4.8	Manhattan plot of genome wide significance, based on 69,636 SNP's in 11,260 patients and 17,056 controls with P-value threshold for statistical significance corrected for 29,258 independent variants (1.71×10^{-6})	151
Fig. 4.9	Haplotypes of significant tag SNP's derived from 1000 genomes data for CEU and GBR populations.	154
Fig. 4.10	Haplotypes of variants near to IP6K1 previously associated to IBD together with novel AxSpA associated variant rs7374183.	155
Fig. 4.11	eQTL for lead SNP rs7374183 in whole blood affecting expression of UBA7 and RBM6.	155
Fig. 4.12	LD block analysis of 641 AxSpA patients and 1,911 Irish controls between rs73090828 and rs1992661.	156
Fig. 4.13	Effect of rs2371108 (chr3:27715527 GRCh38/hg38) on bulk tissue gene expression of LINC02084 in spleen and whole blood.	158
Fig. 5.1	Quality characteristics of RNA expression in cells <i>ex-vivo</i> .	174
Fig. 5.2	Ex-vivo cells from combined patients and controls, clustered by gene expression and annotated.	176
Fig. 5.3	Total cell counts by group.	177

Fig. 5.4	Cell proportions subdivided by group (AxSpA and control) and batch	178
Fig. 5.5	Gene expression profiles of the seven most populous cell clusters, and the aberrant T cells identified in ex-vivo cells.	179
Fig. 5.6	Differential expression of genes by annotated cell cluster.	184
Fig. 5.7	Molecular pathway enrichment by cell type.	188
Fig. S3.1	Genome-wide scan of variants associated with family history of AxSpA, based on 180 patients with and 461 patients without family history of AxSpA.	207
Fig. S3.2	QQ-plot of distribution of significance presented in Fig. S3.1.	208
Fig. S3.3	Genome-wide scan of significance of variants associated with gender assigned at birth in AxSpA, based on 497 male and 138 female patients.	209
Fig. S3.4	QQ-plot of distribution of significance presented in Fig. S3.3.	209
Fig. S3.5	Genome-wide scan of variants associated with modified New York criteria (diagnostic criteria for AS), based on 180 patients with and 461 patients without family history of AxSpA.	210
Fig. S3.6	QQ-plot of distribution of significance presented in Fig. S3.5. Fig. S3.6	210
Fig. S4.1	Manhattan plot reconstructed from IGAS summary data.	211
Fig. S4.2	Tissue expression of AC098614.2.	211
Fig. S5.1	Antibody binding in cell clusters.	214

Chapter 1

General Introduction

1.1 An introduction to axial spondyloarthritis

Axial spondyloarthritis (AxSpA) is a group of closely related seronegative autoimmune disorders which present primarily as chronic inflammation of the axial skeleton. Among these diseases, ankylosing spondylitis (AS) is the best described and most common presentation of AxSpA. Due to the significant overlap between AxSpA subtypes, this introductory chapter will initially focus on AS and expand to include AxSpA as a whole where appropriate.

The term “ankylosing spondylitis” originates from the two major processes of the disease: ‘ankylosis’ meaning fusion of bone tissues, and ‘spondylitis’ indicating vertebral inflammation. Arguably, the first conclusive account of the disease was made in 1691 by the Irish physician, Bernard Connor, when his interest was drawn to local reports of an unusual skeleton found in a nearby cemetery during his stay in France (Blumberg, 1958, Rudwaleit et al., 2009). In his account, Connor describes the anatomical detail: *“All these Bones, which Naturally are 38, each separate and distinct were here so straightly and intimately joyned, their Ligaments perfectly Bony and their Articulations so effaced, that they really made but one uniform continuous Bone”*. While the specimen described by Connor is no longer available for pathological verification, his remarkably concise description and depiction (Fig. 1.1) of the skeletal features resonates sharply with a modern medical textbook definition of untreated AS.

Typically, patients with AS present with lower vertebral or hip pain, spinal stiffness in the morning that is alleviated by exercise, and lethargy. The primary clinical finding in AS is axial inflammation with associated back pain and morning stiffness. Ectopic ossification in the form of syndesmophytes promotes progressive replacement of inter-vertebral cartilage tissue, resulting in restricted movement of the spinal column and eventual vertebral fusion (Romanus and Ydén, 1952, Ball, 1971). Associated non-axial rheumatic manifestations include peripheral arthritis and dactylitis (Song et al., 2007). In addition, AS is often associated with acute anterior uveitis (AAU), inflammatory bowel disease (IBD) and psoriasis (Song et al., 2007). It is worth noting that while these conditions are fully characterised and independent within their own right, they are often considered as extra-articular manifestations (EAMs) in the context of underlying spondyloarthropathy. More rarely, patients develop pathologies of cardiac or renal function (Lange et al., 2007).

AS is incurable, but symptoms can be treated using a variety of anti-inflammatory agents ranging from non-steroidal options to targeted disease-modifying anti-rheumatic drugs (DMARDs). The choice of therapy depends on disease severity, drug availability, and therapeutic responsiveness.

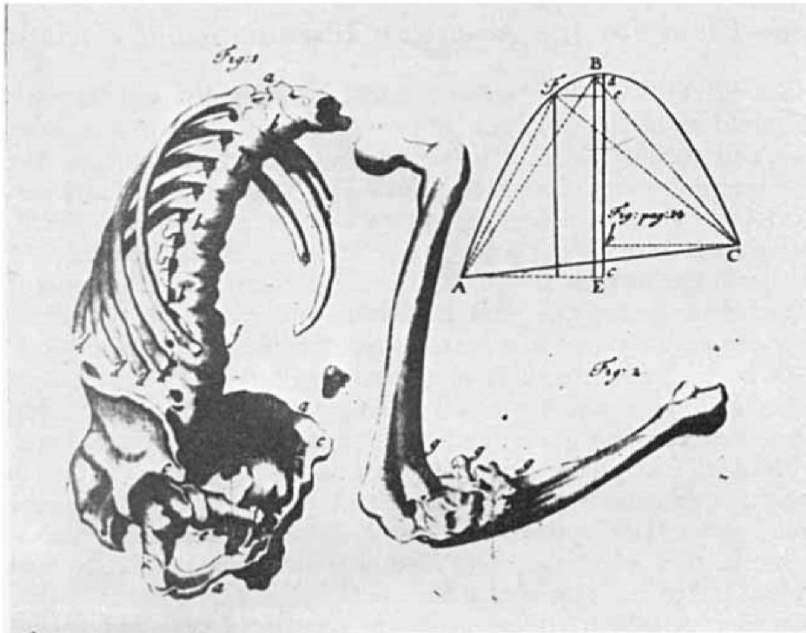


Fig. 1.1: Diagrammatic of the findings made by Bernard Connor in a French cemetery in 1691. His report included detailed physical examination of the spinal column and arthritic knee joint (probably the result of rheumatoid arthritis) from an unrelated skeleton. Image sourced from Blumberg (1958).

1.2 Diagnosis of AS

The provision of rapid and accurate diagnosis of AS has proved challenging due to substantial overlap with related diseases and a lack of clinical awareness of the condition. Non-specific early symptoms and the absence of unique biomarkers have led to the use of rigid diagnostic criteria, which require significant disease progression to satisfy. The first attempts to develop diagnostic criteria for AS were made in 1961, resulting in what became known as the Rome Criteria, shown in Table 1.1, below.

Rome Diagnostic Criteria for Ankylosing Spondylitis 1961

Clinical criteria:

1. Low back pain and stiffness of more than 3 months duration, not relieved by rest.
2. Pain and stiffness in the thoracic region.
3. Limited motion in the lumbar spine.
4. Limited chest expansion.
5. History of evidence of iritis or its sequelae.

Radiological criterion:

6. Radiograph showing bilateral sacroiliac changes characteristic of AS.

Table 1.1: Rome Diagnostic Criteria for Ankylosing Spondylitis. Using these criteria, diagnosis should be made where the radiological criterion together with one clinical criterion is met, or at least 5 clinical criteria are present in the absence of the radiological criterion (Gran and Husby, 1993).

The Rome criteria had the advantage of allowing diagnosis without radiographic evidence, but they were criticized for failing to correct lumbar spine mobility and chest expansion for age and sex. In 1966, the criteria were revised at a meeting in New York. The updated criteria required radiographic evidence of apical involvement, along with supporting mechanical evidence of restricted spinal or chest mobility. These became known as the New York Criteria and were widely used for diagnosing AS until the modified New York (mNY) criteria were introduced in 1984.

The mNY criteria require a combination of three clinical criteria, together with radiographic evidence, as outlined by van der Linden et al. in 1984. The criteria are listed in Table 1.2, below.

Modified New York Criteria 1984

Clinical criteria:

Lower back pain persisting greater than three months, improved with exercise, not relieved by rest.

Limitation of lumbar spine in sagittal and frontal planes.

Restricted chest expansion relative to expected normal for age and sex.

Radiographic criterion:

Bilateral grade 2 or higher sacroiliitis OR

Unilateral grade 3 or higher sacroiliitis

Table 1.2: Modified New York Criteria for diagnosis of ankylosing spondylitis. All three clinical criteria are required in combination with at least one radiographic criterion to make a diagnosis of AS (van der Linden et al., 1984).

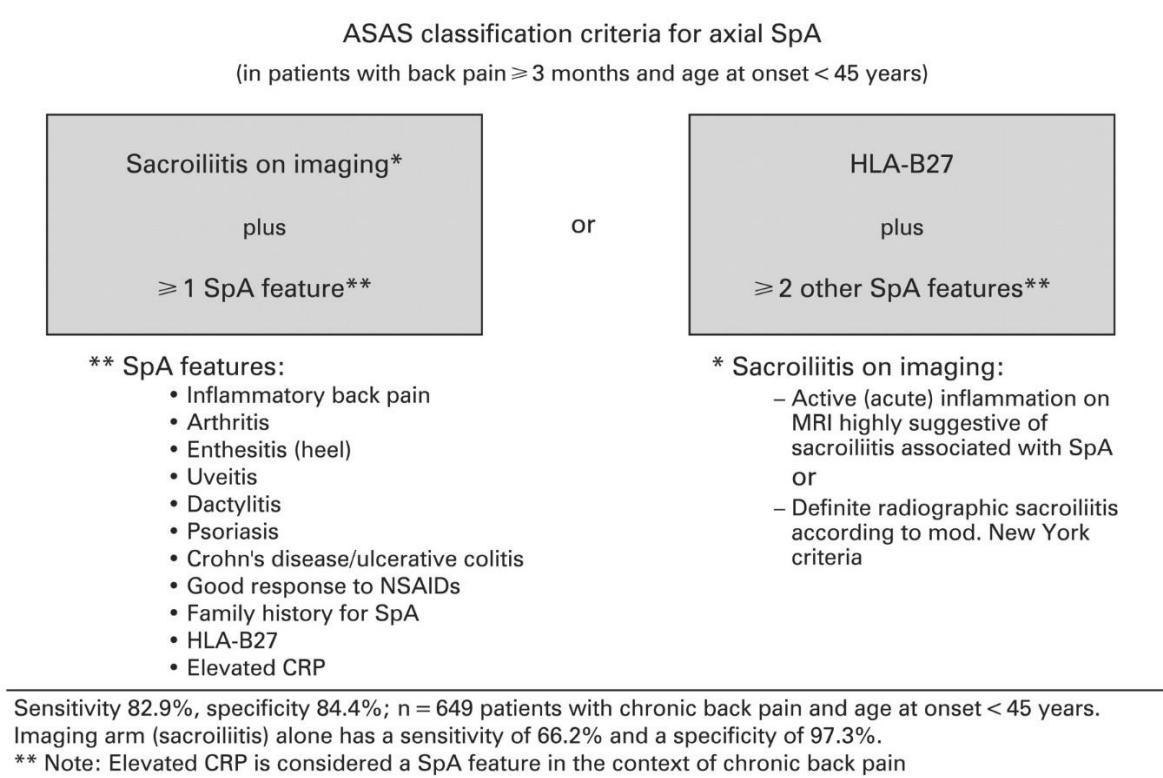
The Modified New York Criteria for the diagnosis of Ankylosing Spondylitis (AS) has high specificity due to the requirement of radiographic evidence of ankylosis and impaired spinal mobility. However, sensitivity is poor as radiographic criteria may take up to a decade to develop, causing significant delay in diagnosis (Mau et al., 1988). The introduction of biological DMARDs, such as TNF- α inhibitors, highlighted this limitation further since patients could not qualify for treatment until radiographic sacroiliitis had developed.

However, the Assessment of SpondyloArthritis International Society (ASAS) developed a classification system for Axial Spondyloarthritis (AxSpA) in 2009 (Rudwaleit et al., 2009), which addressed this issue. The ASAS criteria allow for the diagnosis of radiographic AxSpA in patients with chronic back pain if they have sacroiliitis on imaging and at least one SpA feature. Meanwhile, non-radiographic AxSpA (nr-AxSpA) can be diagnosed in patients with lower back pain where HLA-B27 and at least two additional SpA features are present (Rudwaleit et al., 2009). SpA features include inflammatory back pain, SpA-associated extra-spinal manifestations, response to NSAIDs, family history, elevated CRP or ESR, and HLA-B27. These criteria are shown in Table 1.3.

The ASAS criteria have not only enabled earlier treatment options for potential AS patients but also sparked debate over the classification of this group of diseases. ASAS has classified a relatively heterogeneous group of diseases as nr-AxSpA, which may ultimately progress to classical AS (Robinson et al., 2013). Studies have shown differences between nr-AxSpA and AS

cohorts in patient demographics and HLA-B27 positivity rates (Baraliakos and Braun, 2015, Chimenti et al., 2020). It is currently unclear how many patients diagnosed with nr-AxSpA will eventually progress to AS. However, recent studies suggest progression to AS at a rate of approximately 18% per decade (Wang et al., 2016, Aydin et al., 2012).

While the ASAS criteria have flaws in disease classification, it is the best model for early intervention with modern therapeutic options. The close relationship between these diseases raises questions about how each disease state should be defined and whether each should be considered a separate entity or a spectrum of autoimmune disease of related etiology. These questions are complex and will likely be the subject of much medical debate.



Africa (0.07%) (Dean et al., 2014). The disease has a significant male preponderance of approximately 3.4:1 (Dean et al., 2014). In addition, AS prevalence tends to correlate broadly with the frequency of HLA-B27, worldwide (Khan, 1995).

Environmental factors that affect AS incidence have not been well described. Several studies on the Swedish population have attempted to unravel some of these underlying factors. One such study found an inverse association between AS prevalence and education, suggesting a socioeconomic influence (Exarchou et al., 2015). Another study found that having older siblings increased the risk of AS, while maternal age, type of birth (caesarean/vaginal), birth weight, and smoking status of the mother did not show any association (Lindström et al., 2016). A third study suggested a weak link between respiratory tract infections and AS, although multiple testing was an issue. In France, breast-feeding was found to have a protective association with AS, but the study did not control for socioeconomic status (Montoya et al., 2016).

An association between AS and the pathogenic bacteria, *Klebsiella pneumoniae* has been reported by several studies (Ebringer et al., 2006, Long et al., 2022). However, no evidence of a mechanism for this association has yet emerged.

1.4 Pathogenesis

AS disease progression is contributed to by three primary processes - inflammation, cartilage erosion, and ectopic ossification. The inflammatory nature of the condition is not thoroughly understood due to the difficulty in obtaining enthesal tissue samples of affected vertebrae, but some tentative hints of the process have been revealed (Haroon et al., 2015, Wendling and Claudepierre, 2013). Immunohistological examination of needle biopsies of the sacroiliac and other zygapophyseal joints in AS patients has revealed immune cell infiltration predominated by macrophages and T-cells, including both CD4+ and CD8+ subtypes (Braun et al., 1995, Appel et al., 2006b, Appel et al., 2006a, Bollow et al., 2000), with macrophage and T-cell infiltrates clustered at foci of ectopic ossification (Braun et al., 1995).

Zygapophyseal inflammatory lesions in AS become local centres of ectopic ossification in the form of syndesmophytes (Maksymowych et al., 2009a). Syndesmophytes are described as bony proliferations originating within a ligament. These ossifications may continue to grow to eventually form bridges between physiological bones, resulting in fusion and consequent complete loss of affected joint articulation. However, this process is coupled to an overall loss of bone mass in the vertebral column and hip of AS patients. Increased osteoclast activity appears to persist within the subchondral femoral heads in AS associated hip arthritis until the overlying cartilage is destroyed (Appel et al., 2006a). Supporting this, bone mass loss correlates with the

formation of syndesmophytes (Karberg et al., 2005), suggesting these two seemingly antagonistic processes are intrinsically linked. This oxymoronic coupling of ectopic ossification and overall bone loss within the axial skeleton is, therefore, a key driver of osteoporosis in AS patients (El Maghraoui, 2004).

The Wingless (wnt) pathway plays a role in osteogenesis (Gaur et al., 2005), and recent studies suggest that its aberration may contribute to syndesmophyte formation in AS. The glycoprotein sclerostin, which blocks wnt pathway signalling, is expressed at lower levels in the zygapophyseal joints of both AS and osteoarthritis patients (Appel et al., 2009). Additionally, the wnt pathway inhibitor dickkopf-1 (dkk-1) is decreased in the sera of AS patients (Daoussis et al., 2010). Experimentally, the functional knockout of the *DKK1* gene has been shown to increase bone mass and to activate ankylosis in mouse models (Uderhardt et al., 2010, Li et al., 2011b). Recent research has identified increased expression of miR-29a and miR-146a, which target Dickkopf-1, in human peripheral blood mononuclear cells and AS fibroblasts, respectively (Huang et al., 2019, Di et al., 2018). Overall, this suggests that localised dysregulation of the wnt pathway may play a significant role in facilitating the formation of syndesmophytes in AS.

While the exact mechanisms underlying AS pathogenesis remain unresolved, there is strong evidence for a concerted immunological response driving ectopic ossification at inflammatory foci within the vertebral column. This process is coupled with a more generalised loss of bone density within the axial skeleton.

1.5 Clinical Indicators of Disease Status

The monitoring of disease status in AS relies on patient-reported health instruments and radiographic evidence. The clinical status of AS can be divided into three categories: disease activity, physical function, and axial structural damage (Zochling and Braun, 2005).

Disease activity reflects acute inflammation and rate of change to physical function and structural damage. To assess disease activity, the Bath Ankylosing Spondylitis Activity Index (BASDAI) is used. This self-reported measure employs six visual analogue scale (VAS) measures to assess fatigue, axial and peripheral pain, morning stiffness, and enthesopathy (Garrett et al., 1994). A final score of 4 or higher indicates active disease. Other measures used to assess disease activity include basic laboratory tests for inflammation, such as serum C-reactive protein (CRP) and erythrocyte sedimentation rate (ESR).

Physical disability is measured using the Bath Ankylosing Spondylitis Functional Index (BASFI), which also employs a self-reporting instrument (Calin et al., 1994). The questionnaire

consists of eight questions pertaining to function and two directed at the patient's ability to cope with everyday tasks. Questions are answered on a VAS scale of 1 to 10, with higher scores representing less favourable outcomes. The average score is assessed in the context of the patient's previous scores. BASFI provides a composite indicator of both disease activity and structural damage (Haywood et al., 2005). The Health Assessment Questionnaire for the spondyloarthropathies (HAQ-s) is another physical disability assessment that targets the patient's ability to conduct contrived tasks, such as a finger-to-floor measurement to establish physical limitation (Daltroy et al., 1990).

Structural damage is quantified by radiography to identify sclerosis and ectopic bone structure in the form of syndesmophytes and ankylosis (Zochling and Braun, 2005). Computed tomography (CT) offers a more sensitive alternative to radiography which can detect localised erosions and ankylosis which may not be visible on a standard x-ray (Zochling and Braun, 2005). This allows for greater sensitivity in recognising the disease in the early stages or in cases of atypical clinical presentation, but offers no value in the interpretation of higher grade sacroiliitis (Fam et al., 1985). However, there are issues with specificity due to false interpretation of age or mechanical stress related changes being misconstrued as evidence of AS related structural impairment (Fam et al., 1985, Zochling and Braun, 2005). Magnetic resonance imaging (MRI) has been successfully employed in the detection of both active inflammatory lesions within the axial column, including for the detection of active sacroiliitis, spondylitis and spondylodiscitis (Zochling and Braun, 2005), and in the assessment of chronic spinal structural changes (Braun et al., 2003). An MRI scoring system (ASSpiMRI) developed by Braun et al. has been successfully employed to verify efficacy in clinical trials, but as yet has not been widely adopted in clinical practice (Braun et al., 2003). Separately, a non-patient reported clinical instrument is the Bath Ankylosing Spondylitis Metrology Index (BASMI). This assessment of physical function is performed in a clinical setting and measures lateral lumbar flexion, tragus to wall distance, lumbar flexion (modified Schober), maximal intermalleolar distance and cervical rotation. Similar to BASDAI and BASFI, results are translated to a scale of 1 – 10 and a final average score is taken.

Quality of life is an important consideration in the management of chronic disease. For this purpose the Assessment for Quality of Life (ASQoL) has been developed (Doward et al., 2003). ASQoL contains elements of BASFI and BASDAI in assessing physical disability and pain, but further expands into areas including patient motivation, sleep, social isolation and emotions.

Assessing osteoporosis is more difficult to assess due to interference in dual energy x-ray absorptiometry (DEXA) bone mineral density (BMD) measures from ectopic bone formation (Magrey and Khan, 2010). There is evidence to suggest improved utility in lateral L3 (Lat-L3)

over posteroanterior (PA) DEXA (Gilgil et al., 2005). As an alternative to spinal BMD measurement, it has been shown that osteopaenia develops in the femoral necks of AS patients prior to detectable changes in the spine, and that the observed reduction in BMD at these sites is proportional to the degree of inflammation (Magrey and Khan, 2010). It should be noted, however, that BMD measurements at the femoral neck cannot be performed on patients who have undergone bilateral hip arthroplasty, as is common in patients with hip joint involvement (Magrey and Khan, 2010).

1.6 Serum Biomarkers in AS

AS has a distinct lack of specific biomarkers available for the purpose of diagnosis or monitoring of disease activity. There are, however, a number of inflammatory and metabolic markers which are raised either constitutively or in correlation with disease activity which will be discussed in this section.

1.6.1 Inflammatory Markers

AS patients exhibit elevated levels of acute phase proteins such as erythrocyte sedimentation rate (ESR), C-reactive protein (CRP), haptoglobin, β_2 -microglobulin, and serum amyloid A (Lange et al., 2000, Yildirim et al., 2004). Moreover, several cytokines and chemokines associated with acute phase response are also enriched in the sera of AS patients. Serum TNF- α is consistently raised in AS, but it does not seem to correlate with disease activity (Gratacos et al., 1994). Similarly, IL-6 levels are persistently elevated in AS patients, but they also correlate with acute phase proteins and ESR (Gratacos et al., 1994). Notably, a study of SpA patients found that IL-8, a potent neutrophil chemokine, only increased in active disease, and its levels correlated with CRP and haptoglobin levels (Sonel et al., 2002). In contrast, no association has been found between serum IL-1 β levels and SpA, which is consistent with the lack of activation of macrophage pathogen activated receptor patterns in this condition (Sonel et al., 2002, Gratacos et al., 1994, Weber et al., 2010).

1.6.2 The Adaptive Immune Response

The specific immune response profile in Ankylosing Spondylitis (AS) is characterised by cytokines of the IL-23/TH17 axis and the Th1 system. Increased levels of IL-17 and IL-23 in the

serum of AS patients confirm the signalling of this immune axis (Mei et al., 2011). A study conducted on Chinese patients with AS demonstrated a strong correlation between the BASDAI index (previously described in Section 1.5) of disease activity and the serum levels of both IL-17 and IL-23. This correlation was found to be more clinically useful than CRP and ESR (Chen et al., 2012). A subsequent study on Caucasian patients also reported an increase in IL-23 levels (Przepiera-Będzak et al., 2015). Other cytokines that correlate with BASDAI include IL-33, which has been linked to increased TNF- α and IL-17 levels in rheumatoid arthritis, and the TH1-inducing cytokine, IL-27 (Han et al., 2011, Lin et al., 2015).

1.6.3 Bone Metabolic Markers

Several studies have reported an increase in bone resorption markers, such as urinary D-pyridinoline, C-telopeptide, and pyridinium cross-links, in a proportion of AS patients (El Maghraoui et al., 1999, Toussirost et al., 1999, Grisar et al., 2002). These markers were found to be associated with disease severity and BMI, specifically the patient's fat mass percentage. Patients with osteoporosis were found to have increased CRP and ESR levels, indicating an inflammatory association (Mermerci Başkan et al., 2010). However, there is conflicting evidence regarding the levels of serum osteocalcin in AS, which may be due to the coupling of bone resorption and deposition in the pathological process, as discussed earlier (Franck and Keck, 1993, Toussirost et al., 1999, Grisar et al., 2002)

1.7 The Genetic Nature of AS

The relationship between genetics and autoimmunity is complex, with genetic risk and environmental risk both playing important roles in disease onset. Familial and monozygotic twin studies are used to estimate the relative genetic contribution to autoimmune conditions. Genetic concordance in AS is relatively high in monozygotic twins (44-75%), but considerably lower in first and second-degree relatives at 8.2% and 1%, respectively (Brown et al., 2000, JÄRvinen, 1995, Hohler et al., 1999).

The MHC class I subtype, HLA-B27 is the strongest genetic association observed in AS and is linked with a family history of the disease (Arévalo et al., 2018). HLA-B27 displays unique features and interactions amongst HLA class I molecules, which has given rise to an *arthritogenic peptide model* for AS aetiology (D'amato et al., 1995). The *arthritogenic peptide model* postulates that specific peptide fragments presented on MHC molecules instigate an inappropriate T-cell response. Despite this attractive explanation for disease causation, no specific peptide

sequence or convincing source has yet been identified. However, a web of additional genes centred around antigen presentation and T-cell activity have been linked to AxSpA, supporting the concept of arthritogenic peptides as a putative aetiological mechanism (Cortes et al., 2013b).

1.7.1 The Role of HLA-B27 in AS Pathology

Despite HLA-B27's significant genetic contribution to AS and its involvement in other inflammatory diseases, the pathological nature of this MHC Class I type remains unresolved. HLA-B27 is associated with AS in all populations, although the frequency of HLA-B27 in AS patients varies among ethnic groups. Despite this, the prevalence of HLA-B27 appears to be over 65% in AS patients of all populations and around 90% in Caucasians (Abdelrahman et al., 2012, Londono et al., 2015, Atzeni et al., 2013, Roberts et al., 2013, Bakland et al., 2005). Moreover, the disease risk is further increased in homozygotes for HLA-B27 over heterozygotes (Jaakkola et al., 2006).

HLA-B27 is not a single entity but a group of over 100 related HLA class I molecules, not all of which are associated with AS. Although all HLA-B27 variants share a common structural arrangement, they differ slightly in amino acid sequence, which may affect their unique antigen binding repertoire (Bowness, 2015). Common to all HLA-B27 subtypes is restricted binding of only peptide fragments containing arginine at position 2, while at positions 1 and 9, preferential binding of positively charged amino acids is observed (Jardetzky et al., 1991, Colbert et al., 1994, Madden et al., 1992). Although it has been proposed that these binding restrictions may cause disease through preferential binding of arthritogenic peptides (D'amato et al., 1995), this appears unlikely since not all HLA-B27 variants are linked to disease.

Despite the dominant genetic contribution of HLA-B27 to AS aetiology and its involvement in other inflammatory diseases, the pathological nature of the MHC Class I variant has not yet been effectively resolved.

Notably, among populations of European Caucasian descent, the HLA-B2705 and HLA-B2702 subtypes are the most associated with AS (Bowness, 2015). In contrast, subtypes such as HLA-B2706 and HLA-B2709 are not associated (Robinson and Brown, 2014). While HLA-B2706 differs from HLA-B2705 only at peptide residue 152, HLA-B*2709 differs only at residue 116 (Bowness, 2015). Residue 116 lies at the bottom of the peptide binding groove, while residue 152 appears to affect T-cell interaction (D'amato et al., 1995, García et al., 1998). These findings suggest that HLA-B27 is pathogenic only where certain conditions in its molecular structure are met. It is therefore possible that peptides have the potential to become arthritogenic only when presented on a specific HLA-B27 structural arrangement.

Apart from unusual peptide binding properties, other unique features of HLA-B27 have been proposed as possible disease-causing mechanisms. For instance, HLA-B27 has demonstrated unusual behaviours both in the presence and absence of bound peptides, including a tendency to homodimerise through cysteine residues at positions 67 and 164 (Antoniou et al., 2004). This feature may lead to species destruction through the unfolded protein response pathway of the endoplasmic reticulum, resulting in downstream ER stress.

Another atypical feature of HLA-B27 is its unique ability to maintain sustained association with the chaperone calnexin where β_2m stabilisation is successful (Boyle and Hill Gaston, 2003, Williams et al., 1989). This allows for transient expression of immature HLA-B27 on cells independent of interaction with TAP or tapasin (TAPBP) transporters associated with antigen processing (Boyle and Hill Gaston, 2003). TAP and tapasin (TAPBP) function to transfer peptides to the ER and to optimise antigen binding, respectively (Kleijmeer et al., 1992, Purcell et al., 2000). The possibility to bypass this process raises the possibility of aberrant peptide loading and presentation by HLA-B27 molecules where tapasin has not interacted with the species (Peh et al., 1998).

There is solid evidence to suggest that HLA-B27 is capable of binding a restricted peptidome of endogenous antigens and may present itself in the absence of bound peptide. While it remains unclear which of HLA-B27's characteristics, or indeed combination of characteristics, contribute to disease, this tentatively suggests that endogenous peptide binding and presentation may play a key role in AS aetiology.

1.7.2 The Discovery of Further Genetic Associations through GWAS

The expansion of genetic understanding beyond HLA-B27 faced the standard challenges prior to the introduction of genome wide methods. However, in 2007, novel large-scale genetic studies began to emerge with the advent of genome-wide association studies (GWAS). These studies merged microarray technologies with the recently completed Human Genome Project to allow for the detection of up to one million single nucleotide polymorphisms and deletions simultaneously. GWAS are a hypothesis-free research solution but require large study sizes to obtain statistical power due to the level of multiple testing involved. They employ the use of mapping variants, frequently located close to genes of interest, which are often co-inherited with other nearby variants within their local linkage disequilibrium (LD) blocks (Freedman et al., 2011). Microarrays with genome wide scope are capable of identifying associated loci, however they may not capture the specific variants within associated loci that contribute to disease. More

focussed chip microarrays such as the Illumina ImmunoChip, are designed based on findings from previous GWAS with input from academic and other researchers to investigate genes previously associated with a broad variety of immune mediated diseases and contain high densities of variants around these sites. Thus they interrogate focussed areas of the genome rather than the whole genome. The ImmunoChip has been widely used to investigate immune mediated diseases including in AS studies.

The first GWAS performed on AS was published in 2010 by The Australo-Anglo-American Spondyloarthritis Consortium, followed by a second study in 2013, which has remained the largest such study to date. This latter study identified a total of 25 AS-associated variants. Subsequent meta-studies performed between GWAS of multiple inflammatory diseases suggest this number to be much higher, but as commonly observed in this field, many of these additional variants may be of vanishingly small effect size.

The GWAS era has revealed a pathological process involving multiple molecular pathways and cell groups. Through this work, two major disease associated pathways in AS have emerged: antigen processing and presentation and lymphocyte function (in particular the TH17 pathway).

1.7.2 The significance of the *ERAP1* locus

One of the most intriguing discoveries from genome-wide association studies (GWAS) is the identification of an association between the endoplasmic reticulum aminopeptidase 1 (*ERAP1*) locus and ankylosing spondylitis (AS), but only in patients who have the HLA-B27 variant (Cortes et al., 2013b, Alvarez-Navarro and Lopez de Castro, 2014). This strong epistatic interaction suggests that there is a functional connection between these two molecules.

MHC class I molecules preferentially load nonamer and decamer peptide fragments, a process which occurs in the endoplasmic reticulum (ER) (Adams and Koziol, 1995). Most endogenous peptide fragments do not meet this criteria in their native state. However, ERAP-1 is an enzyme that resides in the ER, where it works with ERAP-2 to trim longer peptide fragments, which can then be loaded onto MHC class I molecules, including the HLA-B27 variants (Saric et al., 2002, Saveanu et al., 2005). The fact that ERAP-1 has a functional relationship with HLA-B27 provides strong evidence that its role in preparing peptides for loading onto HLA-B27 is involved in the association with AS (Haroon et al., 2012, Sanz-Bravo et al., 2015).

More than 30 risk variants within the *ERAP1* region have been linked to AS, with rs30187 and rs27434 being the strongest associations (Cortes et al., 2013b, Harvey et al., 2009). It is worth

noting that many of these variants are in linkage disequilibrium with each other and thus are likely to be found together in risk haplotypes in the *ERAP1* region (Cortes et al., 2013b). The rs30187 (Lys528Arg) arginine variant is believed to be the primary disease-associated variant because it results in reduced enzymatic activity (Goto et al., 2006). However, it is unclear how this variant contributes to disease causation, although several mechanisms have been proposed (Sanz-Bravo et al., 2015).

It has been suggested that variants that result in reduced ERAP-1 activity may alter the kinetics of substrate binding, leading to changes in the output peptide pool and a possible role in AS aetiology (Alvarez-Navarro and Lopez de Castro, 2014, Sanz-Bravo et al., 2015). Another possibility is that reduced ERAP-1 activity lessens the availability of peptides for HLA-B27 binding, allowing the protein more time to misfold or dimerize through its intrinsic instability, which can trigger its premature elimination through the unfolded protein response pathway. Whatever the underlying mechanism, the unique interaction between HLA-B27 and ERAP1 is certainly a key factor in AS disease aetiology.

ERAP-2 operates in unison with ERAP-1 in the function of peptide trimming. ERAP-2 dimerises with ERAP-1 to catalyse accelerated peptide trimming by increasing the latter's substrate binding affinity (Evnouchidou et al., 2014). Due to strong linkage disequilibrium with *ERAP1*, the study of *ERAP2* in isolation has been challenging. However, a large-scale study has shown that a specific *ERAP2* polymorphism (rs2248374 A->G) has a protective effect in both HLA-B27 carriers and non-carriers with AS (Cortes et al., 2013b, Robinson et al., 2015b). This polymorphism generates a truncated protein variant by extending exon 10 and removing a splice site, resulting in the inclusion of two TAG stop codons (Andrés et al., 2010). Therefore, it is likely that ERAP-2's contribution to disease risk lies in its ability to assist ERAP-1 in peptide trimming.

1.7.3 Genetic implication of T-cell function

A large proportion of the non-HLA genetic variants associated with AS are associated with lymphocyte function. Variants affecting T helper subset 17 (Th17) are particularly prominent. Th17 cells are increasingly implicated in a range of diseases of autoimmune aetiology, including Crohn's disease, psoriasis and ulcerative colitis, all of which are often found comorbid to AxSpA (Yasuda et al., 2019, Braun and Sieper, 2007). Evidence is growing to suggest that Th17 cells also play a prominent role in AS pathology (Zhu et al., 2019). Within the IL-23 receptor gene (*IL23R*), which performs a critical role on Th17 cell differentiation and function, a coding variant of high protective impact against AS, rs11209026 (Arg381Gln) has been reported (Cortes et al., 2013b, Hazlett et al., 2012, Zhu et al., 2020, Deveci et al., 2019, Duvallet et al., 2011). This

variant has been shown to affect splicing of mRNA through skipping of exon 9 (Yu et al., 2015), possibly explaining its protective effect. Further non-coding variants within *IL12B*, a critical subunit to the IL-23 receptor, and *IL6*, which is important to Th17 cell differentiation, have been identified (Kimura and Kishimoto, 2010, Cortes et al., 2013b, Yasuda et al., 2019). Genetic implication of Th17 cells is not restricted to cytokines and their direct receptors: further associations within their downstream signalling molecules, including Janus kinase 2 (JAK2), tyrosine kinase 2 (TYK2) and the highly pleiotropic signal transducer and activator of transcription 3 (STAT3) have also emerged (Danoy et al., 2010, Cortes et al., 2013b, Saadi et al., 2019).

Each of these implicated molecules comprise critical components of the IL-23/IL-17 signalling axis, which acts to coordinate differentiation and proliferation of potentially pathogenic Th17 cells and enhance the inflammatory phenotype through direct and indirect up-regulation of pro-inflammatory cytokines, including IL-6, IL-21, TNF- α , interferon gamma (IFN- γ) (Fig. 1.2). Interestingly, in addition to its role in inflammation, the IL23/IL17 signalling axis has been implicated in bone restructuring through regulation of the delicate balance of osteoclastogenesis and osteoblastogenesis via STAT3 (Grivennikov and Karin, 2010, Lubberts, 2015). This points to a possible dual pathological role for this pathway in both inflammation and in the coordination of ankylosis.

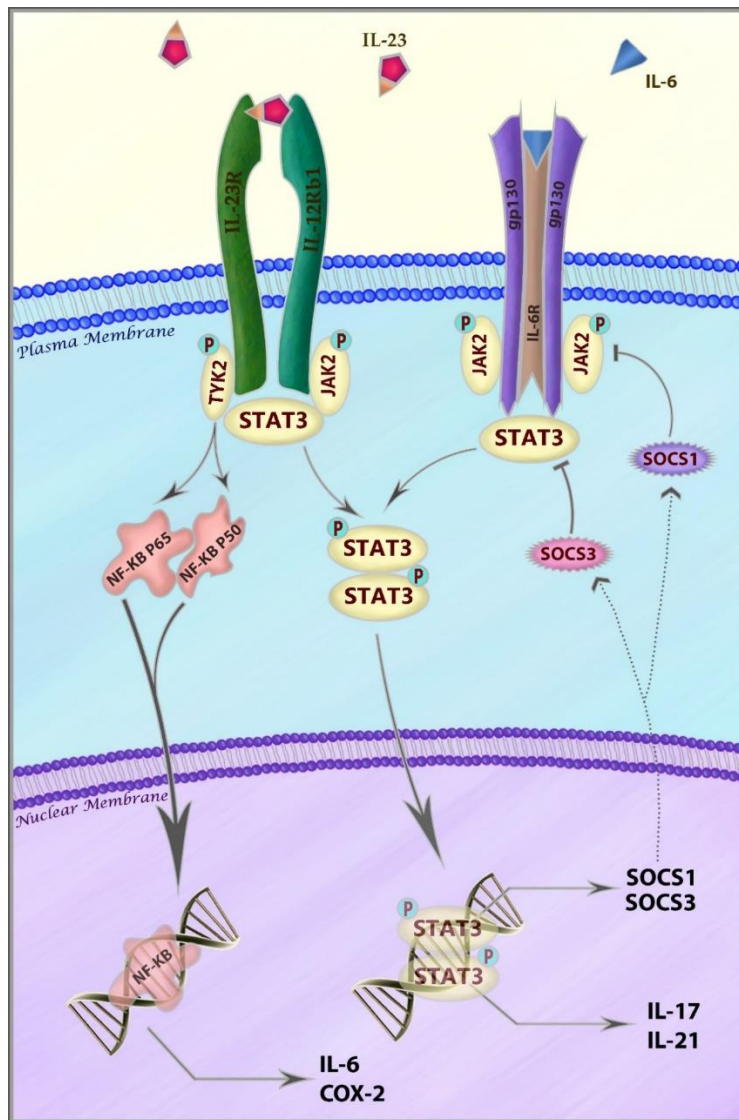


Fig. 1.2: The central role of STAT3 in the IL-17 producing T-cell pathway. STAT3 mediates co-operative signals from IL-23 and IL-6, generating further IL-17 promoting products. Negative feedback is provided by increased expression of SOCS1 and SOCS3, limiting autocrine and paracrine response. Binding of the IL-23 receptor causes activation of both NF- κ B and STAT3. NF- κ B targets a wide audience of genes, including the STAT3 activating IL-6 and Cox-2 (Libermann and Baltimore, 1990, Gauzzi et al., 1996, Lee et al., 2004, Lo et al., 2010). This figure has been included in a published article (Kenyon et al., 2022).

The IL-23/IL-17 axis is also implicated through non-canonical signalling pathways which promote the production of IL-6 and IL-17. An example of this is seen in the prostanoid EP4 receptor (EP4) pathway. Over-expression of EP4 in Th17 cells is linked to increased disease activity in AS patients (Klasen et al., 2019). Genetic association of the upstream *PTGER4* gene, which produces EP4, was established by GWAS in 2013 (Cortes et al., 2013b). Through binding

with its ligand, prostaglandin E2, EP4 triggers activation of the phosphoinositide 3-kinase/protein kinase B (PI3K/AKT) and protein kinase A (PKA) signalling cascades (Yokoyama et al., 2013) (Fig. 1.3). Among other products, this pathway promotes production of IL-6, IL12R β 2 – a component of the IL23 receptor, and cyclooxygenase-2 (COX2) for which prostaglandin E2 is also a known metabolite (Cortes et al., 2013b, Faour et al., 2008, Li et al., 2011a, Narumiya et al., 1999, Sheibanie et al., 2007, Yao et al., 2013). Over-expression of PI3K, has been linked to worse clinical measurements in AS patients, while this pathway has also been linked to osteogenic differentiation through activation by IL-37, which may play a role in the overall spinal osteoporosis previously discussed in *Section 1.4* (Fang et al., 2022b, Fawzy et al., 2016).

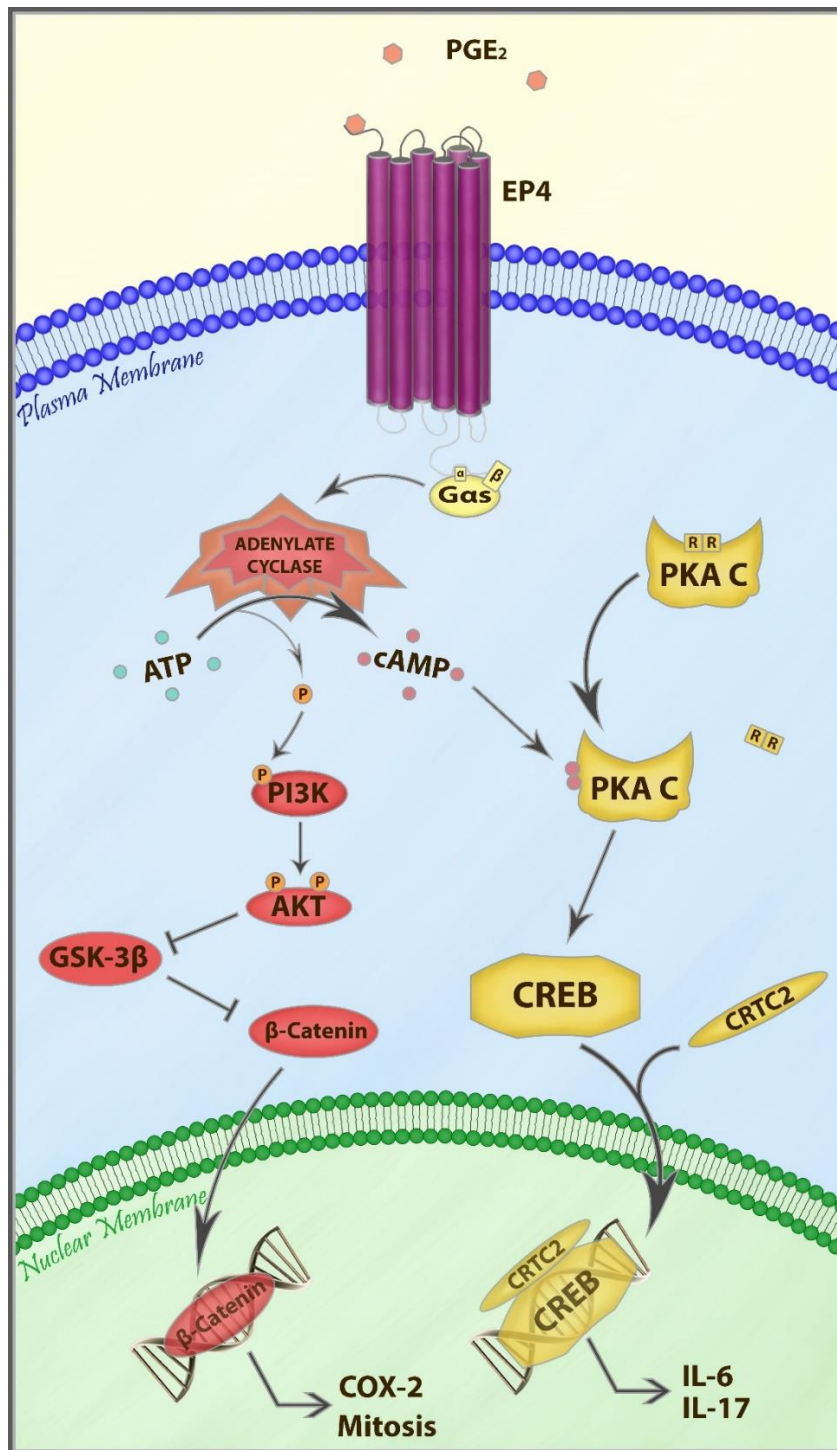


Fig. 1.3: PTGER4 receptor activation promotes the TH17 phenotype through at least two distinct pathways. AKT inhibits the activity of GSK-3 β (Beurel et al., 2015) which allows nuclear migration of β -catenin, promoting the generation of COX-2 which subsequently functions in a positive feedback loop. Activation of cAMP-response element binding protein (CREB) by PKA directly promotes transcription of IL-6 and IL-17, promoting the Th17 phenotype (Yokoyama et al., 2013, Koga et al., 2021, Chava et al., 2009).

A variant found near caspase associated recruitment domain 9 (*CARD9*) highlights another pathway of interest which is implicated in T-cell response. *CARD9* is an adaptor protein which acts as an intermediary for signals transduced by surface t-cell receptors, TLR's and Dectin-1. It plays important roles in fungal immunity and in intracellular bacterial clearance (Dennehy et al., 2009, Hsu et al., 2007). *CARD9* interacts canonically with *MALT1* and *BCL10* and non-canonically via JNK and P38 to promote cell survival, proliferation and immune function (Hara et al., 2007, Colonna, 2007, Hsu et al., 2007, Gross et al., 2006, Rebeaud et al., 2008). *CARD9* interaction with *MALT1* and *BCL10* leads to activation of NF κ B and consequent upregulation of IL-6 and IL-10 (Dennehy et al., 2009). Similar to EP4, *CARD9* is suggested to play a key role in TH17 phenotype development by promotion of IL-23 expression, possibly via JNK and P38 (Sokol et al., 2013, Puel et al., 2010, Dobrev et al., 2009).

1.7.4 Genetic associations found within 'gene desert' regions

GWAS for many diseases have also identified variants associated with disease risk that are located at significant distances from the nearest genes. In AS, several variants have been identified within gene 'desert' (extensive intergenic regions), including rs6759298 in 2p15 and rs2836883 in 21q22 (Cortes et al., 2013b). Interestingly, amongst non-HLA-variants these variants present with relatively high odds ratios (1.28 and 1.30, respectively), despite their distant location from genes of interest (Cortes et al., 2013b).

To establish the contribution of such variants to disease, it is necessary to investigate their potential role in altering cell-specific regulatory elements such as enhancers or repressors. A critical step in this process is to identify if these regions interact with distal genes in the folded chromatin structure. Chromosome conformation capture technologies such as Hi-C, which use a chromatin fixation method to identify the structure of chromatin (Van Berkum et al., 2010) are gaining popularity in this regard. By comparing risk and non-risk alleles for a variety of cell types, chromosome conformation capture can help establish whether alleles of interest are influencing chromatin folding and where the variant localizes within the folded chromatin structure.

1.7.5 The use of fine mapping techniques to establish causal variants

Deciphering the causality of a given associated tag variant near a relevant gene can be challenging due to the phenomenon of linkage disequilibrium (LD). LD is the non-random distribution of alleles at a given locus within a population, resulting mainly from a lack of recombination events in local regions. As a result, LD gives rise to co-inheritance of multiple variants within

haplotypes in a population (Fig. 1.4). While GWAS studies exploit LD by using tag variants to capture haplotypes, this introduces the problem that the tag variant may not be causal within the haplotype(s) it tags.

To address this issue, fine mapping techniques have been developed, including statistical models such as stepwise regression, or shotgun significance evaluation of each neighbourhood across a given locus. Other methodologies include functional annotation through exploiting extra-chromosomal data, such as cellular expression, DNA methylation and chromatin accessibility; and multi-ancestry analysis, in the hope that local haplotypes may differ between populations (Hutchinson et al., 2020). However, multiple causal variants of differing directionality and effect size may act within a given locus (Hutchinson et al., 2020). Therefore, effective unravelling of these variants can be complex and may require multiple approaches and large sample sizes.

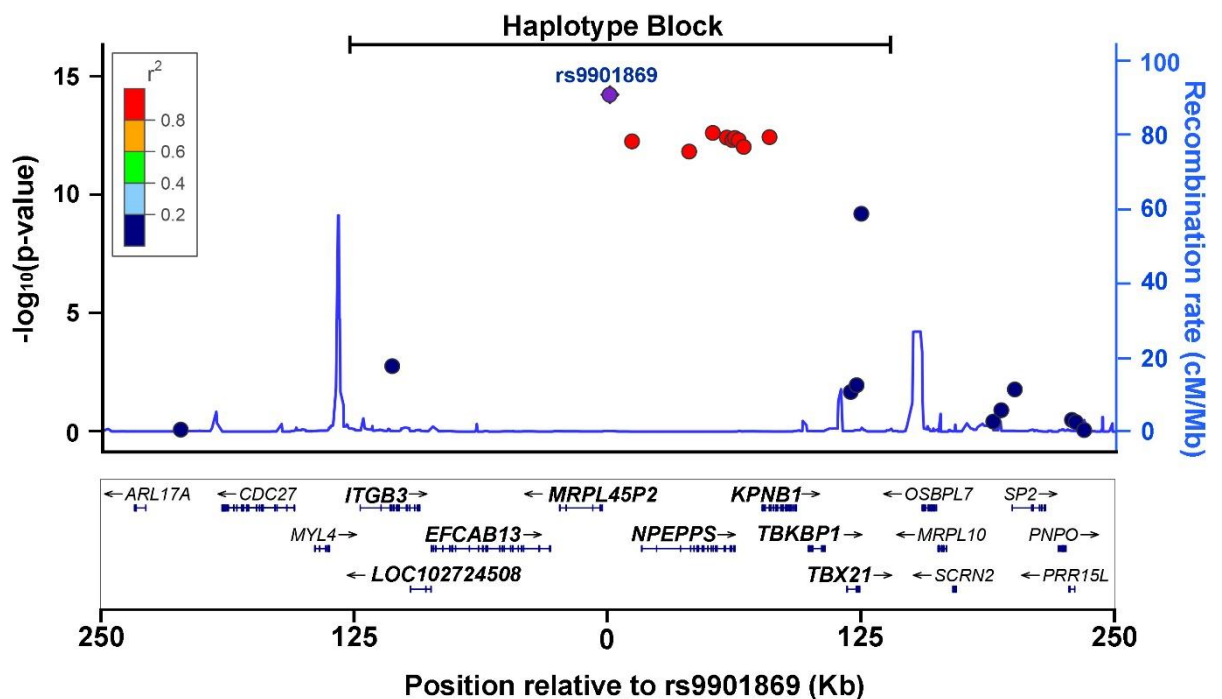


Fig. 1.4: An example of a haplotype block around the SNP rs9901869 on chromosome 17 (chr17:47497740, GRCh38). This variant shares strong LD with several other polymorphisms in its vicinity, shown as red dots. These tend to occur together with rs9901869, and thus share similar frequencies and statistical significance in association analyses. Flanking the region are two hotspots of recombination. Created using LocusZoom (Pruim et al., 2010).

1.8 The relationship between genetics and clinical phenotype

1.8.1 Genetic commonality in disease comorbidity

As previously mentioned, while ankylosing spondylitis (AS) can occur in isolation, it's often associated with other autoimmune conditions known as extra-articular manifestations (EAMs). These EAMs include acute anterior uveitis (AAU), psoriasis (Ps) (including psoriatic arthritis), ulcerative colitis (UC), and Crohn's disease (CD). Each of these are considered discrete diseases in their own right, and these may onset before or after the emergence of classical AS symptoms.

Although each of these EAMs has a unique genetic profile when observed in isolation, there is considerable overlap of genetic risk both with AS and between each other. HLA-B27 is a common finding across all these conditions, although its impact is notably weaker in skin manifestations of psoriasis (Queiro et al., 2015). Additionally, the IL-23/IL-17 axis is linked across all conditions, with many genetic variants in this pathway common to the group (Ellinghaus et al., 2016, Huang et al., 2020).

Acute anterior uveitis (AAU) is the most commonly observed extra-articular manifestation (EAM) in patients with ankylosing spondylitis (AS), with around 30-40% of patients presenting with at least one episode (Robinson et al., 2015a, Cantini et al., 2015). HLA-B27 is the strongest genetic association observed in AAU, seen in almost half of isolated AAU cases (Robinson et al., 2015a, Cantini et al., 2015). Distinct from AS, there are well described major clinical differences between the presentation of AAU depending on whether the case is HLA-B27 positive or HLA-B27 negative. HLA-B27+ AAU presents as an acute unilateral disease, affecting younger persons and with a distinct male preponderance (approximately 2:1) (Chang et al., 2005). HLA-B27 positive AAU also tends to recur. HLA-B27- AAU, by contrast, tends to be chronic, bilateral and affecting middle-aged patients with no perceivable gender bias (Braakenburg et al., 2008, Wakefield et al., 2011).

The first GWAS of AAU, published in 2020 (Huang et al., 2020), revealed genetic associations common to AS, including within or near to the *ERAP1*, *IL6*, *IL23R* and *IL12B* genes, echoing the antigen presentation and IL-23/IL-17 axis mechanisms observed in AS. Genetic risk unique to AAU includes a variant near to *MERTK*, which functions within the anti-inflammatory IL-37 pathway. This pathway has been shown to promote resolution of inflammation through influencing macrophage/myeloid cells (Liu et al., 2016, Nold-Petry et al., 2015).

Most interestingly, the same genetic associations within the the gene ‘desert’ regions of 2p15 and 21q22, previously reported in AS, were also associated with AAU (Huang et al., 2020, Robinson et al., 2015a). To date, the downstream effects of these variants have not been characterised.

Psoriasis, IBD, and AS all share genetic risk in at least 31 loci. Of these, 12 are linked to the IL-23/IL17 axis (Ellinghaus et al., 2016). Psoriasis is present in around 10-25% of AS cases (El Maghraoui, 2011). Genetic concordance between AS and psoriasis has been estimated to be approximately 50% (Ellinghaus et al., 2016). HLA-B27 is a less important marker in psoriasis, but features prominently in psoriatic arthritis (PsA), which may develop during the course of the disease (Queiro et al., 2015). PsA is closely related to AS and presents with a range of overlapping symptoms (Feld et al., 2018). However, pathway enrichment analysis has identified a large number of unique molecular pathways (273 in AS and 67 in PsA) that are specific to each disease, and 37 pathways which overlap (Rahmati et al., 2021). Moreover, among the shared pathways, a significant number have divergent directionality, with inverted allele frequencies of variants between the two diseases. The IL-23/IL-17 axis again features heavily among the overlapping pathways, with common variants in the IL12b and IL17 genes, and their downstream components, including the JAK/STAT pathway and T-cell differentiation pathways such as Th17 (Rahmati et al., 2021).

Surprisingly, despite this common association of the IL-23/IL-17 axis, therapeutic response to IL-23 inhibition has been successful in treating PsA but failed to achieve response in AS (Mease and van den Bosch, 2021), indicating distinct molecular aetiology between the two conditions.

IBD (UC and CD, collectively) presents in under 10% of AS cases (El Maghraoui, 2011). Shared genetic risk is observed in the IL-23/IL-17 axis, including the previously discussed non-synonymous variant in the *IL23R* gene, rs11209026 (Arg381Gln), and other variants affecting common genes such as *PTGER4*, *IL12B*, *CARD9*, and *ICOSLG* (Cortes et al., 2013b). Online methods, such as the Open Targets web resource, have enabled more exhaustive genetic commonality analysis, and a functional summary of genes common to AS, psoriasis, UC, and CD is presented in Table 1.4, with a full description of each of these genes available in supplementary material, Table S1 (Ochoa et al., 2020).

An interesting observation among AS patients is the presence of subclinical bowel lesions (Mielants et al., 1988). These lesions appear to be linked to polymorphisms in the *CARD15* gene, which produces NOD2, a protein associated with Crohn's disease (CD) and implicated in the pathogenesis of AS (Laukens et al., 2005). NOD2 is an intracellular receptor for microbe-associated molecular patterns (MAMPs), and perhaps more importantly a key player in autophagy; it also plays a crucial role in maintaining the integrity of the intestinal barrier (Ferrand et al., 2019). However, research on gut microbiota suggests that the microbiome composition in

AS patients with subclinical lesions differs from that of CD patients (Sternes et al., 2022). This suggests that polymorphisms within the *CARD15* gene may indirectly influence gut microbiota, leading to the emergence of these subclinical lesions.

Table 1.4: Functional overview of susceptibility genes shared between AS, psoriasis, ulcerative colitis and Crohn's.

Gene function	Gene count
Lymphocyte function	19
T cell	15 ^a
B-cell	2 ^a
Involved in IL23/IL17 or JAK/STAT pathway	10
Transcriptional regulation	16
Antigen processing/presentation	5
Metabolism	5
Apoptosis	4
Protein degradation	4
Innate immune cell function	4
Tissue remodelling/TGF beta pathway	3
Cellular proliferation/differentiation	2
Cellular adhesion	2
Other/unknown function	21

^aSome genes involved in multiple lymphocyte cell subsets/functions. Data obtained from Open Targets online methods (Ochoa et al., 2020). Qualifying genes had an Open Targets average genetic association score of 0.7 across at least 4 of AS, AAU, Ps, CD and UC. This table was sourced from published work (Kenyon et al., 2022).

1.8.2 How genetics impact on disease severity

HLA-B27 is both the strongest and most impactful genetic association in AS. Several studies have shown that HLA-B27 is more likely to be present in male patients and those who develop the disease at an earlier age (Akassou et al., 2015, Li et al., 2022, Feldtkeller et al., 2003, Arévalo et al., 2018). A large study reported that HLA-B27 negative patients may be more prone to higher disease activity, as measured by BASDAI, and functional impairment, as measured by BASFI (Arévalo et al., 2018). However, contradicts several other smaller studies (Freeston et al., 2007, Vargas-Alarcón et al., 2002). Also, in Chinese AxSpA patients, a large study found that HLA-

B27 was linked to increased lumbar and thoracic spine involvement but appeared to have no impact on age of symptom onset, peripheral arthritis or uveitis (Yang et al., 2013a).

Familial commonality in disease severity has also been reported in AS patients. Sibling pairs with AS show some correlation in disease activity and radiographic progression, while children of AS parents who reported AAU were more prone to developing uveitis (Brophy et al., 2004). There remains a dearth of studies examining the phenotypic effects of associated variants other than HLA-B27. A Spanish study found that patients who presented with peripheral arthritis at disease onset were more likely to develop the disease at a younger age and had genetic variants near genes including HLA-DQ/HLA-DO, IL23R, and JAK2 (La Borda et al., 2019). However, it is important to note that this study did not adjust for the inflated α of multiple testing.

1.8.3 The polygenic risk approach to predicting disease onset and severity

Following the era of genome-wide association studies (GWAS), interest has grown in the concept of personalized total genetic risk of disease. Currently, the most commonly used approach to achieving this is the use of GWAS data to obtain "polygenic risk scores" (PRS). These are calculated using a variety of similar techniques which all centre around averaging the effect sizes of all known associated genetic variants to obtain a single, personalised score (Duncan et al., 2019). This type of approach can be applied to any non-mendelian disease or trait that carries a known polygenic component. Several studies have demonstrated its predictive value in Alzheimer's disease, myocardial infarction and cancer (Kachuri et al., 2020, Khera et al., 2019, Marden et al., 2016). There is also evidence to suggest that PRS may be able to predict more severe disease in systemic lupus erythematosus (Chen et al., 2020). In AS, a recent study has shown that PRS may be more valuable than HLA-B27 as a predictive marker for the disease (Li et al., 2021). However, no studies have yet examined whether PRS may also offer predictive value in AS or AxSpA disease severity.

In the near term, PRS face some limitations, primarily due to the quality of the GWAS data on which they are based. Effect size inflation of tag variants resulting from sources including minor allele frequency, genotyping quality and population structure are problematic in developing models for PRS calculations (Choi et al., 2020b). Another issue is the dominant focus of GWAS studies on Caucasian populations. A dearth of data involving other groups has led to a comparative lack of representation in genetic association data, consequently affecting the predictive power of PRS in these populations (Lambert et al., 2019). Therefore, the power of PRS can be improved by collecting more robust genetic association data within a broader representation of ethnic groups (Wray et al., 2014, Márquez-Luna et al., 2017).

Despite the roadblocks faced by PRS, it remains a promising solution for a more impactful predictive test than individual genotyping alone. In the future, it may become integrated into the AxSpA diagnostic work-up, possibly as a replacement for HLA-B27 typing. Furthermore, PRS methodologies are likely to advance beyond simple calculation against GWAS data. Solutions are already being developed to account for epistasis, such as the gene pathway analysis package for the PRSice PRS package (Choi and O'Reilly, 2019). Thus, PRS have potential applications in the diagnosis, prognosis and in personalising medicine.

1.8.4 Genetics play a role in therapeutic selection in AS

Targeted biological therapies have become a key component in treating severe AS cases. The first line of these therapies involves TNF- α inhibitors (TNFi), which come in the form of either a recombinant receptor or chimeric IgG antibody. Since FDA approval in 2005, TNF- α inhibition has proven to be a game-changer (Braun et al., 2005) by reducing the infiltration of immune cells into synovial tissue, halting the inflammatory process (Smeets et al., 2003, Baeten et al., 2001). However, around 40% of recipients either do not respond adequately or experience adverse effects (Dougados and Baeten, 2011).

Genetic and phenotypic evidence in AS has pointed directly at Th17 cells, leading to clinical trials of inhibitors of the IL-23/IL17 pathway. IL-17 inhibition trials have reported highly favourable results (Braun et al., 2017), resulting in recent approval for AS treatment. However, IL-23 inhibition has not proven effective. Despite excellent response in clinical trials and subsequent approval for the treatment of psoriasis, direct IL-23 inhibition (risankizumab) and combined IL-23/IL-12 inhibition (ustekinumab) have both proven ineffective in treating AS (Baeten et al., 2018, Reich et al., 2019). The cause for this surprising finding has not yet been resolved. One proposed mechanism is a reduced importance of IL-23 in AS, where stimulation of Th17 cells may originate through alternative pathways such as IL-1 and IL-6 (McGonagle et al., 2021). This hypothesis, however, is at odds with the strong association of the *IL23R* Arg381Gln variant with AS (Cortes et al., 2013b). It is possible that the mechanisms of IL-17 stimulation change throughout disease progression or that IL-23 involvement is present in only a proportion of patients for whom genetic risk in this pathway is highest (Kenyon et al., 2022). Further study is required to examine the specific response of immune cells in AS and psoriasis patients in relation to IL-23.

There is growing interest in targeting downstream signalling molecules of the JAK-STAT pathway. While broad-spectrum Janus kinase inhibitors (JAKinibs) have demonstrated efficacy, they come at the cost of significant adverse effects (Mohanakrishnan et al., 2022, Wishart et al., 2006, Tian et al., 2019). As an alternative, the targeting of specific Janus kinases is being explored. One such molecule is TYK2, which has been directly associated with AS through genetic studies (Cortes et al., 2013b).

Early indications in a murine model suggest that TYK2 inhibition may be effective in treating AS, while human safety trials have indicated improved safety of TYK2 inhibition over broad-spectrum JAKinibs (Hromadova et al., 2021, Singh et al., 2021, Nogueira et al., 2020, Gracey et al., 2020). Additionally, the targeting of STAT3 is being investigated (Jo et al., 2021). However, therapeutically targeting broad-spectrum transcription factors has historically proved challenging (Medeiros et al., 2021).

1.9 Summary

Until recently, AS has been a disease which is defined by clear and characteristic manifestations in the lumbar spine. However, this classical presentation has become increasingly blurred as AS has become absorbed into the AxSpA group of diseases. The term *ankylosing spondylitis* itself is slowly being phased out in favour of *radiographic AxSpA* (rAxSpA). The ASAS criteria for AxSpA have been instrumental in facilitating the early diagnosis and treatment for patients, although this early diagnosis inevitably comes at the cost of including a broader range of aetiologically similar pathologies under the AxSpA group. Already it has been clearly demonstrated that nrAxSpA is more genetically diverse than AS and that nrAxSpA will not necessarily progress to AS. It is probable that genetic profiles of rAxSpA patients are subtly different to traditional AS, due to the highly stringent selection criteria of the formerly used mNY criteria. Early intervention may also prevent detectable sacroiliitis from developing altogether (Maguire et al., 2022a), making it impossible to define which patients would likely progress to rAxSpA in the absence of such treatment. This drift in disease definitions will feed into future studies examining this group of diseases. These changing goalposts provide a difficult environment for both genetic association analysis and clinical studies of patient cohorts, particularly when comparing contemporary work to previous studies.

While HLA-B27 has been the subject of intensive research, it is remarkable that its precise role in disease aetiology remains unresolved. Suspicion lies in the *arthritogenic peptide hypothesis*, which is supported by additional genetic and phenotypic evidence. Studies examining the

phenotypic divergences of HLA-B27 positive and negative patients have been helpful in establishing how genetic risk may influence disease onset and outcomes. However, some evidence remains conflicting between these studies, particularly in relation to the impact of HLA-B27 on disease severity. These conflicts could be the result of study sizes, patient selection biases, the impact of therapeutics on disease progression or study populations. Beyond investigations of HLA-B27, there is a paucity of research linking genotype with clinical phenotype. Study of this area is challenging due to the small effect sizes of common variants in relation to the disease. Possible candidate genes of relatively large effect size include ERAP1, which plays a key role in antigen presentation through its interaction with MHC class I molecules, such as HLA-B27. Meanwhile, the overall impact of genetic variants of small effect size may be better quantified through the use of composite measures such as PRS. While these personalised scores are unable to pinpoint specific genetic associations, they provide a viable option for the initial investigation of how common genetic risk affects outcomes in disease.

While a large number of genetic associations have been identified in AS, there remains a lack of understanding of how these variants may contribute to disease. Genetic variants known to affect lymphocyte function are particularly prominent in AS. T-cell involvement has become well established as a growing body of evidence highlights the roles of all T-cell subsets, with a focus on the Th17 subset. Meanwhile, druggable molecular pathways such as the IL-23/IL-17 axis have been successfully targeted, with the first IL17 inhibitors recently approved for treatment of NSAID/TNFi refractory AS.

Despite these early successes, and genetic implication of the IL-23 pathway, treatment of AS with IL-23 specific DMARDs has proved ineffective. This suggests a fundamental lack of understanding of the molecular aetiology remains. Variants affecting a single gene may impact upon many different biomolecular pathways. Further complexity is added through remote DNA-DNA interactions within the chromatin structure, where variants may exert long distance cis effects through chromatin folding. These interactions and pathways may be limited to particular cell types or activated only under specific stimuli. Therefore, it is critically important that cell specific studies are conducted to understand how gene expression in AS patients differs from that of the healthy population and under what circumstances these potentially pathological differences arise. Recent technologies such as single cell RNA expression and proteomics offer a granular approach to understanding cellular response, which may be transformative in unravelling some of these mysteries.

1.10 Rationale

AxSpA is a complex disease which presents with a spectrum of manifestations and severity. Although the contribution of genetic risk in AxSpA aetiology has been well documented, less is known about how this genetic risk plays into eventual outcomes in affected individuals. Studies have shown that HLA-B27 is linked with earlier disease onset and uveitis in AxSpA patients, but no studies have examined the relationships of other genetic risk variants and disease phenotype. Candidate genes for exploration include *ERAP1* which is strongly linked with AxSpA disease only in HLA-B27 positive individuals. Therefore, it is possible that this gene exerts influence on disease severity through HLA-B27. Although some risk traits, such as HLA-B27 and *ERAP1*, mentioned above, have a relatively higher impact on AxSpA aetiology, many other risk variants of lower impact have also been associated with the disease. Although it is not practical to investigate the impact of these variants individually, it may be possible to assess whether these variants collectively correlate with disease presentation.

Few patients who carry genetic risk for AxSpA will ever develop the disease. This suggests that an environmental trigger may be involved in disease onset. This notion ties in well with the restricted antigen repertoire which can be loaded onto HLA-B27 molecules for subsequent presentation to T cells. To date, no specific agent, infectious or otherwise, has been isolated. However, it has been noted that a significant proportion of AxSpA patients present with subclinical inflammation in the gut, while a smaller proportion present with IBD in the form of ulcerative colitis or Crohn's disease. This hints toward possible involvement of the gut microbiome in AxSpA pathogenesis. Recent studies have revealed associations between viral pathogens and autoimmunity. Whether the immunological response to microbial pathogens is possibly impaired in AxSpA patients is, therefore, an avenue which is worth exploring.

1.10.1 Objectives

1. Perform a comprehensive appraisal of the clinical characteristics of the ASRI AxSpA cohort and to investigate specifically the relationships between common EAMs and establish how EAMs develop over the course of AxSpA disease.
2. Determine whether EAMs have a heritable component aside from genetic risk association attributable to AxSpA.
3. Investigate whether genetic risk variants within the *ERAP1* gene, which is co-opted by HLA-B27 are related AxSpA disease onset and severity.
4. Identify novel AxSpA genetic risk variants using meta-analysis of genetic data from ASRI with previously published genetic studies.
5. Explore the landscape of gene expression of disease relevant immune cells at a single cell level to identify overlap between genetic risk variants and gene expression and to further report the bio-molecular pathways which underpin disease pathogenesis in AxSpA.

Chapter 2

Clinical features of the *ASRI* cohort

2 A clinical appraisal of the ASRI cohort

2.1 Introduction

2.1.1 AS and AxSpA: increasingly blurred definitions

The ASRI cohort differs from other cohorts which have been genetically profiled to date due to the inclusion of non-radiographic AxSpA, which has broader catchment criteria. AxSpA comprises a group of inflammatory diseases with considerable cross-over, both in terms of phenotypic presentation and in the genetic factors underlying diseases within the group. In contrast, traditional cohorts have mainly focused on AS, which is a subtype of AxSpA characterized by radiographic sacroiliitis and progressive formation of syndesmophytes in the vertebral column. However, such studies are becoming increasingly difficult to design, due to the introduction of biologic disease modifying anti-rheumatic drugs (DMARDs) in the first decade of this century. These remarkably effective treatments markedly improve patient outcomes and may even prevent the onset of radiographic change altogether in some patients if administered early in disease course (Maguire et al., 2022a). Therefore, the application of DMARDs may mask the classical features which would otherwise lead to a diagnosis of AS. Due to this and a shift in medical consensus, clinicians are becoming increasingly reliant on the new ASAS criteria for AxSpA (Rich-Garg et al., 2021), which has existed for little over a decade and remains under-evaluated from both clinical and genetic perspectives.

Similar to other complex diseases, the aetiology and pathophysiology of AxSpA result from the intricate interplay of genetic and environmental factors. Consequently, when performing a study which aims to examine the relationship between genotype and clinical phenotype of this condition, it is imperative to first characterise the phenotypic diversity of the study cohort. This approach enables us to not only discern the interconnections among clinical features but also to gauge the extent to which various environmental influences contribute to disease outcomes.

To this end, we performed an in-depth study of the clinical characteristics of ASRI participants, with a specific emphasis on aspects that have received limited attention in prior AxSpA studies. The ASRI AxSpA cohort is a large national registry of AxSpA patients from across Ireland. These participants have undergone thorough phenotyping via a structured interview procedure, encompassing clinical assessments conducted during the registration phase. Within this chapter, we present an outline of the ASRI AxSpA cohort from a clinical vantage point, including the impact of gender, disease duration, and lifestyle factors on disease outcomes.

2.1.2 Measuring disease outcomes in AxSpA

Disease outcomes in AxSpA include the activity of the disease, both in terms of inflammation and spinal radiographic progression and also by the onset of EAMs and comorbidity. In the previous chapter (Chapter 1 Section 1.5) we introduced a series of patient reported outcomes (PROs) for the clinical assessment of disease severity and spinal disability, these including BASDAI, BASFI, BASMI, HAQ-s and ASQoL. All of these measures demonstrate excellent reliability in the monitoring of disease activity, patient functioning, spinal mobility and quality of life, both within patient (i.e. as a follow-up measurement during frequent attendance at the clinic) and by intra-observer correlation, where this has been performed (Zochling, 2011). While each of these measures represents a distinct combination of features, some overlap is observed. For example, both BASMI and BASFI, which measure physical disability, are known to correlate with progressive spinal deterioration in patients (Martindale et al., 2006). Meanwhile, disease activity and quality of life, as measured by BASDAI and ASQoL, have little to no correlation with radiographic spinal progression (Martindale et al., 2006, López-Medina et al., 2018). Collectively, therefore, these measures offer a comprehensive clinical assessment of patient health in AxSpA.

In our introductory chapter (Chapter 1, Section 1.8.1), we discussed extra-axial manifestations (EAMs) that are associated with AxSpA. These EAMs, including acute anterior uveitis (AAU), psoriasis, ulcerative colitis, and Crohn's disease, share common genetic risk factors with AxSpA. Uveitis is the most commonly observed EAM in AxSpA, and its frequency, onset, and relationship to AxSpA disease outcomes have been well-described previously. Uveitis can occur at any point following AxSpA disease onset and may even precede AxSpA in some cases (Gómez-García et al., 2021). Evidence suggests that AAU preceding AxSpA onset may lead to less severe AxSpA disease (Gómez-García et al., 2021), although it is unclear whether this is due to earlier therapeutic intervention resulting from shorter delay to AxSpA diagnosis in these patients.

In contrast, other EAMs are less well described in terms of their relationships with AxSpA outcomes. Psoriasis has been shown to affect around 20% of AxSpA patients and is linked to worse functioning as measured by BASFI (Doran et al., 2003). IBD is found in less than 10% of AxSpA patients, although subclinical gut inflammation is observed in 25-49% of patients (Rudwaleit and Baeten, 2006). Approximately 6% of patients presenting with this subclinical gut inflammation are expected to eventually develop overt IBD (Rudwaleit and Baeten, 2006). Peripheral arthritis (PA) is one of a number of other EAMs and is observed in approximately 25% of patients with AxSpA and has been linked to earlier AxSpA onset, increased prevalence of EAMs, increased disease activity, reduced functioning, and reduced quality of life (Singh et al., 2008, La Borda et al., 2019, Boonen et al., 2003, Heuft-Dorenbosch et al., 2004, Capelusnik

et al., 2021). PA may first develop at any point during the course of AxSpA disease, including periods of spinal remission (Ginsburg and Cohen, 1983). Currently no data is available on the rate of onset of PA following the first symptoms of AxSpA.

Together with PROs, EAMs provide a robust assessment of patient health, which we will use as markers for disease status both within this chapter and in our later investigations of genetic risk.

2.1.2 Differences in presentation between males and females

The male preponderance in AS has been well-defined in many studies, with approximately three times as many men as women developing the condition (Braun and Sieper, 2007). However, there is evidence to suggest that women who develop the condition experience a slightly different syndrome from their male counterparts. Previous studies, including research conducted within the ASRI cohort, have shown that men generally experience increased radiographic change and are more likely to suffer with permanent employment disability, while women score worse in markers of disease activity and quality of life (Landi et al., 2016, Maguire et al., 2022d).

The prevalence of extra-axial manifestations in AS also differs by gender. Women with AS are more likely to develop axial enthesitis and peripheral arthritis, including dactylitis (Landi et al., 2016) and psoriasis (Rusman et al., 2018). Conversely, AAU appears to have a male predilection, particularly amongst those who carry the MHC class I HLA-B27 allele (Linssen and Meenken, 1995). Furthermore, men tend to develop AAU at a younger age than women (Linssen and Meenken, 1995). IBD is possibly associated with female sex in AxSpA, although results have been inconsistent (Rusman et al., 2018, Cunha et al., 2021). While most of these studies specifically concern AS, these features are not well described in AxSpA and have not been reported by studies of the ASRI cohort.

Thus, while AS exhibits a strong predilection towards men, the presentation and progression of the disease differs significantly between men and women. Although these differences have been well documented in AS, with very little data published on AxSpA as a whole. EAMs, in particular are not well described in AxSpA in relation to gender.

2.1.2 The effects of smoking and obesity on AxSpA disease outcomes

Smoking and obesity are both established risk factors for disease severity and comorbidity in AxSpA and are therefore important considerations in the relationship between genotype and phenotype. In AS, smoking has been associated with spinal progression (Poddubnyy et al., 2012)

but it also leads to negative outcomes such as increased disease activity, worse functioning, and reduced quality of life (Jones et al., 2017, Zhao et al., 2019, Poddubnyy et al., 2012). Interestingly, the prevalence of uveitis and peripheral arthritis is reduced in those who have ever smoked (Jones et al., 2017, Zhao et al., 2019). Conversely, the cessation of smoking, is associated with improved outcomes in BASDAI and ASQoL over current smokers (Jones et al., 2017).

Obesity plays a similar role in disease severity, as it has been linked to spinal radiographic changes and worse outcomes in disease activity and functioning in AxSpA, including within the ASRI cohort (Maguire et al., 2022b, Bakirci et al., 2020, Maas et al., 2016). Additionally, obesity has been shown to affect pharmacokinetics of biologic DMARD therapies (Rosas et al., 2017). While smoking is in decline in Ireland (Ipsos, 2021), obesity is of particular concern due to its rising prevalence within Western populations (Janssen et al., 2020).

Although both smoking and obesity are established risk factors in AS, the relationship between them is complex. Smokers are less likely to be obese than non-smokers, but smoking cessation is associated with a significantly increased risk of obesity and consequent comorbidity, including type 2 diabetes (Dare et al., 2015, Morimoto et al., 2012). In AxSpA patients, physical disability can limit daily activity (Nikiphorou and Ramiro, 2020), and a smoking habit may further impact functioning (Villaverde-García et al., 2017), potentially placing patients with a history of smoking at particular risk of post-cessation obesity and associated comorbidity. However, there have been no studies to date examining the additional potential risks of post-cessation obesity in the context of AxSpA. We, therefore, as part of our clinical profiling of the ASRI cohort, specifically aim to investigate the effects of smoking cessation on AxSpA disease outcomes within the ASRI cohort.

2.2 Aims

- To characterise the ASRI AxSpA patient cohort by demographics, disease characteristics, comorbidity, lifestyle and current therapies.
- To model the relationships of clinical measures of disease severity and EAMs with disease duration within the ASRI AxSpA cohort.
- To examine gender-based differences in AxSpA disease presentation.
- To investigate the effects of EAMs on AxSpA disease outcomes.
- To establish whether cessation of smoking presents additional risk related to obesity in AxSpA patients.

2.3 Methods

2.3.1 Data collection

The study participants' data was collected through a structured interview format conducted by a trained professional. Information was obtained through clinical assessment and clinical records at registration or self-reporting, if clinical records were unavailable. Data collected from patients included demographics such as age, sex, ethnicity, and family history of AxSpA. Clinical characteristics such as age of disease onset (self-reported), age of diagnosis, extra-axial manifestations, comorbidities, modified New York (mNY) criteria, ASAS criteria for axial spondyloarthritis, and current therapies were also recorded. HLA-B27 types were obtained from genotyping data, where possible, with the methodologies covered in Chapter 3, Sections 3.38-3.39. Out of the 912 patients included in the study, 641 had genotyping data available to allow for HLA-B27 status assessment. For the remaining patients, HLA-types were taken from ASRI database records, obtained from medical records or self-reported, where available.

Additional data collected was as follows:

- Cardiovascular risk markers, including BMI, waist circumference and blood pressure.
- Comorbidity, including diabetes mellitus (type I or II), ischaemic heart disease, cerebrovascular disease and depression (self-reported).
- Patient and clinic reported measures of disease activity and progression were collected, following their standardised procedures as follows:
 - Bath Ankylosing Spondylitis Disease Activity Index (BASDAI): reported on a scale of 1-10 with higher scores representing higher disease activity (Garrett et al., 1994).
 - Bath Ankylosing Spondylitis Metrology Index (BASMI): reported on a scale of 1-10 with higher scores indicating greater spinal impairment (Jenkinson et al., 1994).
 - Bath Ankylosing Spondylitis Functional Index (BASFI): reported on a scale of 1-10 with higher scores indicating greater functional impairment (Calin et al., 1994).
 - Ankylosing Spondylitis Quality of Life (ASQoL): reported on a scale of 1-18 with higher scores indicating poorer quality of life (Doward et al., 2003).
 - Health Assessment Questionnaire for the Spondyloarthropathies (HAQ-s): reported on a scale of 0-3, with higher score indicating greater disability (Daltroy et al., 1990).
- AxSpA disease duration was defined as the time between first AxSpA symptom onset and time of joining the ASRI AxSpA registry, measured in years.
- Smoking status was self-reported by participants and was categorised as *current*, *past* or *never*.

- Alcohol consumption was self-reported and categorised as current, past or never.
- Employment status was self-reported and categorised by industry.
- Overweight was defined as body mass index (BMI) $\geq 25 \text{Kg/m}^2$. Obesity was defined as BMI $\geq 30 \text{Kg/m}^2$. Patients who met the criterion for obesity were also considered as overweight in our analyses.
- Central obesity was defined as a waist circumference greater than 94cm in men and greater than 85cm in women.

2.3.2 Statistical analysis

The data collected was formatted for analysis using Microsoft Excel. Statistical analyses, including Mann Whitney U or Kruskal-Wallis tests for continuous data and Fisher's exact test for proportions, were performed using Minitab V21.1. Regression analysis, including logistic regression, binary logistic regression, and multivariate regression, were also performed in Minitab V21.1. Results from bilateral analyses and multiple regression models considered $P < 0.05$ as a significant result. All graphical representations of the data were produced using Minitab V21.1.

2.4 Results

2.4.1 Demographics of the ASRI cohort

At the time of analysis, the ASRI AxSpA study had registered 912 patients. A characterization of the demographics of this patient cohort is presented in Table 2.1. Of the participants, 75.1% were male, which is consistent with previous AS studies (Braun and Sieper, 2007). The median age of the participants was 44 years (IQR 36-55). Most participants (694, 76.1%) were diagnosed with AS according to the modified New York criteria. 84.9% of participants with available data were HLA-B27 positive.

The median age at onset of AxSpA symptoms was 25 years, and the median age at AxSpA diagnosis was 34 years. The delay in diagnosis averaged 6 (IQR 2-11) years. The median AxSpA disease duration from the onset of the first symptom was 16 years.

Uveitis (34.4%), peripheral arthritis (30.9%), psoriasis (16.8%), IBD (10.7%), and depression (10.4%) were the most commonly reported EAMs and comorbidities in the cohort.

A total of 531 (58.2%) of patients were receiving anti-TNF DMARD therapy, while a further 14 (1.5%) were currently treated with IL-17 pathway inhibitors. 461 (51.1%) patients were receiving NSAID therapy. A total of 275 (30.2%) were currently receiving multiple drug therapies.

Table 2.1: General characteristics of the ASRI cohort.

Demographics and disease characterisation	Included n	Median (IQR) or n(%)	Range
Male, n (%)	912	75.1	--
Age in years, median (IQR)	912	44 (36 - 55)	18 - 85
Age at AxSpA symptom onset, median (IQR)	906	25 (19 - 33)	7 - 78
Age at diagnosis, median (IQR)	906	34 (26 - 42)	9 - 80
Delay to diagnosis in years, median (IQR)	906	6 (2 - 11)	0 - 52
Disease duration in years, median (IQR)	912	16 (9 - 27)	0 - 57
ASAS criteria for AxSpA, n (%)	912	912 (100)	--
mNY criteria for AS, n (%)	912	694 (76.1)	--
HLA-B27 positive, n (%)	863	733 (84.9)	--
Clinical measures of disease activity and severity			
BASDAI score, median (IQR)	912	3.7 (1.9 - 5.8)	0.0 - 10.0
BASMI score, median (IQR)	912	3.6 (2.4 - 5.6)	0.4 - 9.0
BASFI score, median (IQR)	912	3.2 (1.3 - 5.6)	0.0 - 10.0
ASQoL score, median (IQR)	912	5.0 (1.0 - 11.0)	0.0 - 18.0
HAQ-s score, median (IQR)	912	0.4 (0.0 - 0.8)	0.0 - 3.0
Extra-axial manifestations and comorbidity			
Uveitis, n (%)	895	308 (34.4)	--
Psoriasis, n (%)	895	150 (16.8)	--
IBD ¹ , n (%)	897	96 (10.7)	--
Peripheral arthritis, n (%)	894	276 (30.9)	--
Dactylitis, n (%)	897	60 (6.7)	--
Osteoporosis, n (%)	905	46 (5.1)	--
Depression, n (%)	905	94 (10.4)	--
Current therapies			
Anti-TNF, n (%)	912	531 (58.2)	--
IL-17 pathway inhibitor, n (%)	912	14 (1.5)	--
Methotrexate, n (%)	903	48 (5.3)	--
Sulfasalazine, n (%)	903	33 (3.7)	--
NSAID, n (%)	903	461 (51.1)	--
Multiple therapies, n (%)	912	275 (30.2)	--
No current therapy, n (%)	912	127 (13.9)	--

¹ Defined as ulcerative colitis or Crohn's disease.

2.4.2 The effect of disease duration on severity and extra-articular manifestations

To investigate the relationship between disease duration and the severity of AxSpA, we conducted both univariate and multivariate logistic regression analyses of each patient reported outcome (PRO) (BASDAI, BASFI, BASMI, ASQoL, and HAQ-s) against AxSpA disease duration. The multivariate analysis included gender at birth, BMI, smoking, and HLA-B27 status as covariates. Our results showed that in the univariate analysis, BASMI had the strongest correlation with AxSpA disease duration, followed by BASFI and HAQ-s, while no significant association was observed between BASDAI or ASQoL and disease duration (Figs 2.1A-E). Our multivariate analysis revealed a similar pattern of association across all PROs (Table 2.2).

Interestingly, we noted a substantial difference in the relative strength of association between BASMI and BASFI with disease duration (Figs 2.1B, 2.1C). Poor correlation of these measurements has been previously reported (Zochling, 2011). To confirm this within the ASRI cohort, we conducted a correlation analysis between BASMI and BASFI (Fig. 2.1F). Our findings ($r^2=24.8\%$, Fig. 2.1F) were in keeping with previously published data. Importantly, BASMI and BASFI specifically target different aspects of disability: BASMI is focussed on articular limitation, while BASFI captures ability to engage in daily tasks.

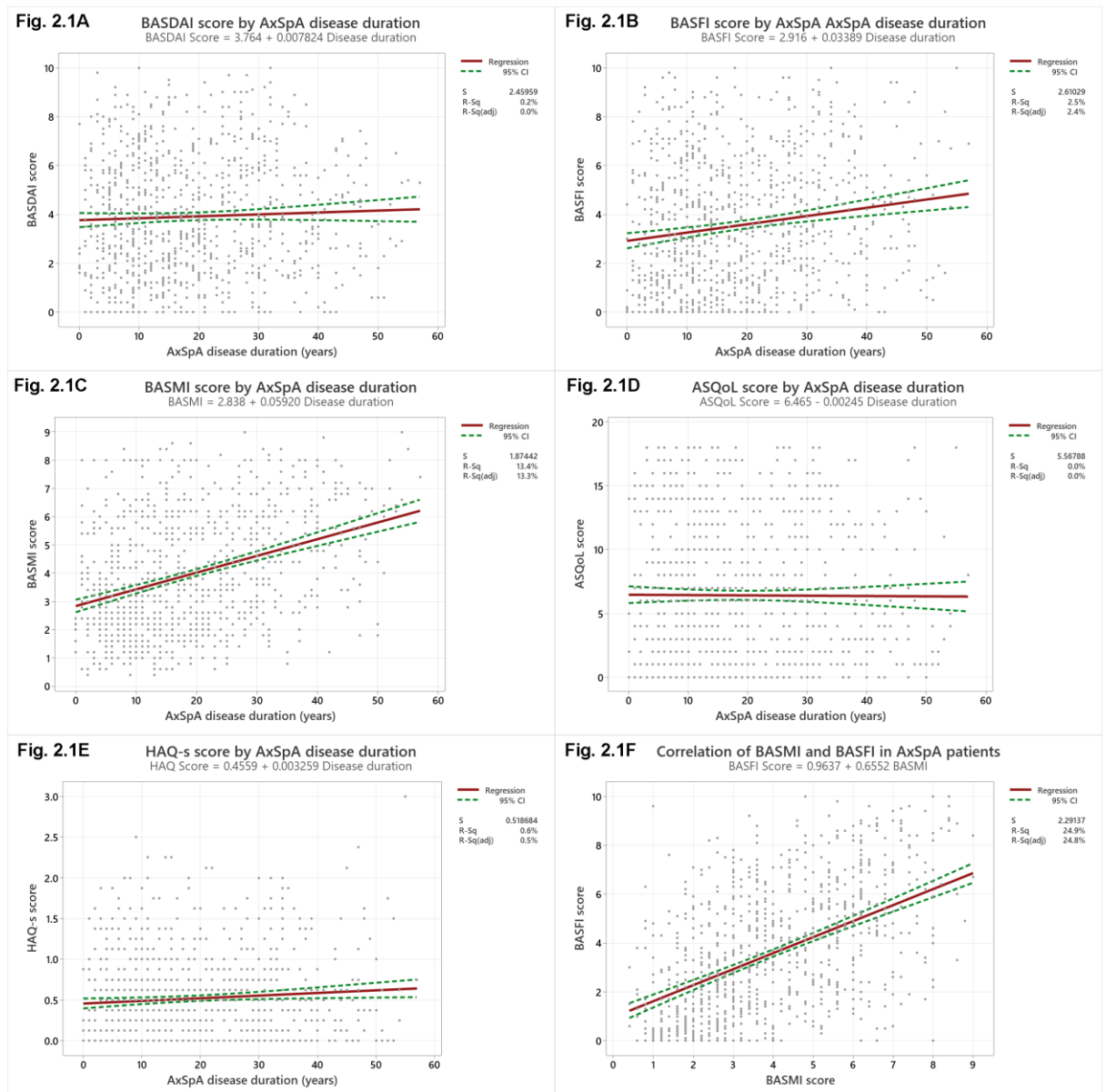


Fig. 2.1A-E: Univariate regression analysis of patient reported outcomes by AxSpA disease duration. Each of BASMI ($r^2=13.3\%$, $P<0.001$), BASFI ($r^2=2.4\%$, $P<0.001$), and HAQ-s ($r^2=0.5\%$, $P=0.019$) were associated with AxSpA disease duration. Neither BASDAI nor ASQoL showed any correlation.

Fig. 2.1F: Correlation between BASMI and BASFI, $r=49.9\%$, $r^2=24.8\%$)

Table 2.2: Univariate and multivariate regression of clinical outcomes against AxSpA disease duration.

	Univariate	Multivariate
Clinical measurement	P	p ¹
BASDAI	0.233	0.18
BASFI	<0.001	<0.001
BASMI	<0.001	<0.001
ASQoL	0.869	0.883
HAQ-s	0.019	0.029

¹Gender at birth, BMI, smoking and HLA-B27 status included as covariates

To describe the relationship between EAMs and AxSpA disease duration in the absence of dates of onset for these manifestations, we used binary logistic regression models to determine the likelihood of an event (EAM) with respect to AxSpA disease duration in the ASRI cohort (Figs 2.2A-E, Table 2.3). To account for covariates, we further performed multivariate analysis, accounting for gender at birth, HLA-B27 status, BMI and smoking status (Table 2.3). Additionally, to confirm the accuracy of our univariate models, a bar chart of uveitis history in ASRI patients by AxSpA disease duration was created (Fig. 2.2F), which closely tracked the mean probability of uveitis as determined by our binary logistic regression model throughout all age ranges tested (Figs 2.2A, 2.2F).

Our results indicate that uveitis and peripheral arthritis are strongly correlated with AxSpA disease duration, with a higher prevalence observed as disease duration is increased (Figs. 2.2A, 2.2D, Table 2.3). In contrast, psoriasis and dactylitis appear more likely to coincide with or precede AxSpA onset and do not show any association with AxSpA disease duration (Figs. 2.2B, 2.2E, Table 2.3). While IBD was found to increase with AxSpA disease duration, the small number of affected patients within the ASRI cohort limited our ability to accurately model the rate of increase (Fig. 2.2C, Table 2.3).

Our data suggests that approximately 20% of AxSpA patients experience at least one episode of uveitis prior to or within one year of AxSpA symptom onset, and approximately 8% of patients experience their first uveitis episode per decade post-AxSpA onset. Interestingly, the prevalence of uveitis in our cohort at 30 years AxSpA disease duration matched that reported by Zeboulon et al. (2008). However, contrary to the findings of Zeboulon et al. (2008), we did not observe a plateau of uveitis first episodes at 20 years post-AxSpA onset. Instead the prevalence of uveitis continued to trend upward to 50% at after 40 years of AxSpA disease duration. Peripheral arthritis also appears to precede or coincide with AxSpA in approximately 20% of cases, with

prevalence increasing by 5% per decade thereafter. We were surprised to find that dactylitis did not follow a similar pattern of onset, given its close association with peripheral arthritis (Singh et al., 2008). However, to our knowledge no previous studies have addressed the relationship between dactylitis and AxSpA disease duration.

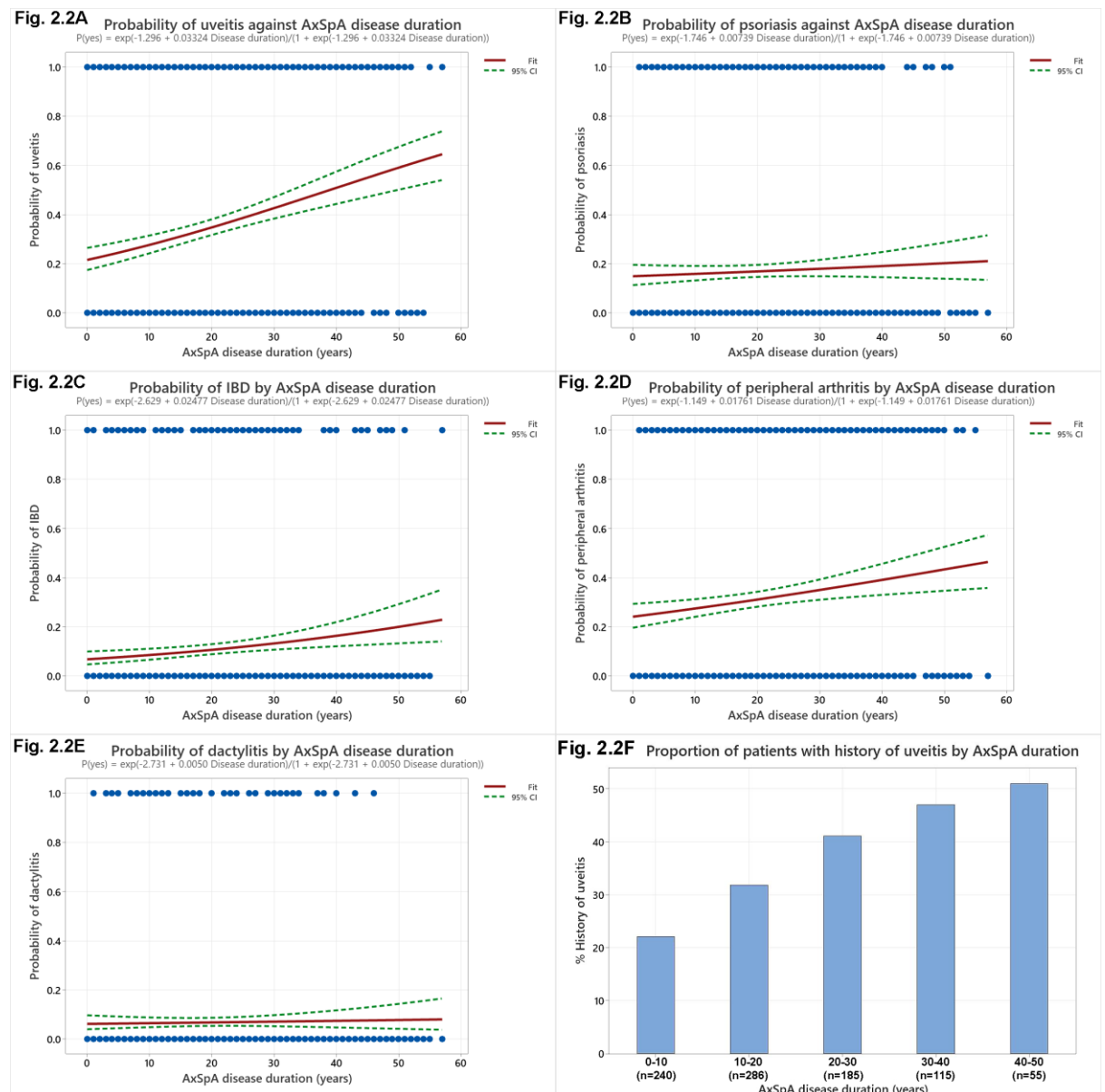


Fig. 2.2A-E: Univariate binary logistic regression of extra-axial manifestations by AxSpA disease duration. Uveitis, IBD and peripheral arthritis are increased with AxSpA disease duration. No relationship between either psoriasis or dactylitis and AxSpA disease duration was observed.

Fig. 2.2F: Bar chart of prevalence of uveitis history in ASRI patients by AxSpA disease duration.

Table 2.3: Univariate and multivariate regression of extra-spinal manifestations against AxSpA disease duration.

	Univariate	Multivariate
Extra-axial manifestation	P	p ¹
Uveitis	<0.001	<0.001
Psoriasis	0.298	0.249
IBD	0.003	0.003
Peripheral arthritis	0.002	0.001
Dactylitis	0.64	0.336

¹Gender at birth, BMI, smoking and HLA-B27 status used as covariates.

2.4.3 Gender differences in presentation

The ASRI cohort included 685 men and 227 women (Table 2.4). Women were younger than men (41 vs 46 years, $P=0.005$), but no difference was found by sex in age at first symptom onset, age at AxSpA diagnosis or delay to diagnosis. Men were more likely to meet the modified New York criteria for AS (78.8% vs 67.8%, $P=0.002$) and had longer disease duration (17 vs 15 years, $P=0.044$) due to their increased median age. HLA-B27 was slightly more prevalent in men (85.8% vs 82.3%), but this was not statistically significant.

Regarding disease outcomes, decreased spinal mobility, as measured by BASMI, was noted in men, but increased indicators of disease activity and reduced quality of life were found in women (measured by BASDAI, HAQ-s, and ASQoL, respectively), which is consistent with previous findings from the ASRI study (Maguire et al., 2022d).

Peripheral arthritis and dactylitis were more likely in women ($P=0.019$ and $P=0.001$, respectively), consistent with previous findings from the RESPONDIA registry (Landi et al., 2016). However, contrary to other studies, AAU was more common in women than in men (41.2% vs 32.1%, $P=0.015$). This was particularly unexpected given AAU's positive correlation with AxSpA disease duration (Fig. 2.2A), which is shorter in women than men (Table 2.4). IBD was also markedly more frequent in women (18.6% vs 8%, $P<0.001$), but no difference was found in the frequency of psoriasis, osteoporosis, or self-reported depression between men and women.

A higher proportion of men were overweight, but this was not statistically significant (68.5% vs 62.2%, $P=0.084$), and no difference was found in the proportion of patients who were obese or centrally obese between sexes. Hypertension and diabetes mellitus were more common in men (23.1% vs 17.3% $P=0.76$ and 5.9% vs 2.7% $P=0.056$, respectively), but these differences did not

reach statistical significance. Men were more likely to be smokers (30.5% vs 22%, $P=0.014$) and unemployed ($P=0.028$), as previously reported in the ASRI cohort (Maguire et al., 2022c).

Therapeutic choice was generally consistent between male and female patients, with a higher proportion of women receiving NSAID therapy (56.7% vs 49.2%), however, this also failed to achieve statistical significance ($P=0.054$).

Table 2.4: General characteristics of the ASRI cohort with comparison of male and female subgroups.

Demographics and disease characterisation	Included n (male)	Male Median (IQR) or n (%)	Included n (female)	Female Median (IQR) or n (%)	P (male vs female)*	Multiple regression**
Total participants	685	100.0	227	0.0	--	--
Age in years, median (IQR)	685	46 (37 - 55)	227	41 (35 - 51)	0.005	--
Age at AxSpA symptom onset, median (IQR)	681	24 (19 - 33)	225	25 (20 - 30)	0.929	--
Age at diagnosis, median (IQR)	681	34 (26 - 42)	225	34 (27 - 41)	0.991	--
Delay to diagnosis in years, median (IQR)	681	6 (2 - 11)	225	5 (3 - 13)	0.58	--
Disease duration in years, median (IQR)	685	17 (9 - 29)	227	15 (9 - 25)	0.044	--
ASAS criteria for AxSpA, n (%)	685	685 (100)	227	227 (100)	1	--
mNY criteria for AS, n (%)	685	540 (78.8)	227	154 (67.8)	0.002	0.004
HLA-B27 positive, n (%)	648	556 (85.8)	215	177 (82.3)	0.237	--
Clinical measures of disease severity						
BASDAI score, median (IQR)	685	3.4 (1.8 - 5.5)	227	4.6 (2.4 - 6.3)	<0.001	0.001
BASMI score, median (IQR)	685	3.8 (2.4 - 5.8)	227	3.0 (2.0 - 4.4)	<0.001	<0.001
BASFI score, median (IQR)	685	3.1 (1.2 - 5.5)	227	3.4 (1.7 - 5.9)	0.095	--
AsQoL score, median (IQR)	685	5.0 (1.0 - 10.0)	227	7.0 (2.0 - 13.0)	<0.001	<0.001
HAQ-s score, median (IQR)	685	0.4 (0.0 - 0.8)	227	0.5 (0.3 - 0.9)	0.002	0.013
EAM's and comorbidity						
Uveitis, n (%)	669	215 (32.1)	226	93 (41.2)	0.015	0.007
Psoriasis, n (%)	670	112 (16.7)	225	38 (16.9)	1	--
IBD ¹ , n (%)	671	54 (8.0)	226	42 (18.6)	<0.001	<0.001
Peripheral arthritis, n (%)	671	193 (28.8)	223	83 (37.2)	0.019	0.017
Dactylitis, n (%)	671	33 (4.9)	226	27 (11.9)	0.001	0.001
Osteoporosis, n (%)	679	31 (4.6)	226	15 (6.6)	0.223	--
Depression, n (%)	679	68 (10.0)	226	26 (11.5)	0.53	--
Cardiovascular health						
Overweight ² , n (%)	664	456 (68.7)	222	138 (62.2)	0.084	--
Obese, n (%)	664	194 (29.2)	222	62 (27.9)	0.797	--
Central obesity, n (%)	685	396 (57.8)	227	132 (58.1)	0.938	--
Hyperlipidaemia, n (%)	679	112 (16.5)	226	31 (13.7)	0.345	--
Hypertension, n (%)	679	157 (23.1)	226	39 (17.3)	0.076	--
Ischaemic heart disease, n (%)	679	23 (3.4)	226	11 (4.9)	0.315	--
Cerebrovascular disease, n (%)	679	14 (2.1)	226	1 (0.4)	0.134	--
Diabetes mellitus (type I or II), n (%)	679	40 (5.9)	226	6 (2.7)	0.056	--
Lifestyle						
Current smoker, n (%)	685	209 (30.5)	227	50 (22.0)	0.014	--
Current consumer of alcohol, n (%)	685	498 (72.7)	227	159 (70.0)	0.444	--
Gainful employment ² , n (%)	586	409 (69.8)	155	122 (78.7)	0.028	--
Current therapies						
Anti-TNF, n (%)	685	406 (59.3)	227	125 (55.1)	0.278	--
IL-17 pathway inhibitor, n (%)	685	11 (1.6)	227	3 (1.3)	1	--
Methotrexate, n (%)	679	37 (5.4)	224	11 (4.9)	0.864	--
Sulfasalazine, n (%)	679	22 (3.2)	224	11 (4.9)	0.249	--
NSAID, n (%)	679	334 (49.2)	224	127 (56.7)	0.054	--
Multiple therapies, n (%)	685	205 (29.9)	227	70 (30.8)	1	--
No current therapy, n (%)	685	100 (14.6)	227	27 (11.9)	0.376	--

¹ Defined as ulcerative colitis or Crohn's disease.² Data excludes retirees, students and homemakers

* Statistical significance calculated by Mann Whitney U or chi-square test of two proportions (Fisher's exact).

** Multiple regression, based on a binary outcome model, using gender, HLA-B27 and AxSpA disease duration as covariates.

To investigate the differences in AAU and IBD between men and women, we performed regression analysis of these EAMs against AxSpA disease duration, split by gender (Fig. 2.3A,B). Our analysis revealed that AAU was positively correlated with AxSpA disease duration in both men ($P<0.001$) and women ($P=0.004$). Interestingly, we observed an increased likelihood of AAU coinciding with or preceding the onset of AxSpA symptoms in women, even though both men and women exhibited a similar rate of first AAU episodes after AxSpA symptom onset.

IBD onset followed a similar pattern, and was linked to AxSpA disease duration in both men ($P=0.034$) and women ($P=0.007$). Our analysis also suggested that women were more likely to have IBD prior to AxSpA symptom onset, and that women developed IBD at a higher rate than men after AxSpA onset within our cohort. Our sample size of patients with IBD (54 men and 42 women) was not sufficient to confirm this observation.

Due to the high burden of uveitis in women within the ASRI cohort, we aimed to investigate whether uveitis may be associated with worse disease outcomes in women. Therefore, we compared disease characteristics, patient-reported outcomes, and EAMs between women with and without a history of uveitis (Table 2.5). We noted no difference in the ages of women with and without AAU. However, women with a history of AAU generally had longer AxSpA disease duration (19 vs 12 years, $P<0.001$), and HLA-B27 was markedly enriched in women with a history of uveitis (95.4% vs 73.2%, $P<0.001$). We found that HLA-B27 was also enriched in men with a history of uveitis, but to a lesser extent (92.3% vs 82.4%, $P=0.001$).

In terms of patient-reported outcomes, women with a history of uveitis had reduced spinal mobility as measured by BASMI score (3.6 vs 2.8, $P=0.019$), while a higher proportion also met the modified New York criteria for AS. However, these were accounted for by disease duration in multiple regression analysis. Interestingly, women with a history of uveitis were more likely to have concomitant dactylitis (18.3% vs 7.5%, $P=0.021$), which remained statistically significant ($P=0.027$) after AxSpA disease duration was accounted for, supporting our previous observation that dactylitis occurs early in AxSpA disease (Fig. 2.2E). Additionally, psoriasis (20.4% vs 13.7% $P=0.182$) and peripheral arthritis (44% vs 32.6% $P=0.123$) were increased in women with a history of uveitis, although these did not reach statistical significance.

These findings suggest that HLA-B27 and the presence of dactylitis at onset are significant risk factors for uveitis in women with AxSpA. It is furthermore likely that uveitis is linked with other EAMs in women.

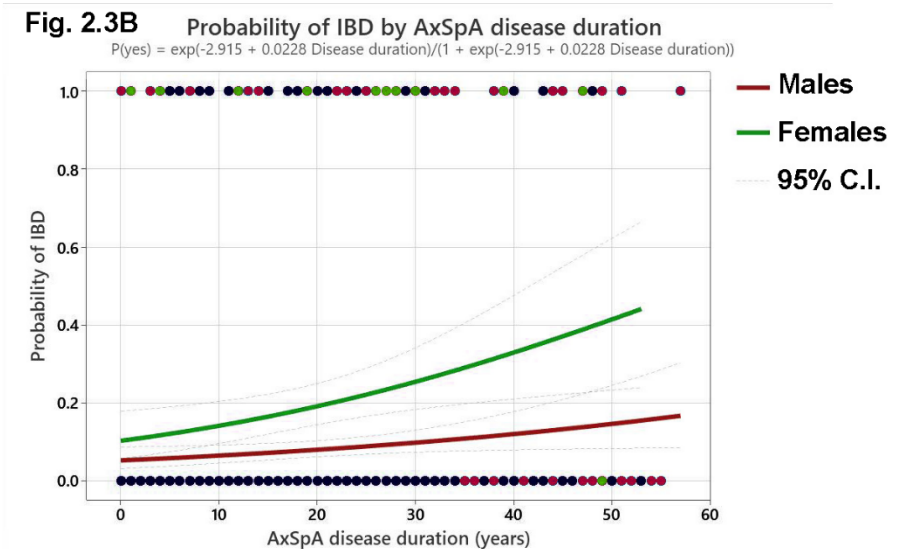
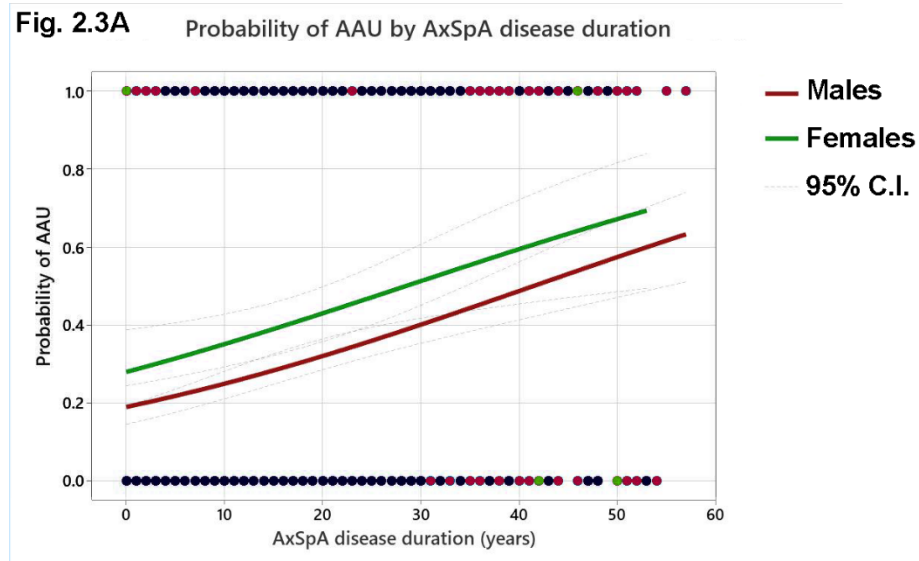


Fig. 2.3: A model of extra-musculoskeletal manifestations with AxSpA disease duration in males and females.

Fig. 2.3A: Relationship between AAU and AxSpA disease duration from first onset in men ($P < 0.001$) and women ($P = 0.004$).

Fig. 2.3B: Relationship between IBD and AxSpA disease duration from first onset in men ($P < 0.034$) and women ($P = 0.007$).

Table 2.5: A comparison of women with and without history of acute anterior uveitis

Characteristic	Women with uveitis		Women without uveitis		P*	Multivariate regression P**
	n	Median (IQR) or n (%)	n	Median (IQR) or n (%)		
Age in years, median (IQR)	93	42 (37 - 52)	133	41 (34 - 51)	0.111	--
Age at AxSpA symptom onset, median (IQR)	92	25 (19 - 28)	132	26 (20 - 33)	0.189	--
AxSpA disease duration in years, median (IQR)	93	19 (11 - 27)	133	12 (6 - 22)	<0.001	--
mNY criteria for AS, n (%)	93	68 (73.1)	133	85 (63.9)	0.152	0.333
HLA-B27 positive, n (%)	87	83 (95.4)	127	93 (73.2)	<0.001	<0.001
Clinical measurements of disease severity						
BASDAI score, median (IQR)	93	4.6 (2.6 - 6.2)	133	4.7 (2.1 - 6.4)	0.862	0.973
BASMI score, median (IQR)	93	3.6 (2.4 - 4.8)	133	2.8 (2.0 - 3.8)	0.019	0.236
BASFI score, median (IQR)	93	3.5 (2.1 - 5.7)	133	3.3 (1.3 - 5.9)	0.396	0.946
AsQoL score, median (IQR)	93	7.0 (2.0 - 12.0)	133	8.0 (2.0 - 13.0)	0.602	0.638
HAQ-s score, median (IQR)	93	0.6 (0.3 - 0.8)	133	0.5 (0.1 - 0.9)	0.776	0.533
EAM's						
IBD ¹ , n (%)	93	20 (21.5)	133	22 (16.5)	0.387	0.631
Psoriasis, n (%)	93	19 (20.4)	131	18 (13.7)	0.204	0.182
Peripheral arthritis, n (%)	91	40 (44.0)	132	43 (32.6)	0.092	0.123
Dactylitis, n (%)	93	17 (18.3)	133	10 (7.5)	0.021	0.027

* Statistical significance calculated by Mann Whitney U or chi-square test of two proportions (Fisher's exact).

** Multivariate regression including AxSpA disease duration in years as a covariate.

2.4.4 The effects of extra-axial manifestations on AxSpA disease severity

We investigated the relationship between disease severity and the most common extra-articular manifestations (EAMs) in our cohort, including AAU, psoriasis, and peripheral arthritis. To account for potential biases in patient-reported outcomes (PROs) and EAMs, we additionally conducted multivariate analyses that incorporated several covariates, including AxSpA disease duration, sex, HLA-B27 status, smoking, and BMI.

Uveitis is the most common EAM within the ASRI cohort, reported in 34.4% of patients. We have previously established that uveitis is associated with AxSpA disease duration (Fig 2.2A), female gender (Fig. 2.3A) and HLA-B27 (Table. 2.5). A history of uveitis was linked with earlier AxSpA disease onset, and positively associated with several other EAMs, including psoriasis, IBD, peripheral arthritis, and dactylitis (Table 2.6). After accounting for covariates, a history of uveitis was most strongly associated with the presence of dactylitis (11.4% vs. 4.1%, $P < 0.001$). However, there was no significant difference in PROs between patients who had experienced episodes of uveitis and those who had not, and no evidence to suggest that episodes of uveitis influenced clinical decision-making for patient treatment. Despite the increased prevalence of HLA-B27 in patients with uveitis, no relationship was observed between uveitis and radiographic disease progression.

Overall, we found that uveitis is associated with increased prevalence of other EAMs and earlier AxSpA onset. However, we observed no link with spinal mobility or PROs for disease severity and quality of life. Additionally, there is no indication that a history of uveitis in any way influences clinical decision making in treatment.

Table 2.6: Characteristics of AxSpA patients with and without concomitant uveitis.

	AxSpA with uveitis		AxSpA without uveitis		P (with vs without uveitis)*	Multiple regression**
	Included n	Median (IQR) or n (%)	Included n	Median (IQR) or n (%)		
Demographics and disease characterisation						
Male sex, n (%)	308	215 (69.8)	587	454 (77.3)	0.015	--
Age in years, median (IQR)	308	47 (39 - 56)	587	43 (35 - 53)	<0.001	--
Age at AxSpA symptom onset, median (IQR)	304	24 (19 - 30)	585	25 (20 - 34)	0.007	--
Age at diagnosis, median (IQR)	304	32 (27 - 40)	585	34 (27 - 43)	0.137	--
Delay to diagnosis in years, median (IQR)	304	6 (2 - 12)	585	5 (2 - 11)	0.41	--
Disease duration in years, median (IQR)	308	21 (12 - 31)	587	14 (8 - 24)	<0.001	--
mNY criteria for AS, n (%)	308	230 (74.7)	587	448 (76.3)	0.622	--
HLA-B27 positive, n (%)	296	276 (93.2)	553	444 (80.3)	<0.001	--
Clinical measures of disease severity						
BASDAI score, median (IQR)	308	3.6 (1.9 - 5.6)	587	3.7 (1.9 - 5.8)	0.617	0.691
BASMI score, median (IQR)	308	3.8 (2.6 - 5.4)	587	3.6 (2.2 - 5.6)	0.265	0.773
BASFI score, median (IQR)	308	3.2 (1.3 - 5.5)	587	3.2 (1.3 - 5.6)	0.869	0.871
AsQoL score, median (IQR)	308	5.0 (1.0 - 11.0)	587	5.0 (1.0 - 11.0)	0.252	0.692
HAQ-s score, median (IQR)	308	0.4 (0.0 - 0.8)	587	0.4 (0.0 - 0.8)	0.549	0.601
EAM's and comorbidity						
Psoriasis, n (%)	307	61 (19.9)	582	86 (14.8)	0.058	0.035
IBD ¹ , n (%)	305	45 (14.8)	585	50 (8.5)	0.006	0.028
Peripheral arthritis, n (%)	303	113 (37.3)	580	154 (26.6)	0.001	0.011
Dactylitis, n (%)	307	35 (11.4)	586	24 (4.1)	<0.001	<0.001
Depression, n (%)	304	33 (10.9)	584	58 (9.9)	0.727	0.687
Current therapies						
Anti-TNF, n (%)	308	176 (57.1)	587	343 (58.4)	0.722	--
Methotrexate, n (%)	302	20 (6.6)	584	26 (4.5)	0.201	--
Sulfasalazine, n (%)	302	11 (3.6)	584	21 (3.6)	1	--
NSAID, n (%)	302	160 (53.0)	584	293 (50.2)	0.436	--
Multiple therapies, n (%)	308	87 (28.2)	587	180 (30.7)	0.489	--
No current therapy, n (%)	308	38 (12.3)	587	86 (14.7)	0.361	--

¹ Defined as ulcerative colitis or Crohn's disease.

* Statistical significance calculated by Mann Whitney U or chi-square test of two proportions (Fisher's exact).

** Multiple regression, based on a binary outcome model, using gender, HLA-B27, AxSpA disease duration, smoking status and BMI as covariates.

Within the ASRI cohort, 16.8% of AxSpA patients reported having psoriasis. Our analysis revealed that patients with psoriasis were not distinguished by age at AxSpA symptom onset or by sex, as shown in Table 2.7. While a greater proportion of patients with concomitant psoriasis met the modified New York criteria for AS (80.7% vs. 74.9%), this difference was not statistically significant. HLA-B27 was less prevalent in patients with psoriasis (80.4 vs. 85.7%), but again, this difference was not statistically significant.

However, psoriasis was found to be strongly associated with peripheral arthritis or dactylitis ($P < 0.001$) and positively associated with uveitis, as discussed in the previous section. No significant link was found between psoriasis and IBD or depression after accounting for covariates.

Interestingly, patients with concomitant psoriasis were more likely to be treated with anti-TNF therapy (70.7% vs. 55.6%, $P < 0.001$) and less likely to be untreated (8.7% vs. 14.9%, $P = 0.03$) than patients without psoriasis. However, we found no significant differences in PROs related to the presence of psoriasis.

Table 2.7: Characteristics of AxSpA patients with and without concomitant psoriasis.

	AxSpA with psoriasis		AxSpA without psoriasis		P (with vs without psoriasis)*	Multiple regression**
	Included n	Median (IQR) or n (%)	Included n	Median (IQR) or n (%)		
Demographics and disease characterisation						
Male sex, n (%)	150	112 (74.7)	745	558 (74.9)	1	--
Age in years, median (IQR)	150	44 (36 - 55)	745	45 (36 - 55)	0.973	--
Age at AxSpA symptom onset, median (IQR)	150	24 (18 - 31)	739	25 (20 - 33)	0.197	--
Age at diagnosis, median (IQR)	150	34 (26 - 40)	739	34 (27 - 43)	0.262	--
Delay to diagnosis in years, median (IQR)	150	5 (2 - 12)	739	6 (2 - 11)	0.858	--
Disease duration in years, median (IQR)	150	19 (10 - 29)	745	16 (9 - 27)	0.232	--
mNY criteria for AS, n (%)	150	121 (80.7)	745	558 (74.9)	0.144	--
HLA-B27 positive, n (%)	138	111 (80.4)	711	609 (85.7)	0.121	--
Clinical measures of disease severity						
BASDAI score, median (IQR)	150	4.0 (1.8 - 6.3)	745	3.6 (1.9 - 5.7)	0.254	--
BASMI score, median (IQR)	150	3.6 (2.6 - 5.4)	745	3.6 (2.2 - 5.6)	0.507	--
BASFI score, median (IQR)	150	3.1 (1.4 - 5.6)	745	3.2 (1.3 - 5.6)	0.528	--
AsQoL score, median (IQR)	150	7.0 (1.0 - 12.0)	745	5.0 (1.0 - 11.0)	0.168	--
HAQ-s score, median (IQR)	150	0.4 (0.0 - 0.9)	745	0.4 (0.0 - 0.8)	0.808	--
EAM's and comorbidity						
Uveitis, n (%)	147	61 (41.5)	742	246 (33.2)	0.058	0.035
IBD ¹ , n (%)	147	23 (15.6)	742	72 (9.7)	0.04	0.082
Peripheral arthritis, n (%)	148	68 (45.9)	736	202 (27.4)	<0.001	<0.001
Dactylitis, n (%)	147	21 (14.3)	744	38 (5.1)	<0.001	<0.001
Depression, n (%)	148	23 (15.5)	740	68 (9.2)	0.026	0.051
Current therapies						
Anti-TNF, n (%)	150	106 (70.7)	745	414 (55.6)	0.001	<0.001
Methotrexate, n (%)	148	12 (8.1)	738	35 (4.7)	0.107	--
Sulfasalazine, n (%)	148	5 (3.4)	738	27 (3.7)	1	--
NSAID, n (%)	148	72 (48.6)	738	380 (51.5)	0.53	0.549
Multiple therapies, n (%)	150	54 (36.0)	745	214 (28.7)	0.079	0.064
No current therapy, n (%)	150	13 (8.7)	745	111 (14.9)	0.051	0.03

¹ Defined as ulcerative colitis or Crohn's disease.

* Statistical significance calculated by Mann Whitney U or chi-square test of two proportions (Fisher's exact).

** Multiple regression, based on a binary outcome model, using gender, HLA-B27, AxSpA disease duration, smoking status and BMI as covariates.

In our study, 30.9% of participants reported peripheral arthritis. As previously noted in Sections 2.4.2 and 2.4.3, patients with peripheral arthritis tended to be older, female, and have longer AxSpA disease duration, as shown in Table 2.8. However, there was no significant difference in age at first AxSpA symptom onset, the proportion of patients who were HLA-B27 positive or meeting the mNY criteria for AS between those with and without peripheral arthritis symptoms.

Peripheral arthritis was linked to worse scores in BASDAI (4.4 vs 3.3, $P=0.003$), BASFI (3.7 vs 2.9, $P=0.007$), ASQoL (7.0 vs 4.0, $P=0.003$), and HAQ-s (0.5 vs 0.4, $P<0.001$), after accounting for covariates, indicating increased disease activity, reduced functioning, and poorer quality of life in patients.

Regarding extra-articular manifestations (EAMs), peripheral joint arthritis was linked to all EAMs except IBD, after accounting for covariates. Notably, patients with peripheral arthritis showed a strong association with dactylitis (15.7% vs 2.9%, $P<<0.001$), which was not surprising, given the close association of these phenotypes, as mentioned above.

Table 2.8: Characteristics of AxSpA patients with and without concomitant arthritis.

	AxSpA with arthritis		AxSpA without arthritis		P (with vs without arthritis)*	Multiple regression**
	Included n	Median (IQR) or n (%)	Included n	Median (IQR) or n (%)		
Demographics and disease characterisation						
Male sex, n (%)	276	193 (69.9)	618	478 (77.3)	0.019	--
Age in years, median (IQR)	276	48 (38 - 57)	618	44 (35 - 53)	<0.001	--
Age at AxSpA symptom onset, median (IQR)	274	25 (19 - 33)	614	24 (19 - 33)	0.956	--
Age at diagnosis, median (IQR)	274	35 (25 - 43)	614	33 (27 - 42)	0.884	--
Delay to diagnosis in years, median (IQR)	274	5 (2 - 11)	614	6 (2 - 12)	0.342	--
Disease duration in years, median (IQR)	276	19 (10 - 31)	618	15 (9 - 26)	0.001	--
mNY criteria for AS, n (%)	276	201 (72.8)	618	477 (77.2)	0.176	--
HLA-B27 positive, n (%)	259	215 (83.0)	587	503 (85.7)	0.349	--
Clinical measures of disease severity						
BASDAI score, median (IQR)	276	4.4 (2.2 - 6.3)	618	3.3 (1.8 - 5.6)	<0.001	0.003
BASMI score, median (IQR)	276	3.8 (2.4 - 5.6)	618	3.6 (2.2 - 5.6)	0.153	0.58
BASFI score, median (IQR)	276	3.7 (1.5 - 6.4)	618	2.9 (1.2 - 5.2)	0.001	0.007
AsQoL score, median (IQR)	276	7.0 (2.0 - 13.0)	618	4.0 (1.0 - 10.0)	<0.001	0.003
HAQ-s score, median (IQR)	276	0.5 (0.1 - 1.0)	618	0.4 (0.0 - 0.8)	<0.001	<0.001
EAM's and comorbidity						
Uveitis, n (%)	267	113 (42.3)	616	190 (30.8)	0.01	0.01
Psoriasis, n (%)	270	68 (25.2)	614	80 (13.0)	<0.001	<0.001
IBD ¹ , n (%)	271	38 (14.0)	615	57 (9.3)	0.045	0.276
Dactylitis, n (%)	268	42 (15.7)	618	18 (2.9)	<<0.001	<<0.001
Depression, n (%)	275	39 (14.2)	613	52 (8.5)	0.012	0.025
Current therapies						
Anti-TNF, n (%)	276	171 (62.0)	618	350 (56.6)	0.143	0.132
Methotrexate, n (%)	273	32 (11.7)	612	16 (2.6)	<0.001	<0.001
Sulfasalazine, n (%)	273	16 (5.9)	612	16 (2.6)	0.03	0.032
NSAID, n (%)	273	152 (55.7)	612	297 (48.5)	0.058	0.051
Multiple therapies, n (%)	276	110 (39.9)	618	157 (25.4)	<0.001	<0.001
No current therapy, n (%)	276	30 (10.9)	618	94 (15.2)	0.094	0.102

¹ Defined as ulcerative colitis or Crohn's disease.

* Statistical significance calculated by Mann Whitney U or chi-square test of two proportions (Fisher's exact).

** Multiple regression, based on a binary outcome model, using gender, HLA-B27, AxSpA disease duration, smoking status and BMI as covariates.

Finally, due to the consistent finding of strong dactylitis association with each of the other EAMs, we examined our population based on whether dactylitis was present (Table 2.9). Patients with dactylitis were more likely to be female (45% vs 23.8%, $P=0.003$), as previously identified in Section 2.4.3. However, there were no clear differences in patient age, age at AxSpA symptom onset, or AxSpA disease duration, based on the presence of dactylitis, supporting our previous finding that dactylitis tends to occur early in AxSpA disease (Fig. 2.2E).

Remarkably, patients with dactylitis were much less likely to meet the mNY criteria for diagnosis of AS (58.3% vs 77.1%, $P=0.003$), despite no observable difference in AxSpA disease duration between these groups. While dactylitis patients were more likely to be women, the proportion of patients with dactylitis meeting the mNY criteria was lower than that of female patients overall (58.3%, Table 2.9 vs 67.8%, Table 2.4), although this was not statistically significant. In women, a smaller proportion of those with dactylitis met the mNY criteria (59.3% vs 68.8%), although this was again not statistically significant.

Dactylitis was strongly associated with increased inflammatory disease, as measured by BASDAI (5.6 vs 3.5, $P<0.001$), and lower quality of life, as determined by ASQoL (8.5 vs 5, $P=0.011$). Patients with dactylitis were also more likely to have other extra-axial manifestations, including uveitis (59.3% vs 32.6%, $P<0.001$), psoriasis (35.6% vs 15.1%, $P<0.001$), IBD (25% vs 9.6%, $P=0.001$), and peripheral arthritis of other joints (70% vs 27.4%, $P<0.001$).

Interestingly, depression was not linked to the presence of dactylitis, despite its known association with peripheral arthritis (Table 2.8). However, our available study population of dactylitis patients was small.

Taken together, these findings suggest that patients presenting with dactylitis at AxSpA diagnosis represent a unique sub-cohort who tend to present with reduced radiographic spinal disease, markedly increased inflammation, and a markedly increased risk of peripheral manifestations.

Table 2.9: Characteristics of AxSpA patients with and without concomitant dactylitis.

	AxSpA with dactylitis		AxSpA without dactylitis		P (with vs	
		Median (IQR) or		Median (IQR)	without	Multiple
Demographics and disease characterisation	Included n	n (%)	Included n	or n (%)	dactylitis)*	regression**
Male sex, n (%)	60	33 (55.0)	837	638 (76.2)	0.003	--
Age in years, median (IQR)	60	46 (37 - 55)	837	44 (36 - 55)	0.618	--
Age at AxSpA symptom onset, median (IQR)	59	24 (19 - 31)	832	25 (19 - 33)	0.6	--
Age at diagnosis, median (IQR)	59	36 (26 - 43)	832	34 (27 - 42)	0.355	--
Delay to diagnosis in years, median (IQR)	59	7 (3 - 18)	832	6 (2 - 11)	0.331	--
Disease duration in years, median (IQR)	60	18 (11 - 28)	837	16 (9 - 27)	0.407	--
mNY criteria for AS, n (%)	60	35 (58.3)	837	645 (77.1)	0.003	--
HLA-B27 positive, n (%)	57	46 (80.7)	795	677 (85.2)	0.343	--
Clinical measures of disease severity						
BASDAI score, median (IQR)	60	5.6 (3.7 - 7.1)	837	3.5 (1.8 - 5.7)	<0.001	<0.001
BASMI score, median (IQR)	60	3.3 (2.6 - 5.3)	837	3.6 (2.2 - 5.6)	0.942	0.844
BASFI score, median (IQR)	60	3.9 (1.7 - 6.4)	837	3.1 (1.3 - 5.5)	0.111	0.078
AsQoL score, median (IQR)	60	8.5 (2.8 - 13.0)	837	5.0 (1.0 - 11.0)	0.011	0.01
HAQ-s score, median (IQR)	60	0.6 (0.2 - 1.0)	837	0.4 (0.0 - 0.8)	0.023	0.043
EAM's and comorbidity						
Uveitis, n (%)	59	35 (59.3)	834	272 (32.6)	<0.001	<0.001
Psoriasis, n (%)	59	21 (35.6)	832	126 (15.1)	<0.001	<0.001
IBD ¹ , n (%)	60	15 (25.0)	833	80 (9.6)	0.001	0.006
Peripheral arthritis, n (%)	60	42 (70.0)	826	226 (27.4)	<0.001	<0.001
Depression, n (%)	59	6 (10.2)	831	85 (10.2)	--	--
Current therapies						
Anti-TNF, n (%)	60	34 (56.7)	837	486 (58.1)	0.893	--
Methotrexate, n (%)	59	10 (16.9)	829	37 (4.5)	0.001	<0.001
Sulfasalazine, n (%)	59	3 (5.1)	829	29 (3.5)	--	--
NSAID, n (%)	59	33 (55.9)	829	421 (50.8)	0.501	--
Multiple therapies, n (%)	60	21 (35.0)	837	247 (29.5)	0.383	--
No current therapy, n (%)	60	5 (8.3)	837	119 (14.2)	--	--

¹ Defined as ulcerative colitis or Crohn's disease.

* Statistical significance calculated by Mann Whitney U or Fisher's exact.

** Multiple regression, based on a binary outcome model, using gender, HLA-B27, AxSpA disease duration, smoking status and BMI as covariates.

2.4.5 Investigating potential outcomes of obesity in past smokers

Within the ASRI cohort, 268 (29.4%) participants were past smokers, while 259 (28.4%) were current smokers and 385 (42.2%) had never smoked. We examined demographics, patient and clinic reported outcomes, EAMs and cardiovascular disease in patients with respect to smoking (Table 2.10). Current smokers scored worse in all of BASDAI, BASFI, BASMI, ASQoL and HAQ-s and were more likely to meet the mNY criteria for AS. Additionally, an inverse association between current smoking and both uveitis and peripheral arthritis was observed, which is consistent with previous studies (Jones et al., 2017). Depression was more common in current smokers than in never or past smokers ($P<0.01$, $P=0.019$, respectively). Current smokers were less likely to be obese or centrally obese than never smokers ($P<0.001$, $P=0.008$, respectively). Fewer smokers were overweight when compared to never smokers (58.8% vs 66.5%), but this did not reach statistical significance ($P=0.053$). No difference was observed in the frequency of hyperlipidaemia, hypertension, ischaemic heart disease, cerebrovascular disease or diabetes mellitus between current and non-smokers.

In comparison to current smokers, past smokers were older (49 vs. 42 years, $P<0.001$) and were older at first AxSpA symptom onset (26 vs. 23 years, $P=0.004$). This correlated with longer AxSpA disease duration in past smokers (19 vs. 15 years, $P=0.004$). Disease activity, as measured by BASDAI, was improved in past smokers when compared to current smokers ($P=0.035$), with past smokers scoring similarly to never smokers ($P=0.207$). Quality of life, determined by ASQoL, was also improved in past smokers when compared to current smokers ($P=0.017$). This, again, did not differ significantly from never smokers ($P=0.283$). However, spinal mobility was worse in past smokers compared to never smokers ($P<0.001$) and was similar to that of current smokers ($P=0.561$). Functioning, as measured by BASFI, was improved in past smokers with respect to current smokers (median score 3.35 vs. 3.7), although this was not statistically significant ($P=0.397$). Past smokers scored worse on BASFI than never smokers ($P=0.001$).

Uveitis was increased in past smokers when compared to current smokers (35% vs 27.4%), although this was not statistically significant ($P=0.072$). Peripheral arthritis was substantially more common in past smokers than in current smokers (35.5% vs 22.4%, $P=0.001$). Interestingly, this was not reflected in the frequency of dactylitis, which was unaffected by smoking status.

Regarding body weight and composition, central obesity was significantly more common in past smokers (69%) than in current smokers (46.7%, $P=0.003$) and never smokers (57.7%, $P=0.008$). Past smokers were also more likely to be overweight (75.6%) than current smokers (58.8%, $P<0.001$) or never smokers (66.5%, $P=0.017$). Additionally, obesity (BMI ≥ 30) was more

prevalent in past smokers (34.1%) than in current smokers (18%, $P=0.067$) and was comparable to the rate in never smokers (32.4%, not statistically significant).

Furthermore, past smokers had higher rates of hyperlipidemia (22.1%) and hypertension (30%) than current smokers (12%, $P=0.002$ and 15.8%, $P<0.001$, respectively) and never smokers (14.1%, $P=0.011$ and 19.8%, $P=0.004$, respectively). The prevalence of diabetes mellitus was also higher in past smokers (8%) than in never smokers (3.4%, $P=0.012$) and current smokers (4.6%, $P=0.114$), although this difference was not statistically significant.

Table 2.10: Demographics and clinical presentation of patients based on smoking status.

	Never smokers		Current smokers		Past smokers				
		Median (IQR) or	Included	Median (IQR) or	Included	Median (IQR) or	Current vs	Past vs	Past vs
	Included (n)	n (%)	(n)	n (%)	(n)	n (%)	Never (P)*	Never (P)*	Current (P)*
Demographics and disease presentation									
Age in years, median (IQR)	385	43 (35 - 54)	259	42 (35 - 52)	268	49 (41 - 59)	0.19	<0.001	<0.001
Age at AxSpA symptom onset, median (IQR)	381	24 (19 - 32)	258	23 (19 - 30)	267	26 (19 - 35)	0.611	0.004	0.039
Disease duration in years, median (IQR)	385	15 (8 - 26)	259	15 (9 - 26)	268	19 (11 - 31)	0.818	0.004	0.002
mNY criteria for AS, n (%)	385	278 (72.2)	259	215 (83.0)	268	201 (75.0)	0.002	0.472	0.025
Clinical measures of disease severity									
BASDAI score, median (IQR)	385	3.4 (1.8 - 5.5)	259	4.4 (2.2 - 6.2)	268	3.6 (2.1 - 5.8)	0.001	0.207	0.035
BASMI score, median (IQR)	385	3.2 (2.0 - 5.2)	259	3.8 (2.4 - 5.6)	268	4 (2.6 - 5.8)	0.003	<0.001	0.561
BASFI score, median (IQR)	385	2.7 (0.9 - 4.8)	259	3.7 (1.7 - 6.3)	268	3.35 (1.5 - 5.9)	<0.001	0.001	0.397
AsQoL score, median (IQR)	385	4 (1.0 - 10.0)	259	7 (2.0 - 12.0)	268	5 (1.0 - 12.0)	<0.001	0.283	0.017
HAQ-s score, median (IQR)	385	0.4 (0.0 - 0.6)	259	0.5 (0.1 - 0.9)	268	0.4 (0.0 - 0.9)	<0.001	0.052	0.071
EAM's and comorbidity									
Uveitis, n (%)	377	146 (38.7)	252	69 (27.4)	266	93 (35.0)	0.004	0.362	0.072
Psoriasis, n (%)	375	52 (13.9)	254	47 (18.5)	266	51 (19.2)	0.12	0.081	0.911
IBD ¹ , n (%)	379	36 (9.5)	252	25 (9.9)	266	35 (13.2)	0.891	0.16	0.274
Peripheral arthritis, n (%)	375	125 (33.3)	254	57 (22.4)	265	94 (35.5)	0.003	0.612	0.001
Dactylitis, n (%)	377	28 (7.4)	253	14 (5.5)	267	18 (6.7)	0.417	0.759	0.589
Depression, n (%)	383	26 (6.8)	259	43 (16.6)	263	25 (9.5)	<0.001	0.236	0.019
Cardiovascular health									
Overweight, n (%)	372	249 (66.9)	255	150 (58.8)	258	196 (76.0)	0.053	0.017	<0.001
Obese, n (%)	372	122 (32.8)	255	46 (18.0)	258	89 (34.5)	<0.001	0.067	<0.001
Central obesity, n (%)	385	222 (57.7)	259	121 (46.7)	268	185 (69.0)	0.008	0.003	<0.001
Hyperlipidaemia, n (%)	383	54 (14.1)	259	31 (12.0)	263	58 (22.1)	0.477	0.011	0.002
Hypertension, n (%)	383	76 (19.8)	259	41 (15.8)	263	79 (30.0)	0.212	0.004	<0.001
Ischaemic heart disease, n (%)	383	14 (3.7)	259	8 (3.1)	263	12 (4.6)	0.826	0.684	0.495
Cerebrovascular disease, n (%)	383	5 (1.3)	259	5 (1.9)	263	5 (1.9)	0.535	1	1
Diabetes mellitus (type I or II), n (%)	383	13 (3.4)	259	12 (4.6)	263	21 (8.0)	0.533	0.012	0.114
Lifestyle									
Current consumer of alcohol, n (%)	385	260 (67.5)	259	207 (79.9)	268	190 (70.9)	0.001	0.39	0.02
Gainful employment ² , n (%)	312	239 (76.6)	220	143 (65.0)	209	149 (71.3)	0.004	0.184	0.179

*Statistical significance calculated by Mann-Whitney U or Fisher's Exact.

** Protective effect after age and sex are accounted for.

¹ Defined as ulcerative colitis or Crohn's disease.² Data excludes retirees, students and homemakers.

To address the significant age difference between past smokers and current/never smokers, we conducted a multivariate regression analysis with age and sex as covariates for PROs, peripheral arthritis, and comorbidities associated with obesity (Table 2.11). After accounting for age and sex, we found that past smokers had improved results across all clinical measures of disease severity. However, BASFI was increased in past smokers compared to never smokers.

We also found that peripheral arthritis increased following smoking cessation ($P=0.039$), similar to the frequency observed in never smokers ($P=0.867$) after accounting for age and sex. Furthermore, obesity ($P=0.001$) and central obesity ($P<0.001$) were both increased following smoking cessation, with no significant difference in frequency between past and never smokers after accounting for age and gender. Coinciding with this, diabetes mellitus was also increased in past smokers when compared to current smokers, but this was not statistically significant ($P=0.068$).

Our study confirms that weight gain post-smoking cessation in AxSpA patients. Although we observed tentative evidence of increased prevalence of diabetes in past smokers, our study size was insufficient to effectively scrutinize the effects of rapid weight gain in this subgroup. There is evidence to support that the cessation of smoking increases the risk of peripheral arthritis, which is suppressed by smoking activity.

Table 2.11: Multivariate regression of clinical presentation by smoking status

	Never smokers Median (IQR) or n (%)	Current smokers Median (IQR) or n (%)	Past smokers Median (IQR) or n (%)	Past vs Never (P)	Past vs Current (P)
Clinical measures of disease severity					
BASDAI score, median (IQR)	3.4 (1.8 - 5.5)	4.4 (2.2 - 6.2)	3.6 (2.1 - 5.8)	0.453	0.002*
BASMI score, median (IQR)	3.2 (2.0 - 5.2)	3.8 (2.4 - 5.6)	4 (2.6 - 5.8)	0.129	0.026*
BASFI score, median (IQR)	2.7 (0.9 - 4.8)	3.7 (1.7 - 6.3)	3.35 (1.5 - 5.9)	0.046	0.002*
AsQoL score, median (IQR)	4 (1.0 - 10.0)	7 (2.0 - 12.0)	5 (1.0 - 12.0)	0.269	0.002*
HAQ-s score, median (IQR)	0.4 (0.0 - 0.6)	0.5 (0.1 - 0.9)	0.4 (0.0 - 0.9)	0.165	0.003*
EAM's and comorbidity					
Uveitis, n (%)	146 (38.7)	69 (27.4)	93 (35.0)	0.152	0.428
Peripheral arthritis, n (%)	125 (33.3)	57 (22.4)	94 (35.5)	0.867	0.039
Cardiovascular health					
Overweight, n (%)	249 (66.9)	150 (58.8)	196 (76.0)	0.195	0.003
Obese, n (%)	122 (32.8)	46 (18.0)	89 (34.5)	0.859	0.001
Central obesity, n (%)	222 (57.7)	121 (46.7)	185 (69.0)	0.119	<0.001
Diabetes mellitus (type I or II), n (%)	13 (3.4)	12 (4.6)	21 (8.0)	0.068	0.98

Covariates include patient age and sex.

*Protective effect of smoking cessation.

2.5 Discussion

Within this chapter we have described the ASRI AxAxA cohort and further provided robust models to show how PROs develop with AxAxA disease duration. We have clearly demonstrated that disease duration is linked with radiographic spinal progression, as evidenced in BASMI and BASFI, but not disease activity or quality of life. To our knowledge, we are the first group to provide robust models detailing the relationships between EAMs and AxAxA disease duration (Fig. 2.2 A-E). Notably, both psoriasis and dactylitis tend to precede or occur early in AxAxA disease, while AAU, IBD and peripheral arthritis present continuing risk factors in established AxAxA cases.

Additionally, we reveal how ASRI differs between men and women. Women have an increased prevalence of peripheral manifestations, increased disease activity, and reduced quality of life (Table 2.4). Men, meanwhile present with increased spinal disability. Notably, despite these findings, we observed no significant differences in the choice of therapeutics used for men and women. Our data clearly indicates that AxAxA disease presents differently based on gender and suggests that women may benefit from a more tailored therapeutic approach.

Furthermore, we show that EAMs are closely related with disease presentation and outcomes. For example, uveitis is linked with earlier AxAxA disease onset (Table 2.6), while peripheral arthritis and dactylitis are associated with increased disease activity and reduced quality of life (Tables 2.8, 2.9). Moreover, EAMs tend to coincide with each other which may further impact on patient outcomes.

2.5.1 The relationship between disease duration and patient outcomes

Our study revealed that AAU, peripheral arthritis, and IBD were associated with AxAxA disease duration. It is widely reported that up to 40% of AxAxA patients will eventually develop uveitis (Bacchiega et al., 2017, Zeboulon et al., 2008). However, our model (Fig. 2.2A) suggests that this figure may be considerably greater over a longer duration of AxAxA disease. Our findings are in line with a previous study of 301 SpA patients, which found that up to 50% of patients will develop uveitis after 50 years post-AS onset (Frantz et al., 2019). We did not find evidence of a previously reported plateau (Zeboulon et al., 2008) in first uveitis onset at 20 years post-AxAxA onset. Instead, we observed a relatively constant increase in uveitis history up to 50 years following AxAxA onset (Fig. 2.2A,F). Survivorship becomes a key limiting factor beyond 50 years post-AxAxA onset. Therefore, for clinical application, it is prudent to consider AAU as a continuing risk in AxAxA patients with no previous uveitis history.

While we lacked data to indicate the incidence of uveitis prior to AxSpA onset, our regression model suggests that it is likely around 20% within this Irish cohort. Previous studies have obtained highly variable results for this metric, but our findings are broadly in line with the expected prevalence of historical uveitis at AxSpA onset (Zeboulon et al., 2008, Sykes et al., 2018). To our knowledge, no previous reports have detailed the relationships between peripheral arthritis or IBD and AxSpA disease duration.

We observed that neither psoriasis nor dactylitis are linked to AxSpA disease duration (Fig. 2.2B-E). The relationship between psoriasis and AxSpA is complicated, with psoriatic arthritis falling into the SpA disease group and resulting in possible misdiagnosis as AxSpA (Baraliakos et al., 2015, Benavent et al., 2020). A prior study showed that the baseline prevalence of psoriasis in AxSpA patients was 16.8%, which increased to 26.8% after six years (Lucasson et al., 2022). In contrast, our results suggest that the prevalence of psoriasis is unlikely to exceed 20% at 10 years post-AxSpA onset with 95% confidence. Furthermore, the overall prevalence of psoriasis in ASRI was 16.8%; this with a median disease duration of 16 years within our cohort. This anomaly may be a feature of patient selection unique to each of the cohorts and, possibly, relating to the aforementioned issues of misdiagnosis raised between psoriatic and AxSpA conditions.

We did not observe an increase in the prevalence of dactylitis with disease duration, which contrasts with our observations in peripheral arthritis (Fig. 2.2E). We are not aware of any studies which have previously reported rate of onset of dactylitis in AS or AxSpA. It has frequently been reported that prevalence of dactylitis in these cohorts averages around 6% (PÉREZ ALAMINO et al., 2011, Paramarta et al., 2013), which is consistent with our own observation within the ASRI cohort (6.7%).

2.5.2 AxSpA disease in men and women

The ASRI cohort differs from many AxSpA cohorts in that women are younger than men within the group. In contrast, many AS cohorts tend to be biased toward younger male participants due to earlier disease onset in men. This anomaly may be explained by a key difference in patient selection for the ASRI AxSpA cohort, which enables the admission of patients who do not meet the mNY criteria. Thus, given that women are less likely to meet the modified New York criteria at an early age due to reduced axial disease progression, it is possible that the ASRI cohort is more selective toward younger women than studies which have focussed on classical presentations of AS. We also noted that, contrary to other studies, women were not older at age

of first symptom onset, suggesting that ASRI is capturing a more phenotypically diverse group of participants than many other registries.

We are the first group to identify an increased incidence of AAU in female patients and to link this to a higher burden of AAU preceding onset of classical AxSpA symptoms. Within our study, HLA-B27 had positive predictive value for uveitis risk in both sexes, although this was much stronger in women, where we observed a remarkable 95% prevalence of B27 in uveitis cases (Table 2.5). Women with dactylitis were also more likely to have concurrent uveitis (Table 2.5). AAU preceding onset of AxSpA has previously been reported in other studies and has been linked to improved functioning and reduced radiographic progression in AxSpA (Gómez-García et al., 2021). Unfortunately, we were not able to investigate whether this was the case within our cohort as we lacked the date of AAU episodes for patients. We did not observe any evidence to suggest that a history of AAU is associated with milder disease presentation in either sex (Table 2.6).

IBD followed a similar pattern to AAU in our models, with onset being more likely to precede AxSpA in women than in men, within our cohort. Additionally, women were also more likely to develop IBD following AxSpA onset, when compared to men. The latter was consistent with previous studies (Stolwijk et al., 2015). Peripheral arthritis and psoriasis were also increased in women, although these did not reach statistical significance.

Our other findings based on gender generally corroborated other studies. Axial symptoms (measured by BASMI and mNY criteria) were more pronounced in men (Table 2.2), consistent with the findings of previous work (Landi et al., 2016, Maguire et al., 2022d). Our findings within patient reported outcomes (BASDAI, ASQoL, HAQ-s) of AxSpA in relation to gender had previously been reported from within the ASRI cohort (Maguire et al., 2022d). Peripheral arthritis and dactylitis are known to be more common in women (Landi et al., 2016). We did not, however find any relationship between psoriasis and gender, in contrast to some previous work (Rusman et al., 2018).

2.5.3 The importance of EAMs in patient outcomes

In our examination of EAMs, we found that they coincided frequently, particularly between peripheral arthritis, psoriasis, and dactylitis (Tables 2.6-2.8). Peripheral arthritis emerged as a key prognostic indicator of disease outcomes, as patients presenting with concurrent peripheral arthritis scored worse in terms of disease activity, functioning, and quality of life (Table 2.8). Peripheral arthritis was also correlated with an increase in depression. Additionally, patients with

peripheral arthritis were more likely to receive methotrexate therapy and more likely to require multiple therapies to control their disease.

Uveitis showed weak associations with both psoriasis and peripheral arthritis. Patients with uveitis tended to develop AxSpA at an earlier age, but there was no evidence of worse disease outcomes in those with a history of uveitis. Although psoriasis was not directly linked to worse disease outcomes, it was strongly associated with peripheral arthritis and dactylitis. Patients with concurrent psoriasis were more likely to receive anti-TNF- α therapy and less likely to be treatment-free.

Interestingly, dactylitis defined a distinct subset of patients apart from other EAMs. It was strongly linked with non-radiographic AxSpA and could potentially act as a predictive biomarker for this subtype of AxSpA. Patients with dactylitis exhibited a marked increase in AxSpA disease activity and markedly reduced quality of life. Moreover, dactylitis was strongly associated with other EAMs, including uveitis, peripheral arthritis, psoriasis, and IBD.

Our data indicates that EAMS, particularly peripheral arthritis are further linked to worse disease outcomes in patients, and it would be beneficial to conduct risk assessments to identify those most at risk of late peripheral arthritis onset. Our models suggest that dactylitis is likely to either precede or occur early in AxSpA disease (Fig. 2.2E) and is strongly linked with the onset of other EAMs (Table 2.9). It may, therefore, be a useful predictive biomarker for the onset of further EAMs in patients.

Ideally, we would have aimed to model the incidence of other EAMs based on the presence of dactylitis, but our subset of dactylitis patients (60) was too small to construct any reliable model. Due to this small study population, we were also unable to investigate the profiles of radiographic versus non-radiographic AxSpA patients with dactylitis.

2.5.4 The effects of past smoking on obesity and comorbidity in AxSpA patients

Our findings indicate that smoking cessation is linked to improved scores in both disease activity (BASDAI) and quality of life (ASQoL) (Table 2.4). Furthermore, spinal disability progression (BASMI) is also reduced after smoking cessation, after accounting for age. However, day-to-day functioning, measured by BASFI, is only partially resolved.

Importantly, the cessation of smoking is linked to increased weight gain in patients, with the BMI of past smokers matching that of never smokers after accounting for age. Tentatively, we also found that these patients may be at an increased risk of diabetes mellitus compared to those who

have never smoked. Weight gain following smoking cessation in the general population is a well-documented phenomenon (Dare et al., 2015). However, amongst AxSpA patients, this is coupled with residual reduced functioning from past smoking habits, as measured by BASFI, and increased frequency of peripheral arthritis, which may result in further disability.

Previous studies have shown that BMI is a significant risk factor in AxSpA patients (Maguire et al., 2022b, Maas et al., 2016). Our findings indicate that smoking cessation yields more favorable outcomes for patients. However, AxSpA patients planning to quit smoking may benefit from additional support to manage potential weight gain following cessation.

As demonstrated in our study, smoking and obesity exhibit substantial impacts on the outcomes of AxSpA disease, which is an important consideration when examining genetic effects on disease outcomes. Nevertheless, it is evident that the interplay between smoking and obesity is intricate. The precise quantification of smoking habits poses a significant challenge, primarily due to the variable nature of cigarette consumption over time. This variability introduces substantial inaccuracies when assessing the total consumption of cigarettes (often reported in *pack-years*) in cross-sectional studies such as ASRI. Furthermore, an important aspect that remains unexplored is the lingering effects of smoking on AxSpA outcomes post-cessation, and a chronological assessment of these effects is currently lacking in scientific literature. Both smoking and obesity are also known to be linked with socioeconomic status, which is another important environmental influence (Min et al., 2019) which was not available within our data. These significant knowledge gaps preclude our ability to effectively account for these factors as covariates within our study of relationships between genotype and phenotype.

Chapter 3

The relationship between genetics and phenotypic characteristics in the ASRI AxSpA cohort

3 The relationship between genetics and phenotypic characteristics in the ASRI AxSpA cohort.

3.1 Introduction

Although the role of genetics in aetiology of AxSpA has been well characterised through GWAS studies performed primarily on AS cohorts (Cortes et al., 2013b, Reveille et al., 2010), exploration into the role of genetics in clinical presentation is limited. AxSpA presents with a broad range of disease severity and clinical manifestations beyond the spine. Extra-articular manifestations (EAMs), such as uveitis, psoriasis, and IBD are frequently reported in AxSpA (Cortes et al., 2013b) and offer strong candidates for genetic study due to their distinct and well-defined presentation in patients. However, the definition of some other characteristics can be complicated. Many features of disease severity and even some EAMs are not clearly dichotomous or may be entangled with other features or present after protracted periods following AxSpA diagnosis. For instance, PsA has become subject to debate on whether it is actually an alternative presentation of AS (Feld et al., 2018) and whether these two diseases belong to a spectrum of disease or constitute clear pathologic entities is yet to be established. This is particularly problematic when examining these traits in the context of a cross-sectional study. For example, we have previously demonstrated that first onset of uveitis may occur at any stage in the course of AxSpA disease (Chapter 2, Section 2.4.2, Fig. 2.2A), and others have reported similar findings (Zeboulon et al., 2008). This is complicated even further by the emergence of novel DMARD therapies which may further delay or prevent altogether the onset of some clinical features and manifestations (Choi et al., 2020a, van Bentum et al., 2019). Nevertheless, since no genome-wide association analysis has yet been performed to examine the potential role of high-impact common variants in specific clinical characteristics, it is sensible to start with GWAS analysis to rule out this possibility.

The ASRI AxSpA cohort benefits from extensive phenotyping, although this study size is small in the context of GWAS. One approach to compensate for reduced statistical power is to leverage variants already known to be associated to the disease through the work of larger GWAS. The most significant genetic biomarker for AxSpA is HLA-B27, which for a long time was thought to be the sole genetic determinant of disease. Numerous studies have investigated the effects of HLA-B27 on the presentation and severity of AxSpA/AS, shedding light on how a single genetic trait can influence disease course and presentation, even in complex disease. While HLA-B27 is linked to earlier disease onset and increased likelihood of AAU in AxSpA patients, its association with radiographic damage and inflammatory bowel disease in AS patients remains controversial (Feldtkeller et al., 2003, Arévalo et al., 2018, Huang et al., 2020). HLA-B27 is inversely

associated with psoriasis in AS patients (Arévalo et al., 2018, Chung et al., 2011). This inverse association may be the result of displacement of HLA-Cw6 by HLA-B27 in AxSpA patients since HLA-B27 and the psoriasis associated variant HLA-Cw6 do not share a haplotype. In contrast, studies of the effects of HLA-B27 on clinical instruments for disease activity (BASDAI), physical impairment (BASMI, BASFI), and quality of life (HAQ-s, ASQoL) have yielded inconclusive results (Arévalo et al., 2018), possibly due to variable cohort sizes and methodologies. This emphasises the need for replication studies to strengthen evidence for HLA-B27 in relation to disease activity and severity.

The *ERAP1* gene is closely linked to HLA-B27 in both function and disease association (Cortes et al., 2013b, Huang et al., 2020). Through its role in antigen trimming, the gene product, ERAP-1, plays a central role in the cell surface expression of HLA-B27 and its presentation of antigenic peptides (Cresswell et al., 1999). *ERAP1* is one of the strongest genetic determinants of AS amongst non-HLA loci and is associated only in HLA-B27 positive individuals (Cortes et al., 2013b, Akkoç et al., 2017). Due to its close relationship with HLA-B27, genetic variation within *ERAP1* could also be associated with disease presentation and severity. The most widely recognised genetic determinant in *ERAP1* is the coding sequence variant, rs30187(Lys>Arg), which bioinformatics work has suggested may impact on enzymatic activity of ERAP-1 (Wang et al., 2017, Cortes et al., 2013b). This is supported by a study demonstrating over-expression of disease relevant cytokines, including TNF- α , IFN- γ and Th17 associated cytokines, IL-17A, IL-17F, and IL-23 in B27 positive patients carrying the risk rs30187-T allele (Babaie et al., 2020). Despite its possible role in enzymatic activity, there is some evidence to suggest that rs30187-T itself may not be a disease-causing variant. It has been previously reported that while tag SNPs often vary between populations, disease-causing variants tend to remain consistent, (Amariuta et al., 2020), yet rs30187 is not associated with AS in Asian patients, despite reports of other *ERAP1* associations (Gao et al., 2020). Meanwhile, high linkage disequilibrium (LD) complicates the identification of further genetic risk variants near to *ERAP1* (Nossent et al., 2016, Costantino et al., 2015). However, it may be possible to tease apart haplotypes to determine if rs30187 is likely to be a disease-causing SNP and to further establish whether other SNPs in the *ERAP1* gene region are independently associated with disease.

Given the polygenic nature of AxSpA, it is likely that different disease associated genetic variants across the genome may be involved in the disease's various clinical characteristics and manifestations. This can be problematic to investigate in small studies due to the statistical implications of multiple testing. To overcome this, polygenic risk scores (PRS) offer a useful means to determine overall genetic risk. As explained in Chapter 1, Section 1.8.3, PRS are

personalised aggregated scores generated from an individual's genetic risk variants, weighted by their relative effect sizes, as determined by a reference study. Although PRS are unable to determine specific risk of individual genes, they offer an indication of whether previously identified disease-associated variants are collectively likely to be influencing the phenotype being analysed. Furthermore, PRS can be tailored to include only specific regions of the genome or variants of particular interest, such as those of high odds ratio or statistical significance. However, it is important to note that PRS are based on the phenotype examined in the reference population and may lose power when alternative traits are examined, even within the context of the same disease (Choi et al., 2020b). Despite this, it is worthwhile to examine the relationship between polygenic risk of AS and the development of additional manifestations and comorbidity.

3.2 Aims

- To perform GWAS analysis to identify genetic variants of high impact on clinical phenotype.
- Investigate whether the *ERAP1* SNP, rs30187, which is closely associated to HLA-B27, may also play a role in disease severity and extra-axial manifestations.
- To examine the aminopeptidase gene region (*ERAP1–ERAP2–LNPEP*) of chromosome 5 for genetic association to AxSpA through haplotype analysis. We further aim to identify whether rs30187 is likely to be the sole disease associated variant within this region.
- To examine the relationship between polygenic risk and clinical phenotype.

3.3 Methods

3.3.1 Data collection

Data of study participants was collected by a trained professional in a structured interview format. Each participant was assigned a study identification number to which all corresponding data is then assigned. Data collection was performed in accordance with the ASRI protocol (www.asri.ie). Clinically relevant information was confirmed from medical records, with exception of date of first symptom onset, HLA-B27 status and current depression (self-reported). Collected data included, but was not limited to the following:

Demographics: age, sex, ethnicity, family history of AxSpA.

Clinical characteristics: age of disease onset, age of diagnosis, extra-articular manifestations, comorbidities, HLA-B27 status, modified New York (mNY) Criteria, ASAS criteria for axial spondyloarthritis.

Clinical measurements: erythrocyte sedimentation rate (ESR), C-reactive protein (CRP), Bath Ankylosing Spondylitis Disease Activity Index (BASDAI), Bath Ankylosing Spondylitis Metrology Index (BASMI), Bath Ankylosing Spondylitis Functional Index (BASFI), Ankylosing Spondylitis Quality of Life (ASQoL), Health Assessment Questionnaire modified for the spondyloarthropathies (HAQ-s).

Socio-economic characteristics: current employment status, smoking, alcohol intake.

3.3.2 Genotyping

Blood samples were collected from participants using K₂EDTA tubes. DNA was extracted manually by the TTMI Biobank facility, and the extracted DNA was eluted into Tris EDTA (TE) buffer (pH 8.0). DNA concentrations of eluates were approximated using a Thermofisher Nanodrop ND-8000 spectrophotometer and samples were then stored at -20°C or -80°C.

Illumina Infinium ImmunoChip V2.0 Chips (Part # 15072709) were prepared in accordance with manufacturer's instructions (Infinium® HTS Assay Protocol Guide, Part # 15045738 Rev. A). The completed chips were sent to *Cardiff University MRC Centre for Neuropsychiatric Genetics and Genomics* for interpretation by Illumina HiScan. Raw data outputs were transferred back to TCD via end-to-end encrypted peer-to-peer file transfer.

The raw data outputs were processed using Illumina's GenomeStudio (V2011.1) software. Raw data were imported in GenomeStudio using cluster methodology with GenCall threshold set to

0.15. DNA reports for genotyping rates and heterozygosity were generated. Samples with <98% SNP call rate or SNP heterozygosity beyond 4 standard deviations from the cohort mean were excluded from further analysis. Samples excluded due to poor call rates were repeated on later Immunochips, where possible. Due to the low number of sex chromosome based markers on the Immunochip V2.0 manifest, sex interpretation by genotype was unreliable. Sex check was performed in both GenomeStudio and Plink V1.9 (Chang et al., 2015), and samples were excluded where divergent results were obtained by both methodologies compared to the ASRI phenotypic database.

Further quality control (QC) was performed in the Plink V1.9 environment. SNPs with call rates <98% and variants with minor allele frequency (MAF) less than 5% were excluded. Variants in the sex chromosomes and variants with unassigned chromosomal location were also removed from the data set. No suitable threshold for Hardy-Weinberg equilibrium (HWE) exclusion criteria was identified, most likely due to disease associated skewness in this patient-only cohort. This is discussed further under heading 3.3.3. Cluster plots of all SNPs used in subsequent determinations were reviewed manually, and examples of satisfactory and unsatisfactory clustering are shown in Fig 2.1.

Cryptic relatedness was assessed using identity by descent analysis in Plink V1.9. Briefly, variants in strong linkage disequilibrium were excluded from the calculation using Plink's *indep-pairwise* function, scanning variants within a 50 kilobase (KB) window, sliding at 5KB increments. A list of excluded variants with LD (r^2) greater than a threshold of 0.2 was generated. The remaining variants were processed using Plink's *genome* function, generating a data dump of cryptic relatedness between every possible pair of individuals within the dataset. Second degree and closer relatives were pruned based on kinship coefficient (PI_HAT) >0.18, with one individual of each pair being excluded at random.

Finally, patients of non-Caucasian ancestry were excluded, based on records from the ASRI database.

An overview of samples filtered by quality control is presented in Table 3.1

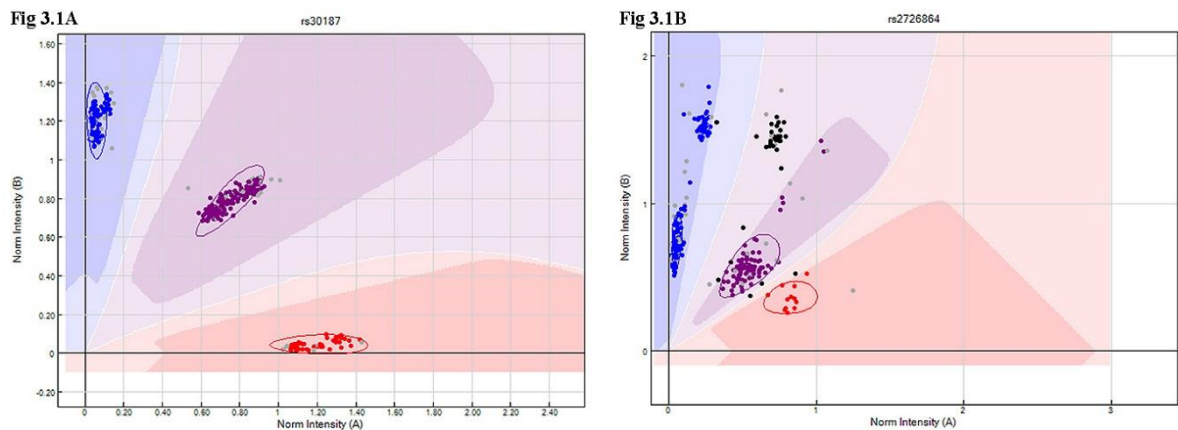


Fig. 3.1 Cluster plots of rs30187 and rs2726864 generated by Illumina GenomeStudio V2011.1. **Fig. 3.1A** Clear clustering of rs30187 genotypes by cartesian coordinates. Homozygous participants cluster tightly within the blue and red regions, heterozygotes align in the mauve region. **Fig. 3.1B** An example of poor cluster quality in variant rs2726864. Multiple clusters are observed in the blue homozygous region and an unidentified cluster falls between the mauve and blue regions. These features may be the result of batch effects and render this variant uninterpretable.

Table 3.1: Sample QC summary

Total samples genotyped	707
Divergent ethnicity	27
Cryptically related	20
Misdiagnosis* or discontinuation of ASRI participation	15
Low heterozygosity	3
Discordant sex	1
Available samples post-QC	641

*Patients re-diagnosed as non-AxSpA subsequent to admission to the ASRI study.

3.3.3 Hardy-Weinberg disequilibrium

During the QC process, we attempted to prune variants based on a Hardy-Weinberg Equilibrium (HWE) departure threshold (P value). However, within a cohort comprised entirely of AxSpA patients, we expected to observe Hardy-Weinberg divergence. Confirming this, we found significant numbers of variants with well-defined genotype clustering were HWE divergent. This was particularly prevalent in the HLA region of chromosome 6, where 29.9% (164 out of 488) of variants were outside of a nominal HWE threshold of $P < 1 \times 10^{-10}$. Notably, this included the

HLA-B27 tag variant, rs4349859 ($P=6.18 \times 10^{-67}$), and the novel tag variant, rs7771432 ($P=9.89 \times 10^{-27}$), which was later used as part of the HLA-B27 tagging strategy. Cluster plots of genotyping fluorescence data for these two variants are shown in Fig. 3.2. We also observed divergence from HWE outside of the HLA region, such as rs7136752 ($P=2.82 \times 10^{-15}$) in chromosome 12 (Fig. 3.3). From our data we determined that no suitable threshold for HWE could be employed as a QC filter. We therefore opted not to include HWE filtering as part of our QC process. Details of the HWE departure counts for each chromosome are presented in Table 3.2.

Table 3.2 Counts of HWE divergence by chromosome

Chromosome	Count of HWE divergent variants
1	41
2	33
3	25
4	13
5	29
6	164
7	18
8	15
9	22
10	27
11	6
12	21
13	5
14	9
15	6
16	11
17	10
18	5
19	14
20	8
21	3
22	3
Total	488

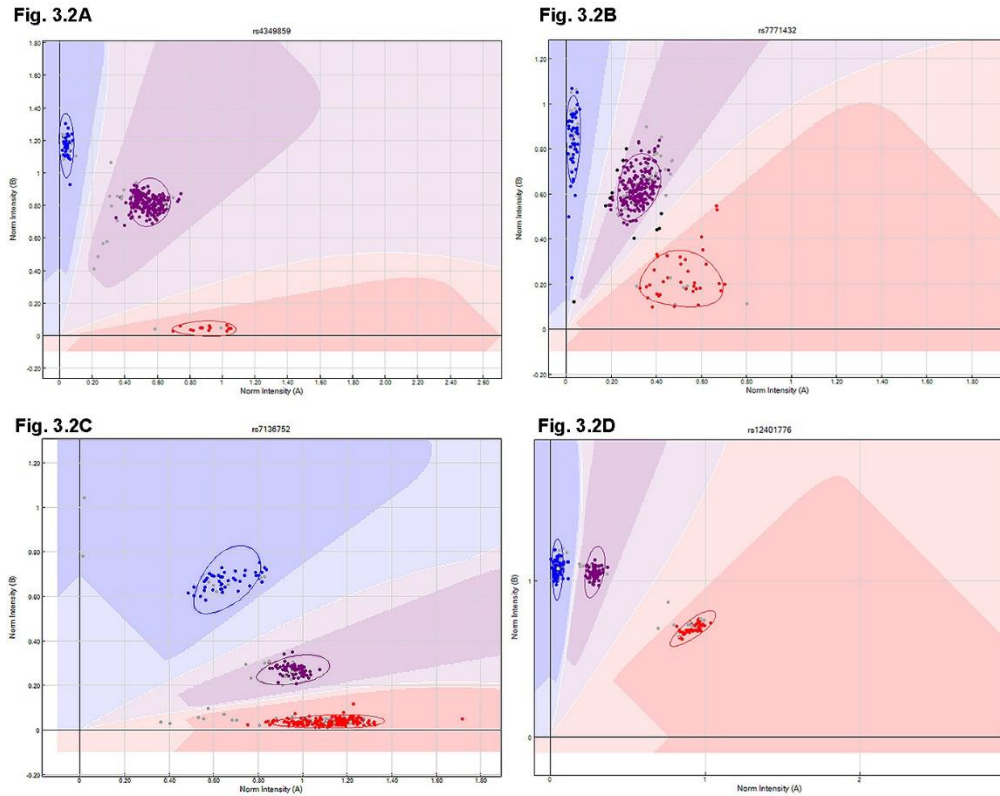


Fig. 3.2: Cluster plots of variants with acceptable genotype clustering failing HWE generated by Illumina GenomeStudio V2011.1.

Fig. 3.2A: HLA-B27 tag variant, rs4349859 ($P=6.18 \times 10^{-67}$).

Fig. 3.2B: Novel tag variant rs7771432 ($P=9.89 \times 10^{-27}$).

Fig. 3.2C: rs7136752 on chromosome 12 ($P=2.82 \times 10^{-15}$).

Fig. 3.2D: rs12401776 on chromosome 1 ($P=2.07 \times 10^{-11}$).

3.3.4 Calculation of threshold for GWAS significance

Following quality control, a total of 179,442 variants were available for analysis. To account for multiple testing in GWAS, we performed Bonferroni correction for a false discovery rate (α) target of 5%, based on the total number of independent variants after accounting for LD. Using the Genetic Type 1 Error Calculator (GEC) (Li et al., 2012), we estimated the number of independent variants to be 86,680 and thus the threshold for statistical significance was determined to be 5.77×10^{-7} . A second threshold line at $P=1.5 \times 10^{-5}$ was also applied to visualise variants of potential interest on Manhattan plots. We also performed statistical power analysis using CaTS Power Calculator (Skol et al., 2006) for a study size of 641 patients, assuming an equal split between comparison groups with allele frequency of 0.3 in the control population. Under these conditions, we expect an 80% likelihood to detect a SNP association with an odds ratio (OR) of 2.32, based on an additive genetic association model. This limited us to the detection of SNPs of high effect size within the GWAS analysis of phenotypes. A summary of

odds ratios detectable at 80% statistical power for various allele frequencies is presented in Table 3.3, below.

Table 3.3: Statistical power to detect associations by allele frequency

Allele freq.	OR	Detection
0.1	2.74	80%
0.2	2.37	80%
0.3	2.32	80%
0.4	2.42	80%
0.5	2.71	80%

3.3.5 Imputation of variants from whole genome sequence data

Plink binary bi-allelic genotype files were converted to variant call file (vcf) format and an index of genome positions was generated using Tabix r942 (Li, 2011). Data was phased with SHAPEIT4 (Delaneau et al., 2019) using Broad Institute's map b37, based on the GRCh37 hg19 human genome reference. Outputs were again indexed with Tabix. Reference data for imputation was collected from the GBR and CEU subsets of 1000 Genomes Phase 3 v5 (Clarke et al., 2012) cohort. These subsets containing totals of 99 and 91 individuals, respectively were extracted from the complete 1000 Genomes data using vcftools (Danecek et al., 2011) and indexed with Tabix. Imputation was performed using Impute5 (Rubinacci et al., 2020) using a positional Burrow-Wheeler transformation (PBWT) depth (number of conditioning neighbours) of 8 (default = 4) and frequency of 0.005 centimorgans (default = 0.02). This increase of PBWT depth and frequency was performed to improve accuracy at the expense of increased computational requirements. The accuracy of imputed SNPs were filtered for info score >0.7 (Verma et al., 2014).

Post-imputation, non-SNP variants and SNPs with MAF < 0.05 were pruned. Multi-allelic variants introduced during imputation were pruned by removal of least frequent allele. Imputed variants which shared the same position and alleles but differed by name were pruned using BASH tools.

A cross check of imputation using the previously reported HLA-B27 tag variant, rs116488202 (Cortes et al., 2013b) which was imputed in our data from the process above, was performed using a subset of 172 patients analysed for this SNP by Taqman assay (method described in section 3.3.8). Concordance in genotyping results for this assay was 98.8%, with two samples providing discordant results. In both of these cases, rs116488202 was determined to be heterozygous by imputation results and homozygous by TaqMan assay. Rs116488202 results

from imputation matched rs4349859 on the Immunochip in both of these cases, suggesting this was due to a genotyping error of either rs4349859 on the Immunochip or TaqMan assay of rs116488202. Notably, TaqMan results for rs13202464 matched the Immunochip rs4349859 and imputed rs116488202 in both cases, suggesting the genotyping error was most likely in the rs116488202 TaqMan assay results.

3.3.6 Conversion of IGAS variant ID's to match ASRI

To facilitate meta-analysis and other investigations, variant ID's listed in the IGAS (Cortes et al., 2013b) summary data (performed on first generation Immunochip) were converted to match the ASRI data reference from chromosome and position data using UNIX tools. Non-matching and ambiguous SNPs were excluded. Of 123,941 variants listed in the IGAS summary data, 77,821 were matched with ASRI variants.

3.3.7 Statistical evaluation of genotyping data

Phenotype files in Plink compatible whitespace delimited format were created based on ASRI data using Microsoft Excel. The phenotypes were presented in case/control format for Plink compatibility. Missing data was excluded from analysis. Plink V1.9 (Chang et al., 2015) was used to perform association analysis by χ^2 and Cochran-Armitage trend test. Manhattan and QQ plots of GWAS summary data were generated using QQman (Turner, 2014) in R V4.1.3 (Team, 2013). Data for individual SNPs of interest were extracted from the data set using Plink V1.9. Regression analyses, including univariate and multivariate were performed in Minitab V21.1 (64 bit).

LD was determined within ASRI data where available using Haploview 4.1 (Barrett et al., 2004). Alternatively, we used the LDPair tool, part of the LDLink suite (Machiela and Chanock, 2015).

Fine mapping using haplotype association analysis, including LD block analysis was performed in the Haploview 4.1 (Barrett et al., 2004) environment.

Line graphs were created in Minitab V21.1 (64 bit). Beanplots were produced in R V4.1.3 using Beanplot package (Kampstra, 2008).

For the comparison of phenotypes based on HLA-B27 and ERAP1 genotypes, Mann Whitney U test, χ^2 and multivariate regression were performed using Minitab V21.1.

Statistical significance was defined at a threshold value of $P < 0.05$, unless otherwise indicated in specific methodologies.

3.3.8 HLA-B27 typing and ImmunoChip genotyping validation

The HLA-B27 status of patients were available for 480 of the 641 individuals remaining post-QC. Verification of HLA-B27 status and typing of additional samples was performed using a combination of tag SNPs identified from the current study data and previously reported tag variants.

Briefly, genome wide association was performed between HLA-B27+ and HLA-B27- patients, using ASRI database reference data as a phenotypic distinction. The strongest HLA-B27 association identified through GWAS (442 B27+, 38 B27-) was observed in tag variant rs7771432 (Fig. 3.3, below). This variant was found to be in linkage disequilibrium ($r^2 = 0.68$, $D' = 1.0$) with a previously reported tag variant for HLA-B27 in the population of Great Britain, rs4349859 (Machiela and Chanock, 2015, Akar et al., 2017) (Fig. 3.4). Other previously reported tag variants, rs116488202 and rs13202464 (Akar et al., 2017, Roberts et al., 2013), were also strongly represented by GWAS (Cortes et al., 2013b, Evans et al., 2011).

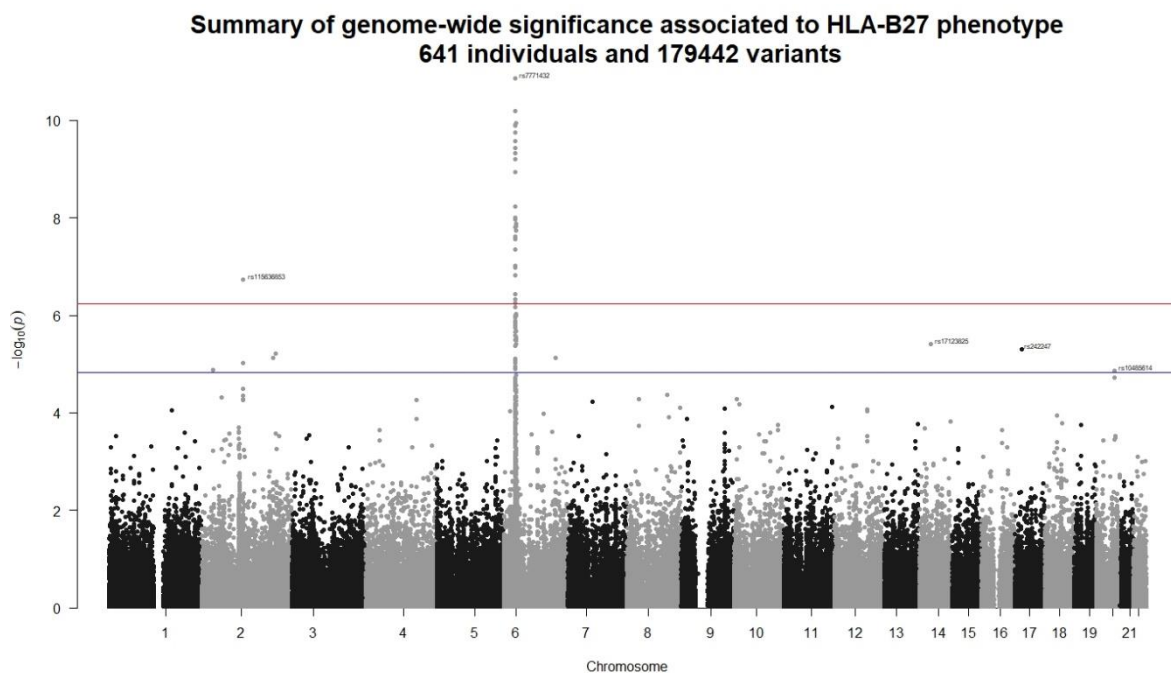


Fig. 3.3: Manhattan plot of GWAS of 480 individuals (442 B27+, 38 B27-). Strong representation in the HLA-region of chromosome 6 is led by rs7771432. Other variants with high association include rs4349859, rs116488202 and rs13202464 (annotation not shown).

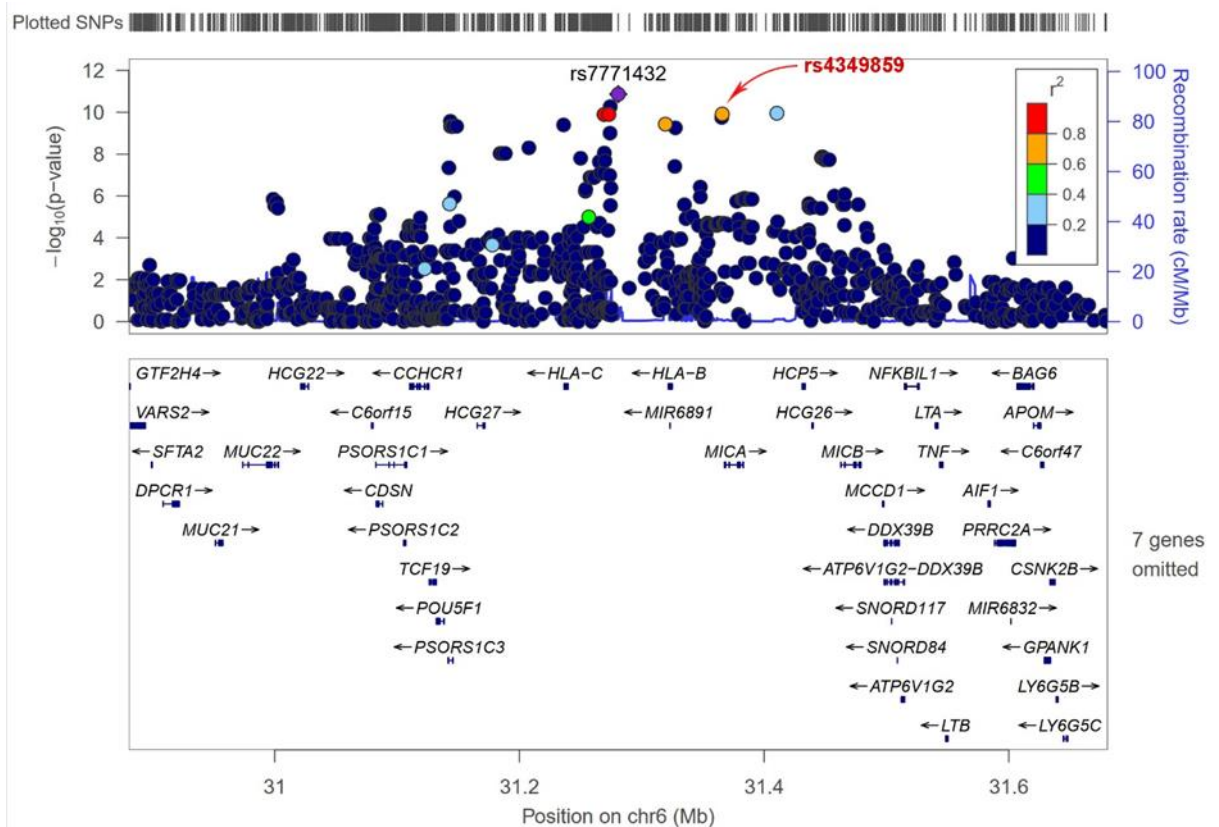


Fig. 3.4: Linkage disequilibrium within 50Kb radius of rs7771432. Annotated are rs771432 and rs4349859. Calculated from ASRI GWAS data using LocusZoom (Pruim et al., 2010).

Due to the critical importance of HLA-B27 in AxSpA disease causation, we used the opportunity of HLA-B27 typing to perform validation of Immunochip and imputed SNP outputs for relevant markers. We selected markers rs13202464 (genotyped on chip) and rs116488202 (imputed as in Section 3.3.5). Rs116488202 is in complete LD with HLA-B27 tag SNP, rs4349859 in European populations and thus offers an alternative typing marker. Rs13202464, meanwhile lies in incomplete linkage disequilibrium with both of these markers.

Validation of Immunochip genotyping results for rs13202464 (on chip) and rs116488202 (imputed) was performed using Taqman assays. This process was intended to confirm Immunochip and imputation results, particularly in cases where genotyping results were discordant with HLA-B27 typing data available from the ASRI database, and to out-rule strand flipping or burrow-wheeler transformation errors during the imputation process. Briefly, Taqman assays for rs13202464 (ThermoFisher cat# C__32201158_10) and rs116488202 (ThermoFisher cat# C_100629578_10) were performed in singlicate on 172 patients. Taqman assays were performed in accordance with manufacturer's instructions using Taqman Genotyping Mastermix (ThermoFisher cat# 4371353) and amplification was achieved using an ABI 7500 qPCR thermocycler.

Patients with missing or genotyping discordant results were preferentially selected for Taqman genotyping. In total, 136 patients with no available HLA-B27 status record, 28 with discordant results between genotype and ASRI record, 8 patients with concordant results between genotyping and ASRI record and 1 patient with isolated rs13202464 positivity were selected. Strong agreement was obtained between the Taqman assay and both the chip genotyped SNP rs13202464 (100%) and the imputed variant, rs116488202 (98.8%). Taqman assay confirmed results for the 8 patients with concordant results between Immunochip and ASRI database. The chip genotyping result for the patient carrying rs13202464 in the absence of rs4349859 or rs16488202 was also confirmed by Taqman results. 27 of 28 genotyping results which were discordant between Immunochip and HLA-B27 typing results available from the ASRI database remained discordant after Taqman genotyping. The remaining patient, carrying the minor allele for rs13202464 in isolation was documented as HLA-B27 positive on the ASRI database. This was in agreement with our results for rs13202464 but contrary to other key HLA-B27 tag SNPs, including rs4349859 and rs16488202. It is likely that this patient was negative for HLA-B27 on the basis that rs13202464 is known to have reduced specificity for HLA-B27 when used in isolation (Martínez-Nava et al., 2021).

Overall, chip-based markers averaged concordance of ~90% with the ASRI database in HLA-B27 status calls. Due to the high predictive power of some of these markers (>95%), particularly rs4349859 and rs16488202 (Martínez-Nava et al., 2021), it is likely that this unexpectedly low concordance was due to self-reporting by some patients included in the study, coupled with data entry errors on the ASRI database. A full cross comparison of results for all SNPs, methodologies and ASRI database entries are summarised in Table 3.4.

Despite high predictive power of rs4349859 and its direct proxy, rs16488202, these two variants alone are not 100% sensitive for HLA-B27 (Martínez-Nava et al., 2021). To maximise our sensitivity and specificity, and to ensure consistency of calls for HLA-B27 status, a standardised 4-SNP system was used to HLA-B27 type all patients irrespective of their previous status according to the ASRI database. All patients were assigned as HLA-B27 positive where at least two of the four HLA-B27 associated alleles of rs4349859, rs13202464, rs116488202 and rs7771432 were present. All other patients, conversely, were tagged as HLA-B27 negative.

Table 3.4: Cross-reference table of the concordance (%) of HLA-B27 tagging variants and methodologies.

Marker/methodology	rs7771432 Chip (n = 629)	rs4349859 Chip (n = 641)	rs13202464 Chip (n = 641)	rs116488202 Imputed (n = 641)	rs13202464 TaqMan (n = 172)	rs116488202 TaqMan (n = 172)	ASRI Database (n = 480)
rs7771432 Chip (n = 629)		85.3	86.3	85.8	90.1	89.5	86.7
rs4349859 Chip (n = 641)	85.3		97.2	99.5	97.1	98.8	92.1
rs13202464 Chip (n = 641)	86.3	97.2		97.7	100.0	95.9	92.1
rs116488202 Imputed (n = 641)	85.8	99.5	97.7		97.1	98.8	92.3
rs13202464 TaqMan (n = 172)	90.1	97.1	100.0	97.1		95.9	25.0*
rs116488202 TaqMan (n = 172)	89.5	98.8	95.9	98.8	95.9		22.2*
ASRI Database (n = 480)	86.7	92.1	92.1	92.3	25.0*	22.2*	

Each block in the table represents the percentage concordance between the methodologies indicated at the top row and leftmost column.

*A total of 36 patients were available for this analysis with intentional selection bias toward discordant results.

3.3.9 HLA-typing of patients from genotypic data

Expanded HLA-typing was performed using the HIBAG (Zheng et al., 2014) R package which imputes HLA-types using extended haplotype structures, based on a trained data set. Briefly, a subset of SNPs were extracted from the imputed ASRI genotyping data set containing all SNP variants within the HLA-region (Chr6:28,000,000 and Chr6:34,000,000, hg19), with this allowing sufficient excess to include nearby markers which may be in LD. HIBAG was trained using a pre-defined European model, based on SNP data corresponding to Illumina Omnichip (hg19) with 4 digit resolution. Prediction was performed against the pre-defined model.

HLA-B27 calls were cross-validated against the ASRI database and the calling methodology described in Section 3.3.8. HIBAG HLA-B27 calls were 100% concordant with our 4-SNP B27 tagging methodology and 92.1% concordant with the ASRI databank, supporting our previous findings in section 3.3.8. 97.1% of HLA-B27 calls were determined to be B27 subtype 27:05, with 27:02 comprising the remaining 2.9%. These data are summarised in Table 3.5.

Table 3.5: HIBAG Report Summary

Total Patients (n)	646
HIBAG B27:05 count (n)	533
HIBAG B27:02 count (n)	16
Non-27:02, non-27:05 HLA-B27 count (n)	0
HLA-B27 positivity (%)	85.0
Concordance with ASRI Databank (%)	92.1
Concordance with 4-marker call methodology (%)	100

HIBAG was also used to predict HLA types in 1,911 Irish controls, which had been genome sequenced and were available for case-control comparisons in this study. Briefly, sequenced control sample variant call format (vcf) data aligned to the GRCH38/hg38 genome reference was converted to GRCh37/hg19 format in order to match ASRI genotyping data by Dr. Deepak Bharti using Lifter (UCSC tools). SNP data was then extracted from the extended HLA region of Chromosome 6 (Chr6:27,000,000 – 34,000,000) by Dr. Deepak Bharti. Data was merged with imputed ASRI data in Plink. A minimum genotyping rate of 98% was accepted for SNPs in the merged data outputs. This process excluded all SNPs which were not included in either data set. Data was then processed for HLA type prediction using HIBAG, as detailed above. To ensure that HIBAG's calls for HLA-Cw6 were consistent in ASRI patients before and after merging with the 1,911 sequenced controls, we compared HIBAG results for HLA-Cw6 between these two data sets. Concordance of HLA-Cw6 calls was 100%.

3.3.10 Interrogation of *ERAP1* haplotypes

3.3.10a Identification of regions of interest near the *ERAP1* gene

Haplotypes can be useful in providing information on genetic associations because an associated SNP can be present on more than one haplotype in a given genomic region. If all haplotypes containing a disease associated allele of a SNP are not associated with disease, this may argue that the SNP is not itself causative, and other interesting information may be gleaned from the pattern of association of the haplotypes and the SNPs from which they are composed.

To create an LD map for the *ERAP1* region, SNPs within a 500KB segment between Chr5:96,000,000-96,500,000 (GRCh37, hg19) were used. The imputed ASRI data from 641 patients was merged with sequencing data from 1911 healthy Irish controls, and output data was thinned for genotyping rate of 98% to remove orphaned SNPs from the merged data. A threshold for HWE $>1 \times 10^{-10}$ was suitable for this combined data, for which controls constituted approximately 75%. The Plink software's --flipscan function was used to identify variants of incongruous allele frequency between the sequenced and genotyped data sets, based on estimated LD. Incongruous variants were pruned by consensus of neighbouring SNPs. In cases where there was no consensus of the strand between a SNP pair, both SNPs were excluded. Out of a total of 1,385 SNPs, 17 were thus pruned. Subsequently, the remaining SNPs were thinned randomly using Plink's --thin-count function to select 250 representative markers. Variants of interest, namely rs30187, rs27434, rs17482078 and rs2287987 previously published in literature (Cortes et al., 2013b, Paladini et al., 2018, Robinson and Brown, 2014, Babaie et al., 2020) were merged back into this dataset. This resulted in a total representative variant pool containing 252 SNPs. Haploview 4.1 was used to process the data using the 95% confidence interval methodology developed by Gabriel et al (2002) for haplotype discrimination (Gabriel et al., 2002). Haplotypes with a frequency of at least 5% frequency were included in the analysis.

Haplotypes were prioritised by AxSpA SNP associations within their respective blocks. SNP associations were defined by residual statistical significance ($\alpha < 0.05$) after Bonferroni correction for multiple testing was accounted for. Briefly, association analysis of 252 variants was performed in Plink using the --assoc function. False discovery rate correction was performed by estimating the total number of independent SNPs from within the 252 SNP sample using clumping based on LD r^2 . In order to obtain a full scan of independent variants within the 500Kb *ERAP1* region, no discrimination for statistical significance for clumping was applied ($P \leq 1$) for this process. An estimated total of 23 independent variants were thus identified. After Bonferroni correction for false discovery, the threshold for statistical significance for α at 5% was calculated to be $P < 0.00217$, which was rounded to $P < 0.002$ for convenience. The most statistically significant and functionally relevant haplotype block regions were selected for further analysis.

These included haplotype blocks 6,7 and 16 (Supplementary Table S3.1). Haplotype blocks 6 and 7 were of particular interest as they contained previously published disease-associated SNPs, while haplotype block 16 was of potential interest because it was not closely linked with the *ERAP1* coding region. Haplotype block 8, while statistically significant, was not investigated further due to its high linkage with blocks 6 and 7.

3.3.10b Analysing statistical significance of haplotypes

We aimed to maximise our ability to capture AxSpA disease risk in each of the haplotype blocks prioritised by the screening methodology detailed above. Thus, we re-built the haplotypes for these blocks, specifically including variants potentially associated to AxSpA disease. We first performed association analysis for each of the regions enclosed by these prioritised haplotype blocks, including all available variants in our analysis. We then merged variants potentially associated with AxSpA at $P < 0.01$ with representative SNPs selected at random from the same region. Thus, each haplotype block contained significant and non-significant haplotypes in terms of AxSpA association. Data was processed in Haploview 4.1 (Barrett et al., 2004).

To examine whether rs30187 is the strongest genetic determinant in *ERAP1*, we followed a similar process, combining 46 representative SNPs from within its respective haplotype block (block 7), together with rs30187 and 3 other previously published SNPs, rs27434, rs17482078 and rs2287987. Data was processed in Haploview 4.1 (Barrett et al., 2004). Statistical analysis of haplotype frequencies based on rs30187 allelotype was performed in Minitab V21.1.

3.3.10c Generation of tag variants for disease association within the *ERAP1* coding region

To screen for independent genetic associations within the *ERAP1* coding region, we generated tag variants to capture genetic association across the *ERAP1* region. First we filtered for SNPs of high statistical significance for AxSpA association ($P < 4.35 \times 10^{-6}$, $\alpha < 0.0001$, based on the aforementioned 23 independent variants) within the region bounded by juxtaposed Haplotype blocks 6-7, which enclose the *ERAP1* coding region. Next, we filtered these using Plink's clump function, based on $r^2 < 0.8$ within a 250Kb window to obtain a pool of variants which were in incomplete LD. Six variants emerged from this analysis. To this selection we added previously the published disease associated variants, rs30187, rs27434, rs17482078 and rs2287987. This total of 10 tag variants were then analysed for AxSpA association and prediction of HLA-B27 positivity using Minitab V21.1.

To examine whether *ERAP2* may be associated with AxSpA separately from *ERAP1*, we generated extended haplotypes including previously published variants in the *ERAP2* region

rs2248374 and rs2549782 and the 10 *ERAP1* tag variants described above. We forced true representation of haplotypes in Haploview 4.1 by using single spine of LD with no threshold for D' .

3.3.11 Calculation of polygenic risk scores (PRS)

Summary statistics data from Cortes et al (Cortes et al., 2013b) were employed as base data for PRS calculation. Variant ID's were harmonised between imputed test and base data. Non-SNP variants and SNPs of low minor allele frequency were previously rejected as covered in *Section 3.3.5*. Variants which were missing or ambiguous between the base and test data sets were pruned. Duplicate variant ID's were pruned from the base data set with UNIX tools. A remaining total of 71,226 unambiguous SNP variants matched to both the imputed ASRI data and GWAS base data sets were considered for PRS calculation. SNPs were filtered for LD using the clumping feature in Plink V1.9. To select a representative population of SNPs for PRS, clumping was performed for all available variants based on the base GWAS summary dataset, with P-value threshold set to consider all variants ($P \leq 1$) and genetic correlation for filtering set to strict criteria with r^2 threshold of 0.1 within a 250Kb sliding window.

Shrinkage (informed selection of SNPs) was performed with the desired P-value thresholds for PRS generated for $P < (1 \times 10^{-8}, 1 \times 10^{-7}, 1 \times 10^{-6}, 1 \times 10^{-5}, 1 \times 10^{-4}, 0.001, 0.05, 0.1, \text{ and } 0.5)$. PRS generation was performed in Plink V.1.9 using the `--score` function, with scoring based on the β statistic for SNP associations in the base GWAS data.

Three sets of PRS were calculated:

- (i) For all available SNPs.
- (ii) HLA-region SNPs only (defined as Chr6: 27,000,000 – 34,000,000, hg19 reference).
- (iii) All SNPs excluding the HLA region.

To account for population stratification, the first 6 principal components were extracted from ASRI data to be used as covariates in regression analysis. Briefly, using Plink's `--indep-pairwise` function, SNPs were pruned based on LD of $r^2 > 0.25$ within a sliding window of 200Kb, with this processed at 50Kb increments. Principal components were then generated from the pruned data using Plink.

The relationship between polygenic risk and clinical phenotype was performed by multivariate regression analysis using R V4.1.3. Covariates for each phenotype included principal

components, sex and, where indicated, AxSpA disease duration or age at first AxSpA symptom onset.

Multivariate regression was performed using R V4.1.3 to examine the relationship between PRS scores and phenotype. Covariates included principal components, sex and, where applicable, AxSpA disease duration or age at first AxSpA symptom onset.

To control for false discoveries, P-values for all outputs were simultaneously processed in R V4.1.3 using Benjamini-Hochberg (Benjamini and Hochberg, 1995) correction for a target false discovery rate (FDR)=0.05.

3.4 Results: Genome wide risk in AxSpA phenotypes

3.4.1 Demographic characteristics of genotyped cohort

Of the 912 patients currently registered in the ASRI study, genotyping was performed on 707 individuals. Following quality control (Section 3.3.2), a total of 641 result profiles were suitable for analysis. A demographic summary of this genotyped sub-cohort of patients is presented in Table 3.6.

Table 3.6: Demographic characteristics of genotyped patients

Characteristic	Included n	Median (IQR) or n (%)
Male, n (%)	641	494 (77.1)
Age, median (IQR)	641	45 (37 - 55)
Age at first symptom onset, median (IQR)	641	24 (19 - 32)
Age at diagnosis, median (IQR)	641	34 (27 - 42)
Delay to diagnosis in years, median (IQR)	641	6 (2 - 12)
AxSpA disease duration in years, median (IQR)	641	18 (10 - 28)
mNY criteria* n (%)	641	496 (77.4)
Uveitis**, n (%)	630	225 (35.7)
Psoriasis, n (%)	631	107 (17.0)
Inflammatory bowel disease, n (%)	632	66 (10.4)
Family history, n (%)	641	180 (28.1)
HLA-B27 Positive (ASRI data), n (%)	480	442 (92.1)
HLA-B27 Positive (post-genotype), n (%)	641	543 (84.7)

* Refers to modified New York criteria, 1984 for AS

** At least one recorded episode of uveitis

3.4.2 Genome wide association analysis of phenotypic presentation

To screen for variants of high impact relating to specific disease features, we conducted GWAS by comparing AxSpA patients with and without each of the major EAMs and other well-defined phenotypes. A summary of the phenotypes examined is presented in Table 3.7. The analysis was performed in 641 AxSpA patients using a case/control format for each phenotype. *Cases* were defined as AxSpA patients presenting with the phenotype of interest, while *controls* were the remainder of the AxSpA cohort. Manhattan plots were generated to identify potential SNPs of

high impact. We also generated QQ plots to examine whether overall genetic risk could be attributed to the analysed phenotypes.

Table 3.7: A brief explanation of phenotypes used in GWAS analysis.

Phenotype	Criteria/details.
Male sex	Male sex defined post-genotyping QC amendments.
Early onset	Early onset of disease (<21 years).
mNY criteria	Patient diagnosed as AS in accordance with modified New York Criteria 1984.
Uveitis	At least one reported episode of AAU.
Psoriasis	Concurrent psoriasis.
IBD	Inflammatory bowel disease, including ulcerative colitis and Crohn's disease.
FHx	Family history of AxSpA.
B27+	HLA-B27 positive in ASRI database. This was used to identify and confirm tagging variants.

3.4.3 Acute anterior uveitis

AAU is the most widely reported EAM of auto-inflammatory pathophysiology reported in AxSpA. 35% of our AxSpA patient cohort included in the GWAS study have reported at least one episode of AAU. As we have previously reported in Chapter 2, Section 2.4.2, episodes of AAU can first occur at any point after AxSpA onset, and it was expected that over 50% of patients will have experienced at least one episode of AAU at 50 years post-AxSpA onset. This pattern was again clear in the genotyped subset of the ASRI cohort (Fig. 3.5). This will have reduced our ability to observe genetic differences based on this phenotype since we are unable to predict which patients without history will later develop AAU.

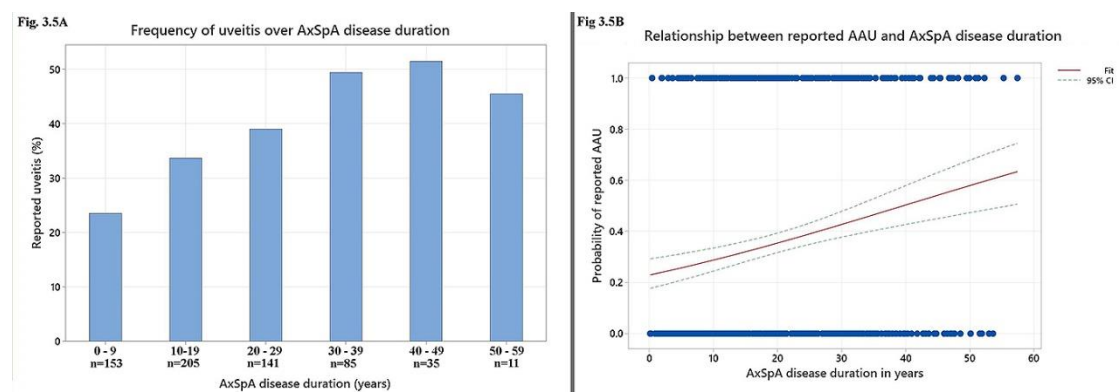


Fig. 3.5: The relationship between AAU and AxSpA disease duration.

Fig. 3.5A: Percentage of patients with history of AAU broken down by AxSpA disease duration, at 10 year intervals.

Fig. 3.5B: Binary fitted line of AAU frequency for AxSpA disease duration with 95% confidence intervals.

Genome wide association returned three markers of interest, namely rs649214, rs1370232 and rs11656399, but none reached genome wide significance (Fig. 3.6). Furthermore, none of these variants were associated with AAU in a recent GWAS (Huang et al., 2020). A QQ plot (Fig. 3.7) showed no clear deviation of P-values from the expected distribution pattern, suggesting that these observations were most likely the result of the normal distribution.

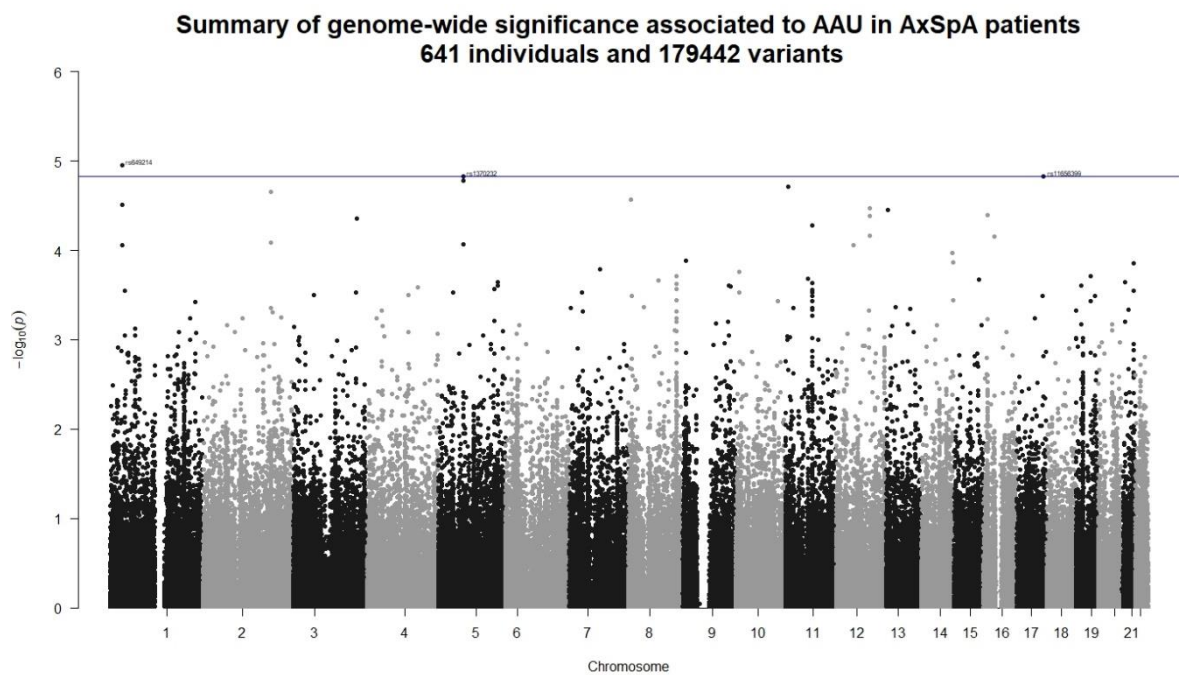


Fig. 3.6: Genome wide significance of AAU in the ASRI patient cohort, based on 225 patients with and 405 without AAU.

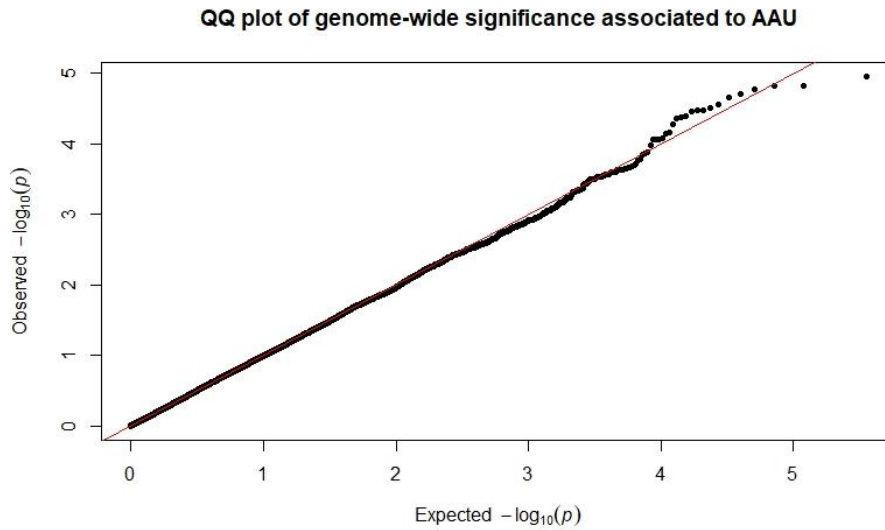


Fig. 3.7: QQ plot of genetic significance associated with AAU in AxSpA patients.

3.4.4 Psoriasis

Our analysis of the ASRI cohort revealed a psoriasis prevalence rate of 17%. No relationship between the prevalence of psoriasis and AxSpA disease duration was observed, suggesting that psoriasis onset tends to precede or shortly follow that of AxSpA (Fig. 3.8). Psoriasis is known to have shared genetic risk with AS of approximately 50% (Ellinghaus et al., 2016). An association between HLA-B27 and psoriatic arthritis has previously been documented, but the predominant HLA-type association in psoriasis is HLA-Cw6, which is linked to the characteristic dermatological manifestations of the disease (Queiro et al., 2015). Although a previous GWAS study (Strange et al., 2010) found a strong association (O/R 4.66) between HLA-Cw6 tag variant rs10484554 and psoriasis, we only observed a weak enrichment of this variant (rs10484554) in the ASRI psoriasis sub-cohort, which did not reach statistical significance ($P=0.093$, OR 1.62, Fig. 3.9). No other markers within the HLA region approached statistical significance. A QQ-plot of the P-value distribution showed only minor deviations from normal distribution, with no major outliers (Fig. 3.10).

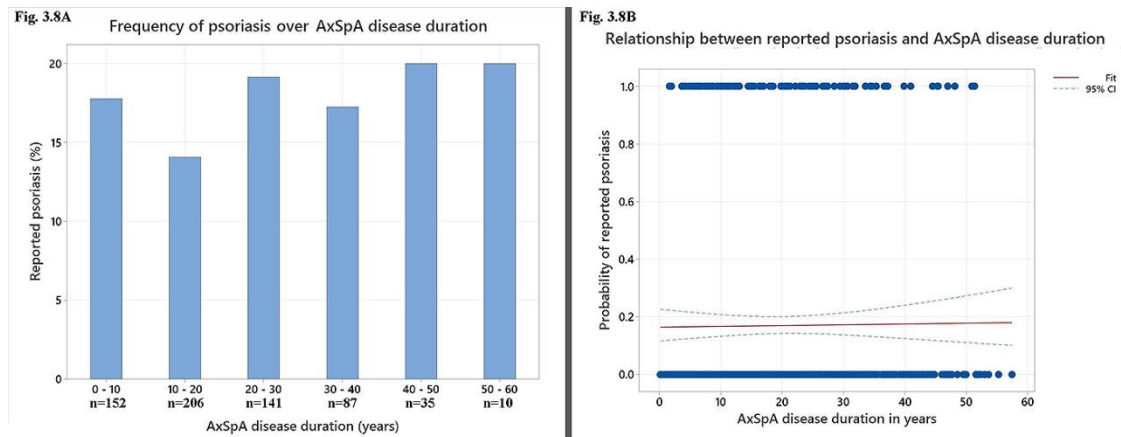


Fig. 3.8: The relationship between psoriasis and AxSpA disease duration.

Fig. 3.8A: Bar chart of psoriasis frequency for AxSpA disease duration binned at 10 yearly intervals.

Fig. 3.8B: Binary fitted line of psoriasis frequency for AxSpA disease duration with 95% confidence intervals.

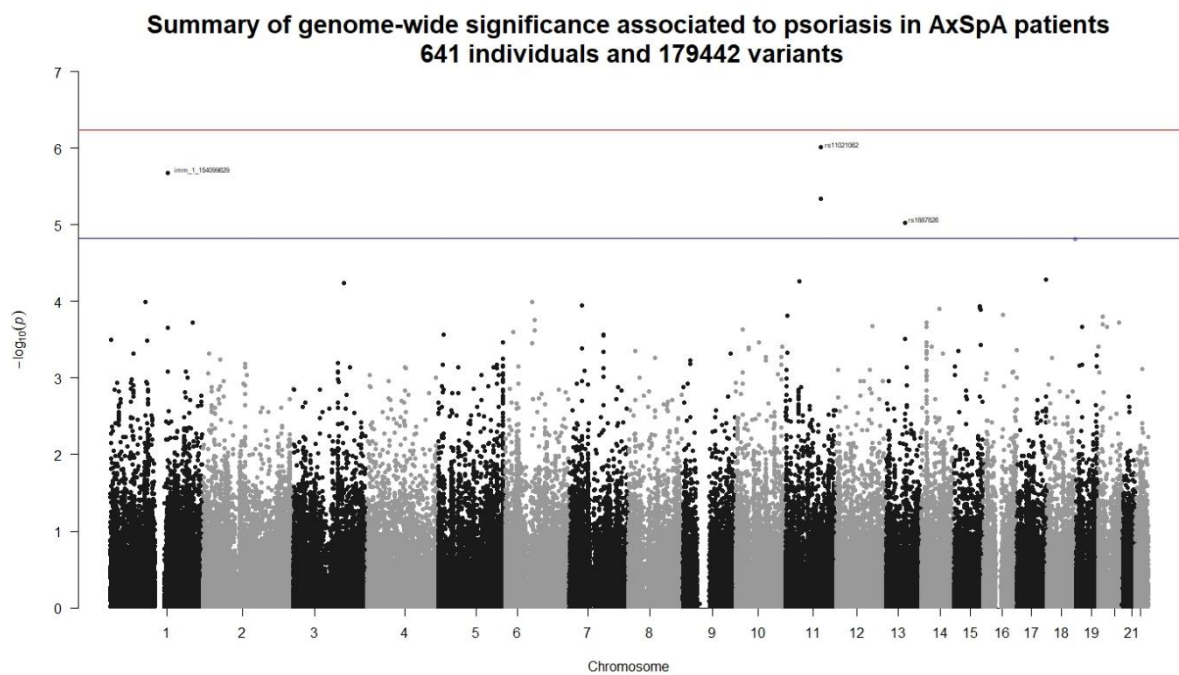


Fig. 3.9: Genome wide significance of psoriasis in the ASRI patient cohort, based on 107 patients with and 524 patients without psoriasis.

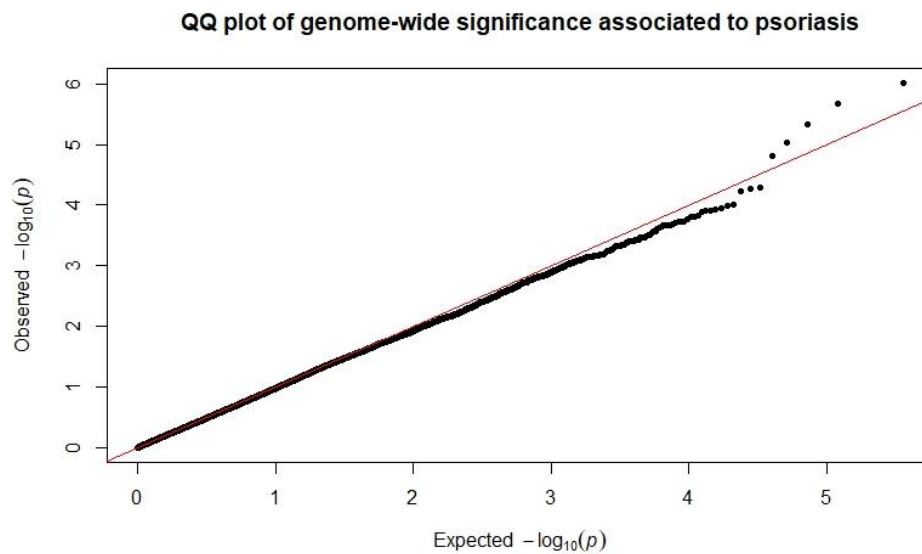


Fig. 3.10: QQ plot of the genetic significance associated with psoriasis in AxSpA patients.

To investigate our findings further, we performed a cross comparison of the allele frequencies of the HLA-Cw6 tag variant, rs10484554, between ASRI and the aforementioned GWAS study performed on psoriasis patients by Strange *et al.* Data was analysed based on a dominant model of association for rs10484554, as reported in the psoriasis GWAS study (Strange *et al.*, 2010). We found markedly reduced prevalence of the rs10484554-T risk allele in our sub-cohort of AxSpA patients with concurrent psoriasis compared to psoriasis patients in the Strange *et al* study ($P < 0.0001$, Table 3.8). Furthermore, the prevalence of rs10484554-T was lower in our psoriasis sub-cohort than in healthy controls from the Strange *et al* study ($P = 1.34 \times 10^{-9}$). Notably, HLA-B27 positivity was also lower in AxSpA with concurrent psoriasis (77.6%), compared to AxSpA patients without psoriasis (86.1%, $P = 0.038$).

Table 3.8: A comparison of rs10484554-T positivity between ASRI patients with and without concurrent psoriasis and cases and controls of the Strange *et al* psoriasis GWAS.

rs10484554-T	Psoriasis			Control (without psoriasis)		
	+	-	% Positive	+	-	% Positive
ASRI AxSpA	8	99	7.5	26	498	5.0
Strange <i>et al</i>	1347	831	61.8	1190	3985	23.0

To rule out the possibility that rs10484554-T is a poor marker for HLA-Cw6 in the Irish population, we used the HIBAG methodology described in Methods Section 3.3.9 to perform expanded HLA-typing of the ASRI patient cohort. This approach considers all available SNPs within the HLA-region to predict haplotypes, offering a more robust HLA-type prediction model than relying on a single tag SNP.

HLA-types determined using the HIBAG prediction model were used to assess the presence of the HLA-Cw6 allele in our psoriasis and non-psoriasis sub-cohorts. This approach indicated that 15 (14%) of AxSpA patients with concurrent psoriasis and 31 (5.8%) of AxSpA patients without psoriasis carried the HLA-Cw6 gene (O/R 2.59, P=0.007). This represents a large increase in the proportion of AxSpA patients with concurrent psoriasis predicted to carry the HLA-Cw6 gene compared to using rs10484554 alone (14% vs 7.5%), suggesting that rs10484554 was under-representative of HLA-Cw6 within our population. We examined the odds of B27 and Cw6 HLA pairings between AxSpA patients with and without psoriasis to identify how these two markers are likely to interact. The results of this examination are presented in Table 3.9. HLA-Cw6 in isolation was positively associated to disease but this did not reach statistical significance (O/R: 2.38, P=0.078), while B27 was inversely associated (O/R: 0.48, P=0.003), as previously described. HLA-B27 when coupled to Cw6 was positively associated with psoriasis (O/R: 2.57, P=0.046). Most interestingly a double negative genotype for B27 and Cw6 was more frequent in the psoriasis sub-cohort (O/R: 1.52), although this did not reach statistical significance (P=0.188).

Table 3.9: Psoriasis HLA-B27 and HLA-Cw6 summary

	Psoriasis			No psoriasis		O/R	P ¹
	n	%		n	%		
Total patients	107	100		524	100	N/A	N/A
B27+	83	77.6	✓	451	86.1	0.56	0.038
Cw6+	15	14.0	✓	31	5.9	2.59	0.007
B27+ Cw6+	8	7.5	✓	16	3.1	2.57	0.046
B27+ Cw6-	75	70.1	✓	435	83.0	0.48	0.003
B27- Cw6+	7	6.5	✓	15	2.9	2.38	0.078
B27- Cw6-	17	15.9	✓	58	11.1	1.52	0.188

¹Fisher's exact.

We noted that despite the observed increase in HLA-Cw6 within our psoriasis sub-cohort using the HIBAG methodology for predictive typing, HLA-Cw6 was still much less frequent in AxSpA patients with concurrent psoriasis than in psoriasis patients reported within the Strange *et al* study (14% vs 61%, $P < 0.0001$). Therefore, as a final investigation, we drew upon Irish control genomic sequencing data of 1,911 individuals (Chapter 4, Methods section 4.3.1) to identify whether the findings in the comparison of our patient cohort to the results of the psoriasis GWAS study by Strange *et al* could be confirmed. We found a prevalence of 11.59% for HLA-Cw6 tag SNP, rs10484554, in the Irish control population, compared to 8.68% in AxSpA patients with psoriasis. This was not found to be statistically significant ($P = 0.122$). We noted, however, that the prevalence of rs10484554 was much lower in the Irish control population than that reported in controls for the psoriasis GWAS study by Strange *et al* (11.59% vs 22.9%, $P < 0.0001$). We therefore used HIBAG to predict HLA-Cw6 in both ASRI patients and 1,911 Irish controls, the results for which are presented in Table 3.10. We found that HLA-Cw6 was not enriched in AxSpA patients with psoriasis, compared to Irish controls (frequency 0.1402 vs 0.1664, $P = 0.449$). We further showed that estimated HLA-Cw6 prevalence in Irish controls was increased using the HIBAG methodology as opposed to using the tag variant, rs10484554 alone (16.64% vs 11.59%, $P < 0.001$).

Table 3.10: HLA-Cw6 carriage in AxSpA Patients and Irish controls

Group	n	Cw6+	Cw6-	% Cw6	P (vs Irish controls)
AxSpA (All)	641	47	594	7.33	<0.001
AxSpA with concurrent psoriasis	107	15	92	14.02	0.449
AxSpA without concurrent psoriasis	530	31	499	5.85	<0.001
Irish controls	1911	318	1593	16.64	N/A

HLA-types determined using HIBAG.

Note: Psoriasis data was unavailable for 4 AxSpA patients.

In summary, our data indicates that HLA-Cw6 is enriched in AxSpA patients with psoriasis, compared to AxSpA patients without psoriasis. However, HLA-Cw6 prevalence in AxSpA patients with psoriasis is lower than that of psoriasis patients in the Strange *et al* GWAS, irrespective of the methodology used to predict HLA-Cw6 ($P < 0.0001$). The level of HLA-Cw6 in AxSpA patients with concurrent psoriasis was similar to that found within the Irish control group (14.0% vs 16.6%, $P = 0.449$).

3.4.5 Inflammatory bowel disease

In the genotyped ASRI cohort, IBD (collectively UC and CD) was reported in 10.4% of patients. Our data did not allow for differentiation of UC and CD within the cohort. IBD is likely to be linked to the duration of AxSpA, as previously reported in Chapter 2 ($P=0.003$, Table 2.3). This pattern was again evident within the genotyped sub-cohort (Fig. 3.11), but due to the small sample size, particularly among those with longer disease duration, the correlation was not statistically significant (Fig. 3.11). Given the differences in genetic risk between UC and CD and their sub-phenotypes (Cleynen et al., 2016), it was not expected that markers of large effect size would be identified, and Manhattan and QQ plots (Fig. 3.12 and 3.13) show only a small number of isolated SNPs deviating from the expected P-value distribution. The isolated top SNP, rs7136752, was found to have poor genotype clustering. None of the remaining SNPs have been previously linked to UC or CD, and were deemed to be likely artefacts of stochastic variation. As a result, no further investigations were conducted.

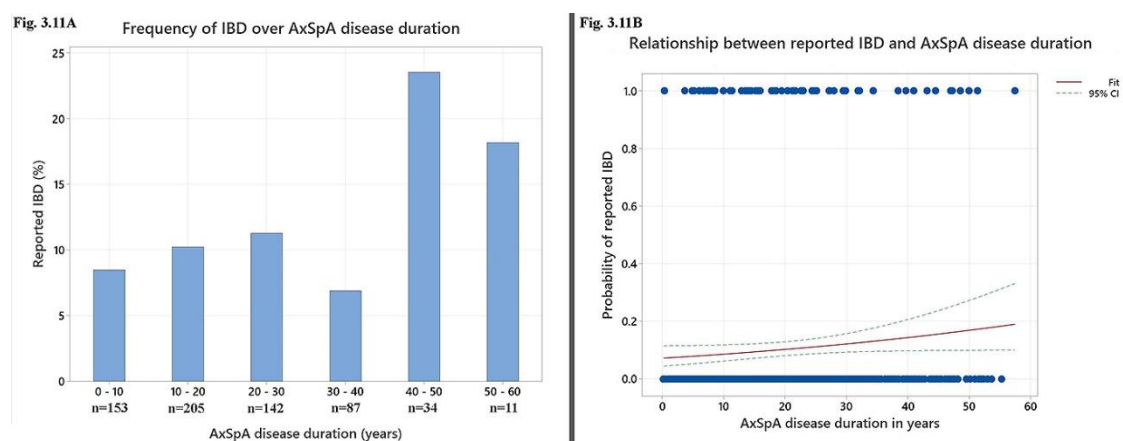


Fig. 3.11: The relationship between IBD and AxSpA disease duration.

Fig. 3.11A: Bar chart of frequency by AxSpA disease duration grouped by 10 yearly intervals.

Fig. 3.11B: Binary fitted line of IBD by AxSpA disease duration.

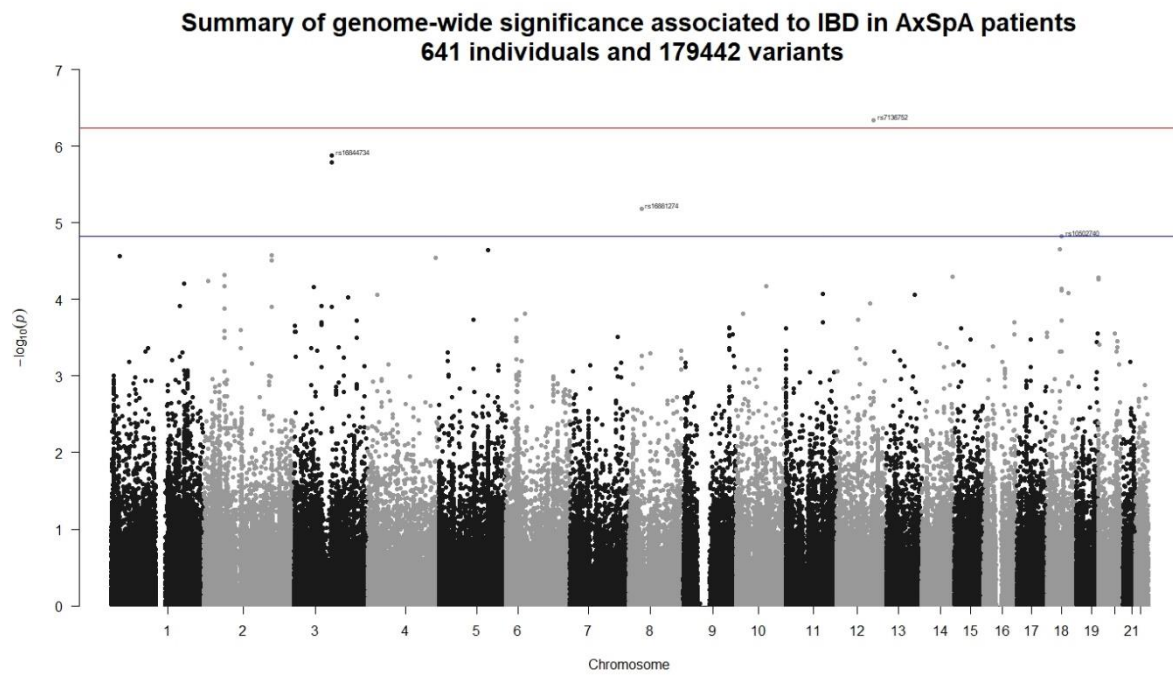


Fig. 3.12: Genome wide significance of IBD in the ASRI patient cohort, based on 66 patients with and 566 without IBD.

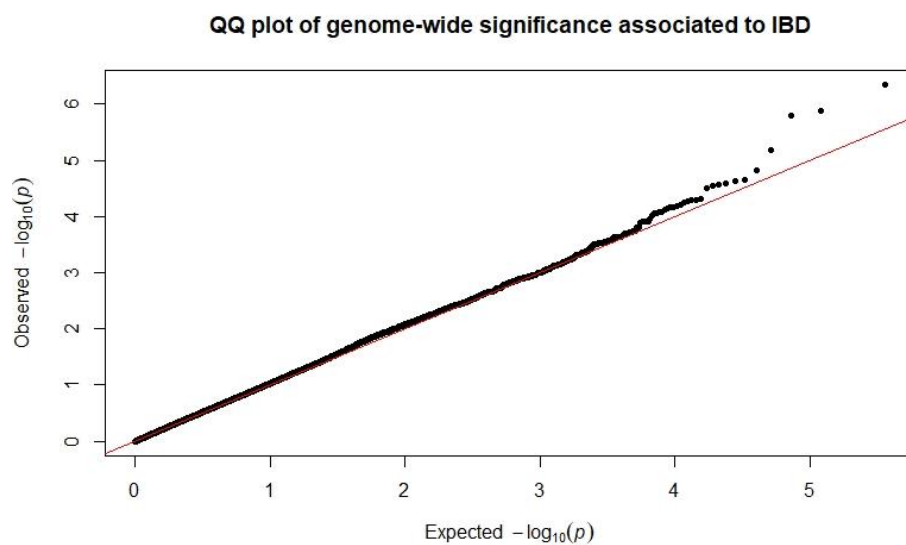


Fig. 3.13: QQ plot of the genetic significance associated with psoriasis in AxSpA patients.

3.4.6 Early disease onset

Previous research has linked the onset of AxSpA before the age of 21 to differences in radiographic disease progression (Riley et al., 1971). We investigated whether the effects of

early disease onset had a genetic component by comparing patients who developed AxSpA before 21 years of age to those who developed it at 21 years or older. We identified a lead variant on chromosome 3, rs73057527 (1kg_3_28037611, Fig. 3.14), which was statistically significant and accompanied by several other SNPs reaching into the significant range. A QQ-plot of the data supported a genetic component to early AxSpA disease onset (Fig. 3.15). The lead SNP and linked variants map to the intergenic region between *EOMES* and *CMC1* where a series of deletions commonly occur. *CMC1* has previously been implicated in AxSpA through meta-analysis of inflammatory diseases, while *EOMES* was implicated in our own meta-analysis of AS and AxSpA (Chapter 4, Section 4.4.1). We found no evidence of significant linkage disequilibrium between rs73057527 and either the previously reported SNP near *CMC1*, rs10510607 ($r^2=0.001$, $D'=0.213$) (Ellinghaus et al., 2016) or the novel SNP, rs13093489 ($r^2<0.001$, $D'=0.137$), identified near the *EOMES* gene in our own meta-analysis. We furthermore found no evidence of previously recorded associations of rs73057527 to any disease type or previously reported eQTL's (Carithers and Moore, 2015) at this locus. Rs73057527 was not available within our meta-analysis data, and no SNPs in significant LD were present. Due to a lack of supporting evidence, no further investigation was performed on this SNP.

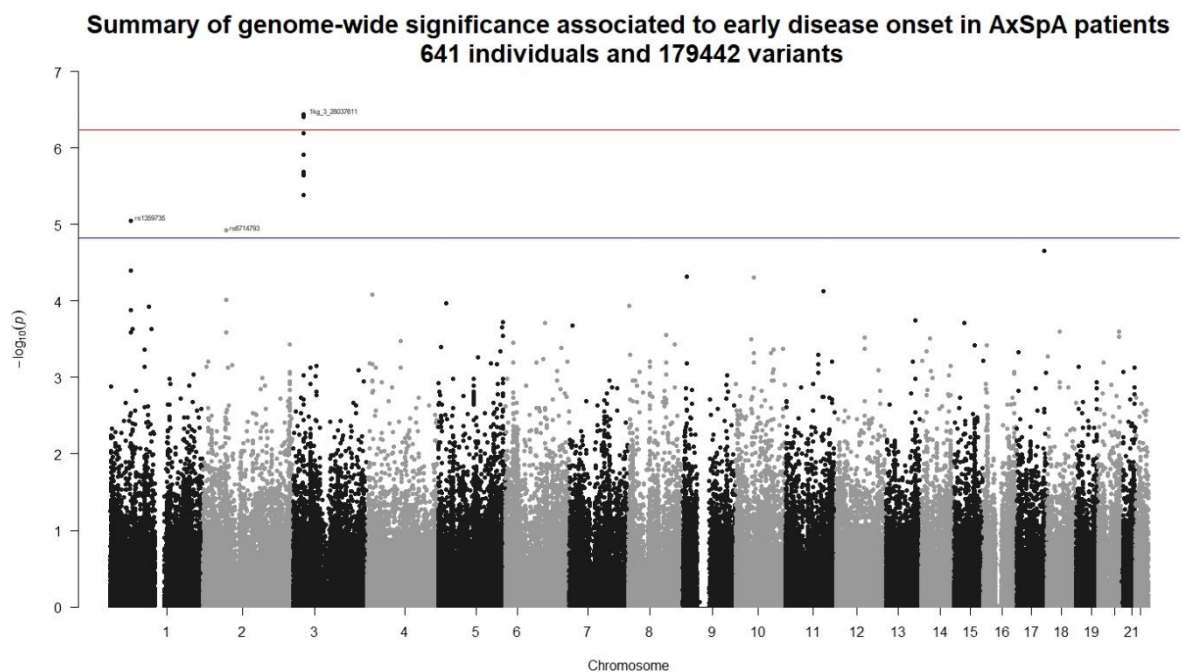


Fig. 3.14: Genome wide significance of early disease onset (<21 years) in the ASRI patient cohort, based on 238 patients with early onset and 403 with later onset.

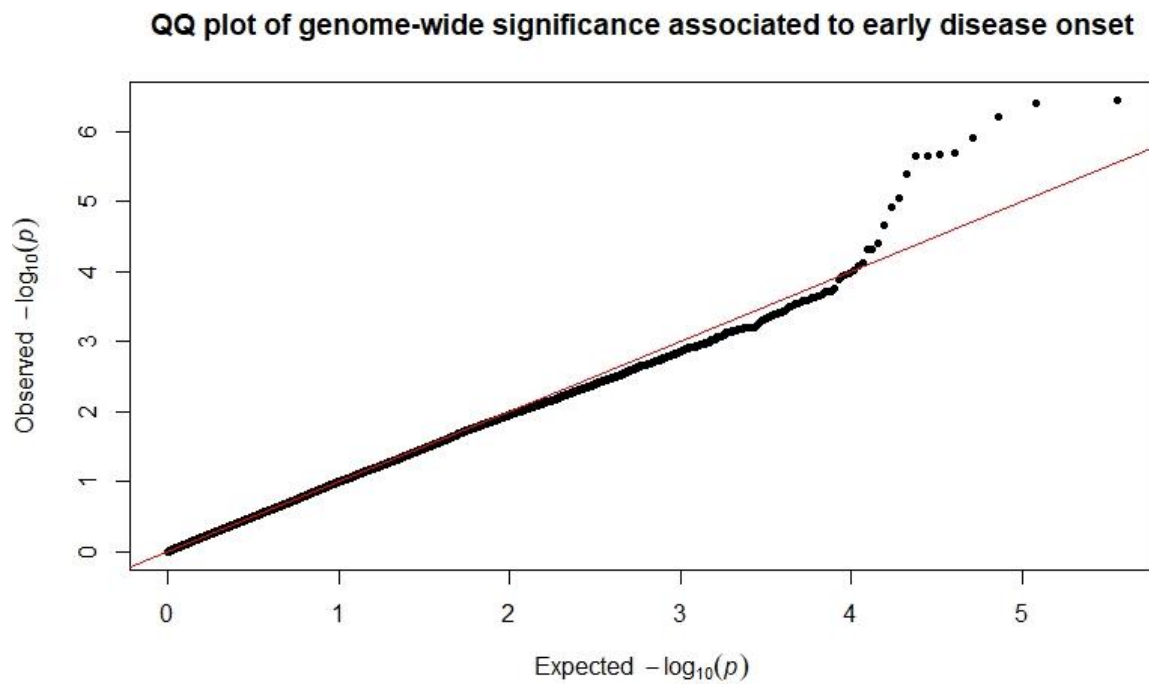


Fig. 3.15: QQ plot of the genetic significance associated with early disease onset (<21 years) in AxSpA patients.

We did not find any clear pattern of genome wide association in relation to gender, family history, or mNY criteria. Manhattan and QQ-plots for these are available in the supplementary data for this chapter (Figs S3.1 – S3.6).

3.5 Results: The role of the ERAP1 gene in clinical phenotype

A total of 641 participants were included for analysis following quality control, with 497 (77.5%) of them being male. Among these participants, 496 (77.4%) met the modified New York criteria for AS. The median age of the participants was 45 years, ranging from 18 to 85 years old. The median reported age of symptom onset (IQR) was 24 years (19-32), and the average age at diagnosis (IQR) was 34 years (27-42). On average, there was a 6-year delay to diagnosis (3-12). Extra-articular manifestations were reported in 319 (49.8%) of participants in the form of uveitis, inflammatory bowel disease, or psoriasis. These data are summarised in Table 3.11.

Table 3.11: General characteristics of participants genotyped for *ERAP1* study.

Characteristic	Included n	Median (IQR) or n (%)	Range
Male n (%)	641	497 (77.5)	
Age in years, median (IQR)	641	45 (37 - 55)	(18 - 85)
Age at symptom onset, median (IQR)	638	24 (19 - 32)	(6 - 77)
Age at diagnosis, median (IQR)	641	34 (27 - 42)	(9 - 80)
Delay to Diagnosis in years, median (IQR)	638	6 (3 - 12)	(0 - 52)
Disease duration in years, median (IQR)	641	18 (10 - 28)	(0 - 57)
HLA-B27 positive, n (%)	641	543 (84.7)	
ASAS criteria for AxSpA, n (%)	641	641 (100)	
mNY Criteria for AS, n (%)	641	496 (77.4)	
Uveitis*, psoriasis or IBD, n (%)	641	319 (49.8)	
Psoriasis, n (%)	631	107 (17.0)	
Uveitis*, n (%)	630	225 (35.7)	
IBD, n (%)	632	66 (10.5)	
BASDAI score, median (IQR)	617	3.9 (2.1 - 5.8)	(0 - 10)
BASMI score, median (IQR)	640	3.6 (2.4 - 5.4)	(0 - 9)
BASFI score, median (IQR)	598	3.5 (1.6 - 5.6)	(0 - 10)
ASQoL score, median (IQR)	641	5.0 (1.0 - 11.0)	(0 - 18)
HAQ-s score, median (IQR)	634	0.5 (0.1 - 0.8)	(0 - 3)

*At least one reported episode of acute anterior uveitis.

3.5.1 Phenotypic characteristics associated with HLA-B27 in AxSpA

Due to the close relationship between HLA-B27 and ERAP1, we first investigated the effects of HLA-B27 on clinical phenotype within the ASRI cohort. The impact of HLA-B27 on clinical outcomes has been well documented in AS, but less so in AxSpA as a whole. To evaluate the effects of HLA-B27 on clinical outcomes in the ASRI cohort, we performed univariate and multivariate regression analysis to compare clinical characteristics between HLA-B27 positive (B27+) and HLA-B27 negative (B27-) AxSpA patients. Our analysis included 641 participants, of which 543 (84.7%) were HLA-B27 positive (B27+). Our results, presented in Table 3.12, show that B27+ patients were more likely to have a recorded family history of AS (30% vs 17%, $P=0.01$). Although no difference in HLA-B27 positivity by sex was noted in univariate analysis, multivariate analysis suggested increased HLA-B27 positivity in males ($P=0.019$). B27+ participants were more likely to suffer from uveitis ($P<0.001$), but psoriasis was more likely in B27- patients ($P=0.038$), suggesting that the presence of HLA-B27 plays a protective role against concomitant psoriasis in AxSpA patients. No significant differences were observed in patient reported outcomes (PROs) and clinical measurement scores (BASDAI, BASMI, BASFI, ASQoL or HAQ-s), and laboratory measurements of inflammation (CRP and ESR) were similar between the groups. Although HLA-B27 was associated with an increased likelihood of AAU, we found no difference in the age of B27+ and B27- patients with reported AAU episodes ($P>0.05$), as shown in Fig. 3.16A. AS disease duration was lower in HLA-B27- patients with reported AAU (Fig. 3.16B), but this did not reach statistical significance ($P=0.057$). Since there was no overall difference in the age of patients with a history of AAU, this may suggest that AAU onset is independent of concurrent AxSpA. Some studies have found that patients who met the modified New York (mNY) criteria for AS were more likely to be HLA-B27+ than their non-radiographic counterparts (Glintborg et al., 2017, Hong et al., 2019) but we saw no evidence for this within our own cohort.

Table 3.12: Univariate and multivariate analysis of HLA-B27+ and HLA-B27- AxSpA patients.

	HLA-B27 Positive		HLA-B27 Negative		ChiSq** or Mann Whitney U (P)	Multivariate (P)***
	Included n	Median (IQR) or n (%)	Included n	Median (IQR) or n (%)		
Male (%)	543	425 (78.3)	98	72 (73.5)	0.295	0.019
Family history, n (%)	543	163 (30.0)	98	17 (17.3)	0.01	0.034
Age in years, median (IQR)	543	45 (37 - 54)	98	48 (38 - 58)	0.049	--
Age at first symptom onset, median (IQR)	541	22 (18 - 30)	97	32 (24 - 42)	<0.001	--
Age at diagnosis, median (IQR)	543	33 (26 - 41)	98	41 (33 - 49)	<0.001	<0.001
mNY criteria* n (%)	543	416 (76.6)	98	80 (81.6)	0.297	0.408
Uveitis**, n (%)	533	209 (39.2)	97	16 (16.5)	<0.001	<0.001
Psoriasis, n (%)	534	83 (15.5)	97	24 (24.7)	0.038	0.007
Inflammatory bowel disease, n (%)	535	51 (9.5)	97	15 (15.5)	0.102	0.103
Depression, n (%)	542	56 (10.3)	98	10 (10.2)	1	--
Clinical Instruments:						
BASDAI score, median (IQR)	521	3.8 (2.1 - 5.7)	96	4.2 (2.3 - 6.0)	0.198	--
BASMI score, median (IQR)	543	3.6 (2.4 - 5.6)	97	3.6 (2.4 - 5.0)	0.251	--
BASFI score, median (IQR)	506	3.5 (1.6 - 5.6)	92	4.0 (2.0 - 6.1)	0.255	--
ASQoL score, median (IQR)	543	5.0 (1.0 - 10.5)	98	7.0 (1.0 - 12.0)	0.285	--
HAQ-s score, median (IQR)	537	0.5 (0.1 - 0.9)	97	0.4 (0.0 - 0.6)	0.137	--
CRP, median (IQR)	462	10 (5 - 20)	82	10 (5 - 20)	0.102	--
ESR, median (IQR)	485	3.0 (1.0 - 8.0)	87	2.0 (0.9 - 6.0)	0.908	--

*Refers to modified New York criteria 1984 for AS.

**Fisher's exact.

***Multivariate by binary logistic regression with HLA-B27 as response.

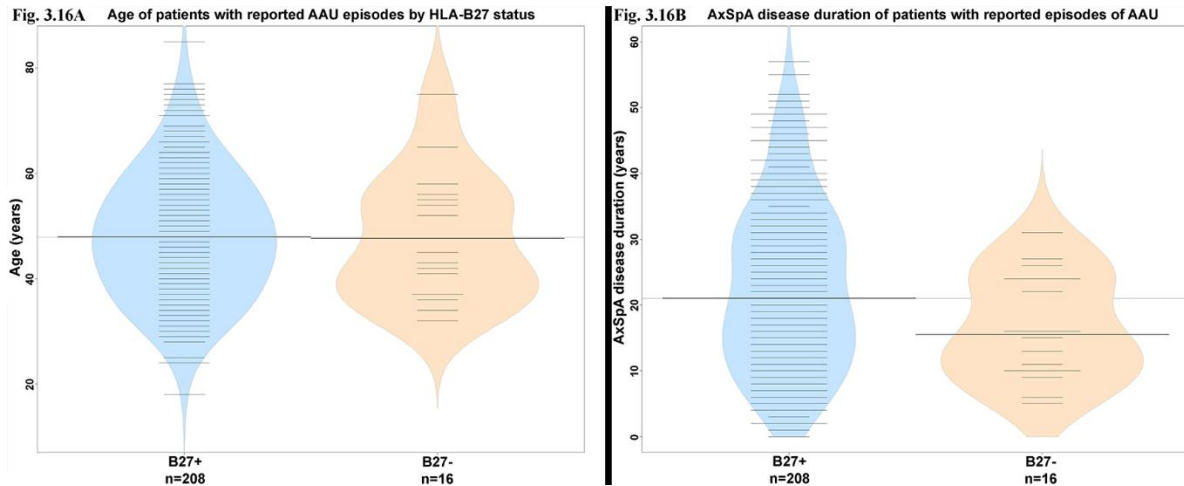


Fig. 3.16: Distribution of patient age and AxSpA disease duration in AxSpA patients with a history of AAU.

Fig. 3.16A: Age of patients with reported AAU episodes by HLA-B27 status

Fig. 3.16B: AxSpA disease duration of patients with reported AAU episodes by HLA-B27 status.

3.5.2 Phenotypic characteristics associated with ERAP1 variant rs30187

To investigate the effects of the *ERAP1* variant, rs30187, we examined the clinical characteristics of HLA-B27 positive individuals by rs30187 genotype. The results of this investigation are presented in Table 3.13.

Among the 543 B27+ participants, 362 (66.67%) carried the previously reported risk variant, rs30187 (rs30187-T), with 84 being homozygous for the risk T-allele. The *ERAP1* rs30187-T allele was enriched in B27+ compared to B27- participants (41.1% vs 28.6%, $P < 0.001$). Rs30187-T carriers were younger at AxSpA diagnosis compared to non-carriers (23.3 vs 25.7 years, $P < 0.001$), with a similar trend observed for age of first symptom onset, but this was not statistically significant.

In terms of extra-articular manifestations (EAMs), AAU was more frequent in rs30187-T homozygotes (48.8%) compared to heterozygotes (40%) and non-carriers (33.5%), $P = 0.017$. The age of AxSpA patients with reported episodes of AAU was slightly lower in patients carrying the rs30187-T allele, but this was not significant (Fig. 3.17A). There was also no trend observed in the AxSpA disease duration of patients with reported AAU episodes based on rs30187 genotype

(Fig 3.17B). No significant differences were observed in other EAMs based on rs30187 genotype. There were no significant differences in PROs and inflammatory measurements (BASDAI, BASMI, BASFI, ASQoL, HAQ-s, ESR, or CRP) between rs30187 carriers and non-carriers.

Table 3.13: Univariate and multivariate analysis of rs30187 allelotypes in HLA-B27+ patients.

Characteristic	Genotype CC		Genotype CT		Genotype TT		Cochran-Armitage or Ordinal Logistic Regression (P)	Ordinal Multivariate Analysis (P)
	Included n	Median (IQR) or n (%)	Included n	Median (IQR) or n (%)	Included n	Median (IQR) or n (%)		
Male (%)	181	140 (77.3)	278	216 (77.7)	84	69 (82.1)	0.434	0.561
Family history, n (%)	181	61 (33.7%)	278	80 (28.8%)	84	22 (26.2%)	0.171	0.219
Age in years, median (IQR)	181	46 (38 - 56)	278	44 (37 - 53)	84	43 (36 - 53)	0.059	--
Age at first symptom onset, median (IQR)	181	24 (18 - 31)	278	23 (19 - 30)	84	22 (19 - 27)	0.066	--
Age at diagnosis, median (IQR)	181	37 (27 - 45)	278	32 (26 - 39)	84	32 (26 - 36)	<0.001	<0.001
mNY criteria* n (%)	181	135 (74.6%)	278	219 (78.8%)	84	62 (73.8%)	0.844	0.465
Uveitis**, n (%)	179	60 (33.5%)	270	108 (40.0%)	84	41 (48.8%)	0.017	0.004
Psoriasis, n (%)	179	28 (15.6%)	272	46 (16.9%)	83	9 (10.8%)	0.472	0.456
Inflammatory bowel disease, n (%)	178	19 (10.7%)	273	23 (8.4%)	84	9 (10.7%)	0.821	0.703
Depression, n (%)	181	21 (11.6%)	278	30 (10.8%)	83	5 (6.0%)	0.218	
Clinical Instruments:								
BASDAI score, median (IQR)	174	4.0 (2.3 - 6.0)	264	3.8 (1.9 - 5.6)	83	3.4 (2.0 - 5.5)	0.203	--
BASMI score, median (IQR)	181	3.6 (2.8 - 5.6)	278	3.6 (2.2 - 5.6)	84	3.4 (2.2 - 5.1)	0.284	--
BASFI score, median (IQR)	170	4.0 (1.8 - 5.7)	259	3.3 (1.7 - 5.6)	77	2.8 (1.1 - 4.8)	0.051	0.168
ASQoL score, median (IQR)	181	4.0 (1.0 - 11.0)	278	5.0 (1.0 - 11.0)	84	5.0 (1.0 - 7.3)	0.544	--
HAQ-s score, median (IQR)	178	0.50 (0.13 - 0.75)	276	0.50 (0.13 - 0.88)	83	0.38 (0.00 - 0.63)	0.351	--
CRP, median (IQR)	159	3.1 (1.0 - 8.0)	252	3.0 (1.0 - 7.8)	74	2.0 (1.0 - 8.3)	0.483	--
ESR, median (IQR)	147	10.0 (5.0 - 19.5)	244	10.0 (5.0 - 19.3)	71	13.0 (5.0 - 21.5)	0.331	--

* Refers to modified New York criteria 1984 for AS.

** At least one reported episode of acute anterior uveitis.

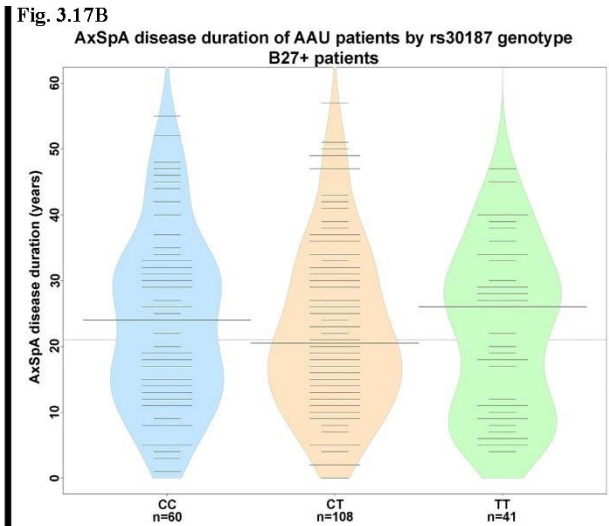
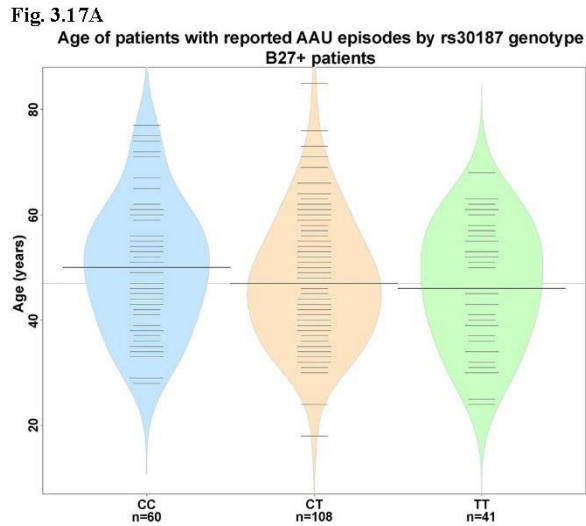


Fig. 3.17: Distribution of age and AxSpA disease duration subdivided by ERAP1 rs30187 genotype.

Fig. 3.17A: Age of patients with reported AAU episodes by rs30187 genotype.

Fig. 3.17B: AxSpA disease duration of patients with reported episodes of AAU by rs30187 genotype.

We additionally investigated the potential influence of rs30187-T on the clinical phenotype of B27- AxSpA. Of the 98 B27- individuals, 54 (55.1%) carried the rs30187-T allele, with two homozygous for the risk allele. We observed no significant differences in age at diagnosis, incidence of AAU, or any clinical measurements. No trends were observed based on rs30187-T allele dosage. This data is available in Supplementary table S3.2.

Overall, our study underscores the close relationship between HLA-B27 and *ERAP1*. HLA-B27 and *ERAP1* rs30187-T work together to promote earlier AxSpA disease onset and the development of comorbid uveitis. However, our findings indicate that not all the characteristics associated with HLA-B27 are mirrored by *ERAP1* rs30187-T. For example, rs30187 was not associated with a family history of AxSpA and did not exhibit the inverse association with psoriasis that was observed with HLA-B27, although this may be due to the limited sample size of our study. Finally, we did not find evidence supporting the role of either HLA-B27 or *ERAP1* in disease activity, physical disability, or quality of life.

3.6 Results: Investigation of causative haplotypes within the ERAP1 gene

We processed 252 representative SNPs within the *ERAP1-ERAP2-LNPEP* region (Chr5: 96,000,000 – 96,500,000, GRCh37, hg19) for linkage analysis. A visualization of LD across the region is presented in Fig. 3.18. We identified a total of 16 logical haplotype blocks within this region (Fig. 3.19). *ERAP1*'s coding region was located between blocks 6-7, with previously reported variants, rs30187, rs27434, rs17482078, and rs2287987, all located in block 7. We used haplotype analysis to investigate AxSpA association, with permutations (n=1,000) to correct for multiple testing. Blocks 6, 7, 8 and 16 exhibited the strongest AxSpA association (Supplementary Table S3.1). Here, we focussed on variants within the coding region of the *ERAP1* gene and therefore did not investigate potential risk variants in blocks 8 and 16.

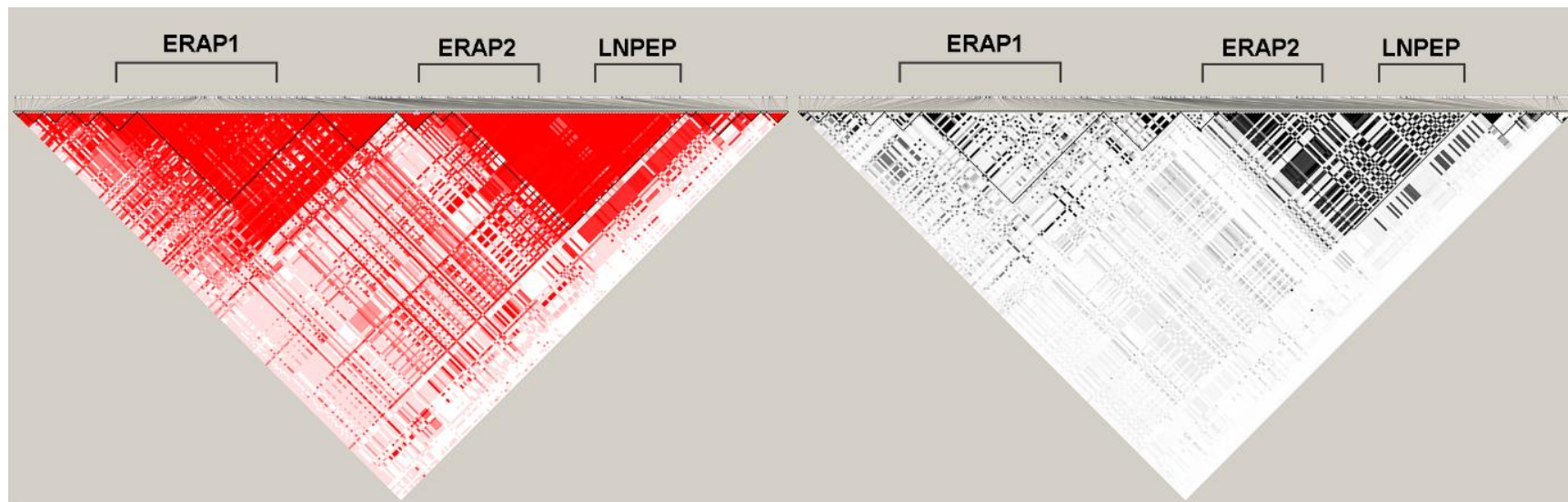


Fig. 3.18: LD structure within the ERAP1-LNPEP region extending from Chr5: 96,000,000 – 96,500,000, (GRCh37, hg19). Left: D' represented by colour intensity. Right: r^2 represented by greyscale intensity.

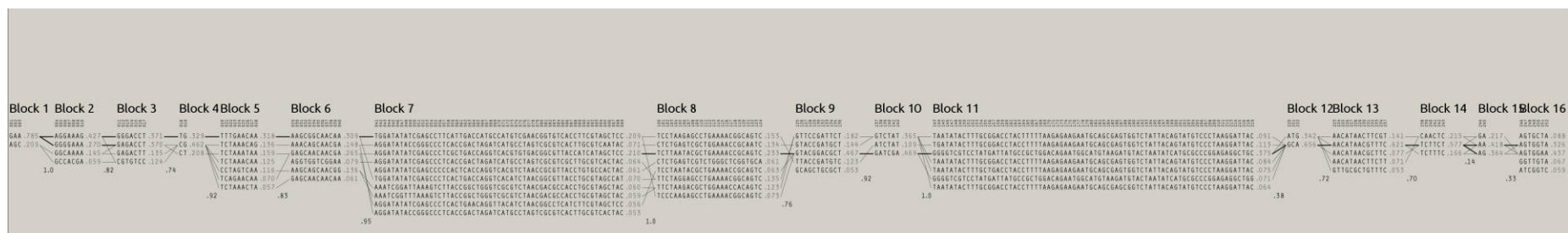


Fig. 3.19: Haplotype blocks within ERAP1 500KB region as determined by Gabriel et al methodology (Gabriel et al., 2002) using 95% confidence, based on 248 representative SNPs chosen at random plus 4 SNPs from previously published work.

3.6.1 Investigation of rs30187 as a causative SNP in AxSpA

To assess the underlying association of rs30187 with AxSpA, we constructed haplotypes based on 50 representative SNPs, including rs30187, to decipher the pattern of AxSpA genetic risk within the *ERAP1* gene. These SNPs were taken from within Haplotype Block 7 (Fig. 3.19), which captured the region enclosing variants previously associated with AxSpA. Our expectation was that haplotypes containing the rs30187-T risk allele should exhibit a pronounced tendency of positive association with AxSpA, and vice versa. However, we found that rs30187 association to AxSpA was inconsistent among the derived haplotypes (Table 3.14). Some haplotypes with rs30187-T were associated while others showed no directionality of association.

To determine if this inconsistency was significant, we directly compared pairs of the most common haplotypes to establish whether their relative association to AxSpA was consistent, based on the presence of rs30187-T and rs30187-C (Table 3.15). We found that rs30187-C was not consistently protective against AxSpA ($P=0.001$), and tentatively, that rs30187-T is not consistently causative, although the latter did not reach statistical significance ($P=0.084$). This suggests that either rs30187 is not the sole disease-causing allele at this locus, or that there is interplay with other disease-associated SNPs sharing haplotypes with rs30187 in the region.

Table 3.14: Association of haplotypes containing rs30187 to AxSpA, based on 641 patients and 1,911 healthy controls.

Haplotype	Haplotype variants	Haplotype frequency	Haplotype ratio counts		Case,Control haplotype frequencies	Chi-square	O/R	P
			Cases	Controls				
1	CCAGGCCTCTTACCGGTCT C GGACTGCGAGTAGGTGTGAATCTAGGGTC	0.217	(224 : 1042)	(871 : 2842)	0.177, 0.234	18.216	0.70	1.97*10⁻⁵
2	CCAGGCCTCCTATAGGTCT C GGGCTGTGAGGGAGTGTGGATCCAGGGTC	0.215	(286 : 980)	(797 : 2917)	0.226, 0.215	0.722	1.07	0.396
3	CCAGGCCTGCTGCAATCTG T GGGCGATGAGTAGATGTGAACTTGAGGTC	0.128	(195 : 1071)	(450 : 3263)	0.154, 0.121	9.027	1.32	0.003
4	TTGCATCCGCTGCAATCTG T AAGATGTGGATAGGTCGGAATCTAGAGTC	0.116	(151 : 1115)	(434 : 3279)	0.119, 0.117	0.052	1.02	0.820
5	CCAGGCCTCTACAATCTG T GGGCGATGAGTAGATGTGAACTTGAGGTC	0.074	(123 : 1143)	(248 : 3465)	0.097, 0.067	12.626	1.50	4*10⁻⁴
6	CCAGGCTTCCAACCGGTCT C GGGCTGTGAGTAGGTGTAAATCTAGGGCT	0.071	(74 : 1192)	(286 : 3427)	0.058, 0.077	4.857	0.74	0.028
7	CCAGGCCTCTACAGGTCT C GGGCTGTGAGTAGGTGTGAACTTGAGTTC	0.066	(84 : 1182)	(246 : 3467)	0.066, 0.066	0	1.00	0.991
8	CCAGGCCTCTACAGGTCT C GGGCTGTAAGTAGGGGTGATTCTAGGGTC	0.059	(78 : 1188)	(220 : 3493)	0.062, 0.059	0.093	1.04	0.760

Haplotypes constructed from 49 representative variants. rs30187 allelotype enlarged in bold.

Haplotype counts rounded to nearest integer.

Table 3.15: Comparison of haplotype frequencies between 641 AxSpA patients and 1,911 healthy controls, based on rs30187 allelotype.

Haplotype	Haplotype variants	Counts of haplotypes			Frequencies (%)		O/R	Chi-square (two proportions) P
		AxSpA cases	Controls	Total	AxSpA cases	Controls		
1	CCAGGCCTCTTACCGGTCT C GGACTGCGAGTAGGTGTGAATCTAGGGTC	224	870.5	1094.5	17.7	23.4	0.70	0.001
2	CCAGGCCTCCTATAGGTCT C GGGCTGTGAGGGAGTGTGGATCCAGGGTC	286	796.5	1082.5	22.6	21.5	1.07	
3	CCAGGCCTGCTGCAATCTG T GGGCGATGAGTAGATGTGAACTTGAGGTC	151	450	601	15.4	12.1	1.32	0.084
4	TTGCATCCGCTGCAATCTG T AAGATGTGGATAGGTCGGAATCTAGAGTC	123	434	557	11.9	11.7	1.02	

Haplotypes constructed from a representative population of 49 SNP's. Rs30187 allelotype enlarged in bold.

To investigate whether additional loci within the *ERAP1* region were contributing to disease in the Irish population, we generated six partially independent tag variants which captured the strongest AxSpA disease association within the *ERAP1* coding region (Methods Section 3.3.10c). LD analysis between these tag variants and previously published SNPs, including rs30187, rs27434, rs17482078, and rs2287987, revealed that tag variant rs10050860 was strongly correlated with the previously published SNP, rs17482078 ($r^2=0.92$, $D'=1$). Unexpectedly, neither rs30187 nor any strong proxy for this SNP were selected as candidates during our tag SNP selection process for strong disease association.

To identify which SNPs were most linked to AxSpA, we performed association analysis and found that neighbouring variants rs149189 (O/R: 1.45, $P=1.06 \times 10^{-6}$) and rs149544 (O/R: 1.44, $P=1.93 \times 10^{-7}$) had the strongest positive association with AxSpA. Rs149544 is in strong LD with a previously reported missense variant, rs26653 ($r^2=0.97$, $D'=1$), and was also strongly represented in our meta-analysis (Chapter 4) for AxSpA association through a closely correlated SNP, rs34760 ($P=9.8 \times 10^{-25}$, $r^2=0.99$, $D'=0.99$). The genetic correlation between rs149544 and rs30187 is weak ($r^2=0.284$, $D'=0.605$, Fig. 3.21), supporting previous reports of an independent risk allele at this locus (Maksymowych et al., 2009).

Unexpectedly, rs30187 demonstrated a more modest association with AxSpA (OR=1.32, $P=2.88 \times 10^{-5}$) within this population, which was below the threshold used for our tag SNP selection process ($P<4.35 \times 10^{-6}$). Therefore, it was not represented in the generated tag SNPs. Overall, these findings tentatively suggest that either rs30187 is not the strongest disease-causing allele at this locus, or there is interplay with other disease-associated SNPs sharing haplotypes with rs30187 in the region.

The strongest protective associations, as measured by odds ratio, were observed for rs10050860 ($P=2.16 \times 10^{-6}$) and rs7063 ($P=3.05 \times 10^{-7}$). It is worth noting that these two SNPs were found to have incomplete linkage disequilibrium ($r^2=0.69$, $D'=0.98$, Fig. 3.20), indicating a potential degree of statistical independence between these associations. Additionally, the low LD ($r^2=0.02$, $D'=0.32$) between the positively associated marker, rs149544, and the protective SNP, rs26481, which flank the analysed region, further suggests that the protective and causative effects observed in haplotypes are not functionally linked to a single SNP in the *ERAP1* coding region.



Fig. 3.20: Tag variants (6) within the ERAP1 coding region generated by clumping methodology with $r^2 < 0.8$. Left: D' represented by colour intensity. Right: r^2 represented by greyscale intensity. Previously published SNPs added to this data are demarcated in blue. Note: rs30187 is represented by “ccc5-96150086” in this data.

Table 3.16: Novel and previously published tag SNP's associated with AxSpA

SNP	Base pair	Effect allele	Frequency AxSpA	Frequency control	Odds ratio	P
rs26481	96,103,352	T	0.31	0.39	0.71	6.83×10^{-7}
rs7063	96,110,211	T	0.24	0.31	0.69	3.05×10^{-7}
rs1065407	96,112,083	G	0.29	0.36	0.72	1.89×10^{-6}
rs17482078*	96,118,866	T	0.18	0.23	0.70	1.71×10^{-5}
rs10050860	96,122,210	T	0.18	0.25	0.68	2.19×10^{-6}
rs30187*	96,124,330	T	0.39	0.33	1.32	2.88×10^{-5}
rs27434*	96,129,512	A	0.26	0.20	1.42	3.09×10^{-6}
rs2287987*	96,129,535	C	0.18	0.24	0.70	1.14×10^{-5}
rs149189	96,137,804	A	0.25	0.19	1.45	1.06×10^{-6}
rs149544	96,157,854	C	0.33	0.26	1.44	1.93×10^{-7}

*Previously published variants.

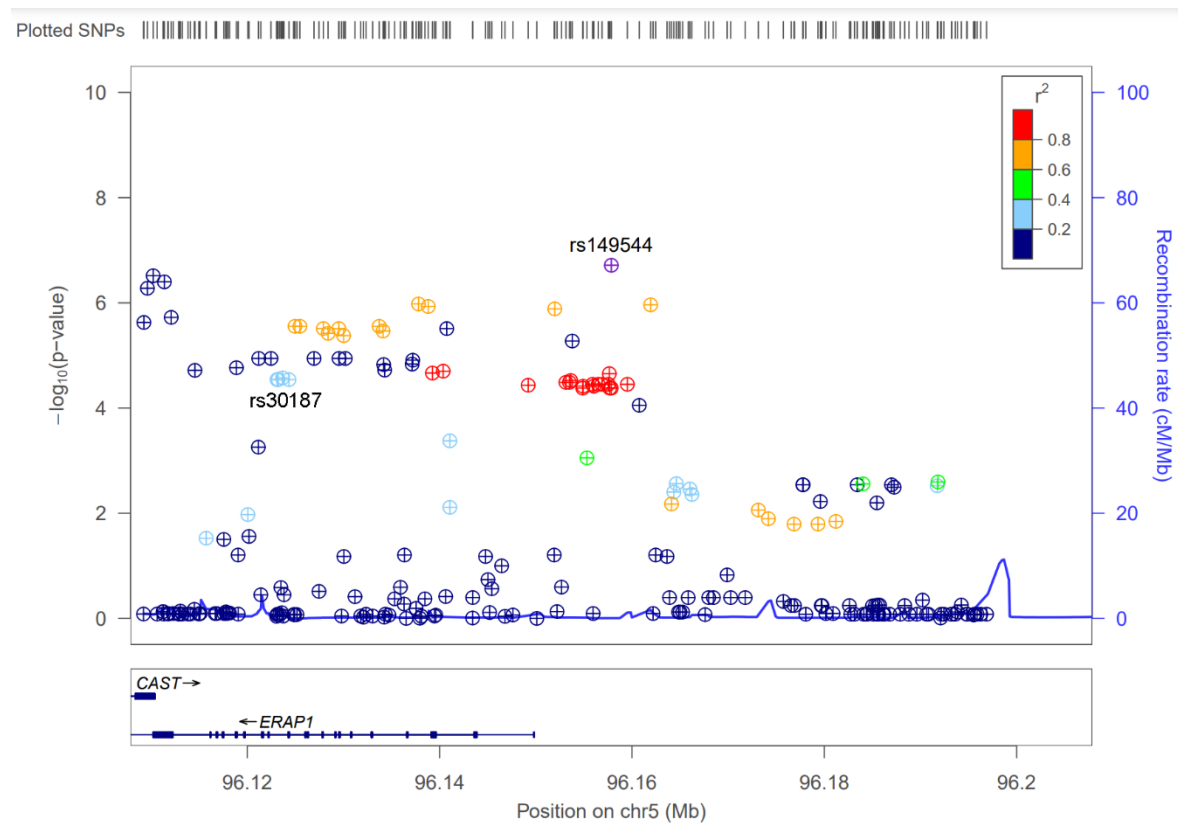


Fig. 3.21: Locus Zoom of r149544 region indicates AxSpA association and partial genetic correlation with rs30187.

To establish the degree of statistical independence of each of our associated tag SNPs, we split positively and inversely associated SNPs into two groups. We then performed association analysis for every SNP with conditioning for each of the remaining SNPs from its respective group. The results of this analysis are presented in Tables 3.17 and 3.19.

Surprisingly, we found that rs149544-G could account for disease causation attributable to all of the tag SNPs tested (Table 3.17). In contrast, rs30187-T showed no residual association after accounting for any of rs27434, rs149189, or rs149544. The SNPs rs27434 and rs149189 showed no residual association after accounting for each other as covariates, as expected due to their high LD ($r^2=0.91$, $D'=0.99$, Fig. 3.20).

Table 3.17: Statistical significance of AxSpA disease associated *ERAP1* tag SNPs after conditioning.

		Statistical significance of SNP associations (P)			
		rs30187	rs27434	rs149189	rs149544
Conditioned by SNP	Unconditioned	2.87×10^{-5}	3.06×10^{-6}	1.01×10^{-6}	1.55×10^{-7}
	rs30187	NA	0.020	0.007	2.89×10^{-4}
	rs27434	0.223	NA	0.161	0.008
	rs149189	0.258	0.948	NA	0.022
	rs149544	0.072	0.319	0.263	NA

Residual statistical significance of tag SNPs (top row) after conditioning by SNPs listed in the left column.

For example, the residual statistical significance of rs149189 after accounting for rs27434 as a covariate, $P=0.161$.

Blocks colour coded by relative statistical significance with lighter colours indicating greater significance.

Association analysis performed by logistic regression in Plink V1.9.

Due to high LD between rs149544 and the previously published sequence variant, rs26653 (Maksymowych et al., 2009b) in the ASRI cohort ($r^2=0.94$, $D'=0.99$), we aimed to determine if rs26653 was likely to be responsible for the observed genetic risk at this locus. We therefore performed genetic association analysis with conditioning between SNPs rs149544 and rs26653, as presented in Table 3.18. Despite the strong LD, a statistically significant residual association was found when either SNP was conditioned by the other. We found no issues in the genotyping quality of either of these SNPs, both of which were included on the gene chip. Notably, rs149544 was associated with larger effect size (O/R 1.44, $P=1.93 \times 10^{-7}$) than rs26653 (O/R 1.35, $P=2.16 \times 10^{-5}$) in AxSpA association analysis.

Table 3.18: Residual statistical significance of *ERAP1* variants rs26653 and rs149544 after conditioning.

		Statistical significance of SNP associations (P)	
		rs26653	rs149544
Conditioned by SNP	Unconditioned	2.16×10^{-5}	1.55×10^{-7}
	rs26653	NA	3.88×10^{-6}
	rs149544	3.61×10^{-4}	NA

Residual statistical significance of tag SNPs (top row) after conditioning by SNPs listed in the left column. Blocks colour coded by relative statistical significance, with lighter colours indicating greater significance.

Association analysis performed by logistic regression in Plink V1.9.

We performed a similar analysis to investigate protective associations in the *ERAP1* coding region by conditioning for our tag variants that were inversely associated with AxSpA, namely rs26481, rs7063, rs1065407, rs17482078, rs10050860, and rs2287987. The results of this analysis are presented in Table 3.19. Interestingly, rs7063 showed the strongest association with AxSpA and accounted for the genetic association of all variants tested. It is worth noting that this SNP has previously been reported as a protective splice variant which may be relevant in AS (Hanson et al., 2018)

Table 3.19: Statistical significance of AxSpA disease protective *ERAP1* tag SNPs after conditioning.

		Statistical significance of SNP associations (P)					
		rs26481	rs7063	rs1065407	rs17482078	rs10050860	rs2287987
Conditioned by SNP	Unconditioned	6.71×10^{-7}	3.58×10^{-7}	2.44×10^{-6}	1.64×10^{-5}	2.08×10^{-5}	1.0×10^{-5}
	rs26481	NA	0.06225	0.082	0.121	0.020	0.071
	rs7063	0.185	NA	0.700	0.559	0.126	0.365
	rs1065407	0.019	0.036	NA	0.156	0.026	0.094
	rs17482078	0.004	0.006	0.021	NA	0.036	0.341
	rs10050860	0.008	0.018	0.042	0.345	NA	0.152
	rs2287987	0.004	0.007	0.023	0.967	0.041	NA

Residual statistical significance of tag SNPs (top row) after conditioning by SNPs listed in left column.

Blocks colour coded by relative statistical significance with lighter colours indicating greater significance.

Association analysis performed by logistic regression in Plink V1.9.

In this Irish population, the strongest genetic risk in the *ERAP1* coding region is located near rs149544. Despite strong LD between this SNP and a previously reported sequence variant, rs26653 ($r^2=0.94$, $D'=0.99$) in our patient cohort, we observed residual association of rs149544 after accounting for this SNP. This suggests there may be additional risk at this locus attributable to rs149544 itself, or another nearby variant. We found no evidence to support independent association of rs30187 in this population.

On the other hand, our analysis of protective associations within the ERAP1 coding region, which involved conditioning for several tag variants, showed that a single protective variant located at or near rs7063-T accounts for the strongest protective association (Table 3.19). Rs7063 remains a strong candidate for causation at this locus due to its known role as a splice altering variant (Hanson et al., 2018).

3.6.2: Phenotypic characteristics associated with rs149544 and rs7063

In Section 3.5.2, we reported that rs30187 was associated with earlier age at diagnosis and increased uveitis burden in AxSpA patients (Table 3.13). Building on this, we focused on two tag variants introduced above, rs149544 and rs7063, to explore their potential contribution to AxSpA disease severity. Tables 3.20 and 3.21 present our findings on the relationship between these SNPs and age of AxSpA onset and uveitis prevalence in HLA-B27 positive patients.

Although not statistically significant, carriers of the rs149544-C minor allele tended to have a younger age of first AxSpA symptom onset and diagnosis in a dosage-dependent manner (Table 3.16). However, rs149544-C was significantly associated with meeting the mNY criteria for AS ($P=0.023$), indicating its potential role in radiographic disease progression. Rs149544-C was also a significant risk factor for uveitis ($P=0.032$), with homozygotes for the minor allele having approximately 50% higher risk compared to non-carriers.

The minor allele of rs7063 was linked to disease characteristics in patients who were homozygous for the minor allele, while no differences were observed between heterozygotes and homozygotes for the major allele across all characteristics tested. Despite the large proportion of heterozygotes (216) in our B27+ cohort, the small number of minor allele homozygotes (12) was surprising. Further investigation showed that this SNP was in Hardy-Weinberg disequilibrium in B27+ patients ($P=0.0015$), but not in B27- patients ($P=0.94$), suggesting that this SNP is functionally active only in B27+ patients and follows a recessive pattern of association. Therefore, we compared homozygotes for the minor allele against all others combined. Homozygotes for the minor allele showed significantly increased age at diagnosis (45 vs 32 years, $P=0.003$) and markedly increased age at first symptom onset (32 vs 23 years), though this difference was not statistically significant. Uveitis was also much less common in homozygotes for the minor allele, despite their older age (8.3% vs 39.8%, $P=0.034$).

Both rs149544 and rs7063 are associated with AxSpA, but in opposite directionality. This was clearly observed in the likelihood of uveitis episodes where rs149544 was positively associated,

while rs7063 was protective. However, there were also differences in which disease parameters were affected by each of these SNPs. Specifically, rs7063 had a stronger influence on age at disease onset, while rs149544 was associated with radiographic spinal disease. Rs149544 shared directionality with and had a similar pattern of association with disease characteristics as rs30187 (Table 3.13). Finally, our analysis revealed that the protective effect of rs7063 was limited to homozygotes for the minor allele in HLA-B27 positive individuals.

Table 3.20: Univariate and multivariate analysis of patient characteristics based on rs149544 allelotypes in HLA-B27+ patients.

Characteristic	Genotype GG			Genotype GC			Genotype CC			Trend test (P)*
	Included	n	Median (IQR) or n (%)	Included	n	Median (IQR) or n (%)	Included	n	Median (IQR) or n (%)	
Male, n (%)	221		171 (77.4)	266		211 (79.3)	58		43 (74.1)	1
Age in years, median (IQR)	221		44 (37 - 53)	266		45 (37 - 55)	58		45 (37 - 54)	1
Age at first symptom onset, median (IQR)	221		24 (19 - 30)	265		23 (19 - 31)	57		21 (19 - 25)	0.468
Age at diagnosis, median (IQR)	221		34 (27 - 42)	265		32 (26 - 40)	57		32 (26 - 36)	0.579
Uveitis**, n (%)	216		74 (34.3)	262		107 (40.8)	57		28 (49.1)	0.032
mNY criteria, n (%)	221		156 (70.6)	266		215 (80.8)	58		47 (81.0)	0.023

*Ordinal logistic regression of minor allele count or Cochran-Armitage trend test

**At least one episode of uveitis in patient history.

Table 3.21: Univariate and multivariate analysis between rs7063 allelotypes in HLA-B27+ patients.

Characteristic	Genotype AA			Genotype AT			Genotype TT			Genotype TT vs AT and AA combined (P)*
	Included	n	Median (IQR) or n (%)	Included	n	Median (IQR) or n (%)	Included	n	Median (IQR) or n (%)	
Male, n (%)	317		253 (79.8)	216		162 (75.0)	12		10 (83.3)	1
Age in years, median (IQR)	317		44 (37 - 54)	216		45 (37 - 53)	12		49 (46 - 69)	0.036
Age at first symptom onset, median (IQR)	315		23 (19 - 31)	216		23 (19 - 30)	12		32 (20 - 37)	0.168
Age at diagnosis, median (IQR)	315		32 (26 - 40)	216		32 (26 - 41)	12		45 (40 - 47)	0.003
Uveitis**, n (%)	313		128 (40.9)	210		80 (38.1)	12		1 (8.3)	0.034
mNY criteria, n (%)	317		246 (77.6)	216		162 (75.0)	12		10 (83.3)	1

*Fisher's exact

**At least one episode of uveitis in patient history.

3.6.3: Investigation of the *ERAP2* gene as a possible independent AxSpA association

In our earlier haplotype analysis, we detected a minor association with AxSpA within Haplotype Block 11 ($P=0.019$, Supplementary Table S3.1). This region is important because it contains part of the coding region of *ERAP2* and includes the region where previously reported SNPs linking *ERAP2* to AS are located. However, there has been a debate on whether *ERAP2* is independently associated with AS since 2013 (Cortes et al., 2013b, Robinson et al., 2015b, Martín-Esteban et al., 2017, Tsui et al., 2010, Gao et al., 2022). To investigate this in our own data, we examined which variants within this large haplotype block, covering a 120KB (Chr5:96,229,225 – 96,350,270) region, were most strongly associated with AxSpA. Despite testing all variants within the Haplotype Block 11 region, we found no statistically significant SNPs after accounting for multiple testing (threshold $P<0.002$, see Methods Section 3.3.10). This included previously reported SNPs, such as rs2248374 and rs2549782 (Robinson et al., 2015b), which are shown in Table 3.22.

Table 3.22: SNP associations in the *ERAP2* coding region (Haplotype Block 11).

SNP	Base pair	Effect allele	Frequency affected	Frequency unaffected	P*	Odds ratio
rs2549782	96231000	T	0.46	0.47	0.3148	0.94
rs2248374	96235896	G	0.46	0.47	0.3389	0.94

*Threshold for statistical significance $P<0.002$.

3.7 Results: The relationship between polygenic risk and clinical phenotype

After quality control, we analysed a total of 641 genotyped patients, whose demographics were previously reported in Table 3.11. To explore the contribution of polygenic risk to disease-related demographics, we performed regression analysis of polygenic risk scores (PRS) against various demographic factors, using gender and genetic principal components as covariates. The results of this analysis are presented in Table 3.23. We found that both HLA and non-HLA PRS were inversely associated with age at first AxSpA symptom onset and age at diagnosis, with younger patients having higher PRS scores, suggesting an earlier onset of disease is associated with a higher genetic burden. The strongest correlation was observed outside of the HLA compartment, particularly at lower PRS thresholds. Notably, statistical significance in the HLA compartment was retained in all PRS thresholds for age at AxSpA symptom onset, with a similar pattern seen in age at diagnosis, although this did not reach statistical significance after correction for false discovery (Fig. 3.22). These findings suggest that a few highly significant loci outside of the HLA region are key drivers of disease onset. Additionally, while we have previously described the impact of HLA-B27 on age at onset (Table 3.12), our results here indicate that other loci within the HLA region also play a role.

Within the HLA region, PRS was strongly influenced by HLA-B27 status as expected due to its dominance in genetic risk within the MHC. Interestingly, we also observed positive PRS association outside of the HLA compartment, suggesting that HLA-B27 may interact with non-HLA variants in the disease state (Table 3.22, Figs. 3.23C – 3.23D). We hypothesized that *ERAP1*, which is known to have an epistatic interaction with HLA-B27, may be responsible for this finding. To test this hypothesis, we re-constructed PRS post-hoc, excluding both the HLA compartment and the *ERAP1* gene and surrounding region (Chr5: 95,000,000 – 97,000,000), and found that the association remained unchanged (Table 3.24). One possible explanation for these findings is fundamental differences between B27+ and B27- AxSpA, which is exposed by the PRS model which was trained on a population of mostly B27+ AS patients. However, we cannot rule out previously undocumented epistatic interactions of non-HLA loci with HLA-B27.

We found some indication that increased PRS within the non-HLA compartment was related to spinal disease progression required to meet the mNY criteria, although this was not statistically significant after accounting for false discovery and was not reflected in BASMI scores. We did not find evidence for an increased genetic burden in patients with a family history of AxSpA, despite a known association between family history and HLA-B27 in AxSpA (Arévalo et al., 2018).

Table 3.23: Polygenic risk associated to patient demographics.

	PRS		All variants		HLA region			Non-HLA Region		
	Threshold	R ²	P	P*	R ²	P	P*	R ²	P	P*
Age at AxSpA symptom onset	1x10 ⁻⁸	0.0149	0.002	0.023	0.0112	0.008	0.058	0.0155	0.001	0.020
	1x10 ⁻⁷	0.0160	0.001	0.020	0.0112	0.008	0.058	0.0164	0.001	0.020
	1x10 ⁻⁶	0.0156	0.001	0.020	0.0122	0.005	0.044	0.0124	0.005	0.043
	1x10 ⁻⁵	0.0152	0.002	0.021	0.0120	0.006	0.045	0.0103	0.010	0.066
	1x10 ⁻⁴	0.0159	0.001	0.020	0.0124	0.005	0.043	0.0107	0.009	0.062
	0.001	0.0173	0.001	0.016	0.0126	0.005	0.042	0.0124	0.005	0.043
	0.05	0.0145	0.002	0.026	0.0130	0.004	0.039	0.0047	0.085	0.177
	0.1	0.0145	0.002	0.026	0.0128	0.004	0.040	0.0051	0.073	0.165
	0.5	0.0139	0.003	0.030	0.0127	0.004	0.040	0.0044	0.095	0.187
Age at AxSpA diagnosis	1x10 ⁻⁸	0.0157	0.001	0.020	0.0094	0.014	0.078	0.0154	0.001	0.020
	1x10 ⁻⁷	0.0158	0.001	0.020	0.0094	0.014	0.078	0.0133	0.003	0.035
	1x10 ⁻⁶	0.0154	0.001	0.020	0.0094	0.014	0.078	0.0110	0.008	0.058
	1x10 ⁻⁵	0.0156	0.001	0.020	0.0096	0.013	0.075	0.0107	0.008	0.061
	1x10 ⁻⁴	0.0153	0.001	0.020	0.0098	0.013	0.074	0.0086	0.019	0.092
	0.001	0.0158	0.001	0.020	0.0103	0.010	0.066	0.0077	0.026	0.102
	0.05	0.0105	0.010	0.065	0.0108	0.008	0.061	0.0008	0.468	0.578
	0.1	0.0097	0.013	0.074	0.0107	0.008	0.061	0.0007	0.512	0.627
	0.5	0.0082	0.022	0.098	0.0107	0.008	0.061	0.0002	0.700	0.760
Family history of AxSpA	1x10 ⁻⁸	0.0036	0.126	0.227	0.0024	0.211	0.331	0.0023	0.223	0.333
	1x10 ⁻⁷	0.0035	0.132	0.236	0.0024	0.211	0.331	0.0016	0.312	0.423
	1x10 ⁻⁶	0.0034	0.135	0.241	0.0024	0.211	0.331	0.0016	0.306	0.415
	1x10 ⁻⁵	0.0034	0.138	0.245	0.0022	0.228	0.338	0.0018	0.274	0.386
	1x10 ⁻⁴	0.0038	0.114	0.216	0.0021	0.241	0.352	0.0032	0.148	0.256
	0.001	0.0040	0.105	0.203	0.0024	0.211	0.331	0.0029	0.171	0.288
	0.05	0.0071	0.031	0.113	0.0023	0.223	0.333	0.0074	0.028	0.107
	0.1	0.0077	0.025	0.100	0.0023	0.223	0.333	0.0083	0.019	0.092
	0.5	0.0059	0.049	0.141	0.0023	0.223	0.333	0.0044	0.090	0.183
HLA-B27 status	1x10 ⁻⁸	0.1583	2.9x10⁻²⁶	6.5x10⁻²⁴	0.1397	1.2x10⁻²⁴	3.7x10⁻²³	0.0237	9.2x10⁻⁵	0.002
	1x10 ⁻⁷	0.1566	5.6x10⁻²⁶	6.5x10⁻²⁴	0.1397	1.2x10⁻²⁴	3.7x10⁻²³	0.0197	3.6x10⁻⁴	0.007
	1x10 ⁻⁶	0.1576	3.8x10⁻²⁶	6.5x10⁻²⁴	0.1413	6.2x10⁻²⁵	2.5x10⁻²³	0.0205	2.7x10⁻⁴	0.006
	1x10 ⁻⁵	0.1564	6.0x10⁻²⁶	6.5x10⁻²⁴	0.1393	1.4x10⁻²⁴	4.1x10⁻²³	0.0191	4.4x10⁻⁴	0.008
	1x10 ⁻⁴	0.1555	8.7x10⁻²⁶	7.5x10⁻²⁴	0.1399	1.1x10⁻²⁴	3.7x10⁻²³	0.0197	3.6x10⁻⁴	0.007
	0.001	0.1517	3.7x10⁻²⁵	1.8x10⁻²³	0.1422	4.4x10⁻²⁵	1.9x10⁻²³	0.0132	0.004	0.036
	0.05	0.1237	1.7x10⁻²⁰	4.3x10⁻¹⁹	0.1447	1.6x10⁻²⁵	9.0x10⁻²⁴	0.0044	0.092	0.185
	0.1	0.1242	1.4x10⁻²⁰	3.8x10⁻¹⁹	0.1445	1.7x10⁻²⁵	9.0x10⁻²⁴	0.0067	0.039	0.128
	0.5	0.1202	6.3x10⁻²⁰	1.5x10⁻¹⁸	0.1446	1.6x10⁻²⁵	9.0x10⁻²⁴	0.0073	0.031	0.112
Modified New York criteria	1x10 ⁻⁸	<0.0001	0.863	0.902	0.0013	0.358	0.460	0.0073	0.031	0.112
	1x10 ⁻⁷	0.0001	0.837	0.881	0.0013	0.358	0.460	0.0074	0.029	0.110
	1x10 ⁻⁶	0.0001	0.821	0.871	0.0015	0.332	0.435	0.0080	0.025	0.100
	1x10 ⁻⁵	0.0002	0.692	0.753	0.0015	0.323	0.430	0.0120	0.005	0.045
	1x10 ⁻⁴	0.0002	0.715	0.768	0.0015	0.335	0.435	0.0087	0.018	0.089
	0.001	0.0003	0.682	0.744	0.0015	0.332	0.435	0.0079	0.025	0.100
	0.05	0.0002	0.707	0.766	0.0015	0.332	0.435	0.0030	0.167	0.285
	0.1	0.0002	0.742	0.796	0.0015	0.332	0.435	0.0024	0.211	0.331
	0.5	0.0006	0.545	0.661	0.0015	0.335	0.435	0.0037	0.126	0.227

Sex and genetic principal components used as covariates

*P values corrected for multiple testing by Benjamini-Hochburg procedure (FDR=0.05).

Positive associations in red text. Inverse associations in blue text.

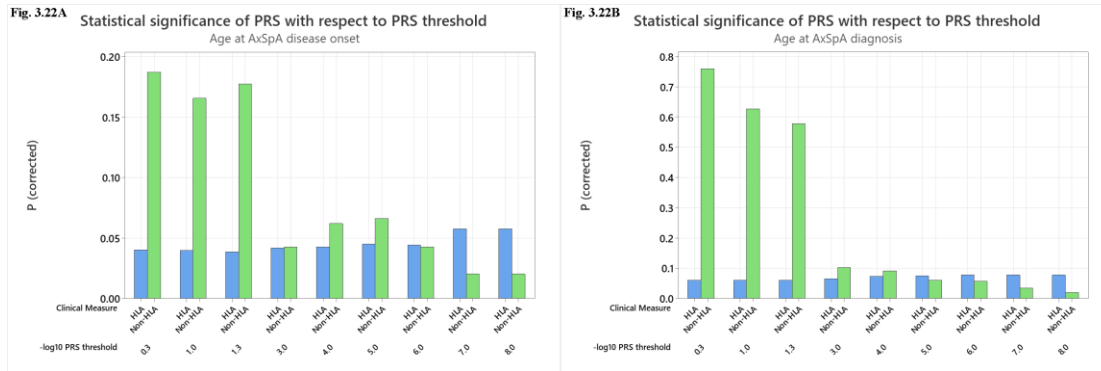


Fig. 3.22: Statistical significance of PRS in relation to age at first symptom onset and diagnosis at various PRS thresholds.

Fig. 3.22A: Statistical significance of PRS in age at first AxSpA symptom onset for each PRS threshold.

Fig. 3.22B: Statistical significance of PRS in age at diagnosis for each PRS threshold.

Table 3.24: Polygenic risk excluding the HLA region and ERAP1 associated with HLA-B27.

Non-HLA Region		
PRS Threshold	R ²	P
1x10 ⁻⁸	0.0273	0.004
1x10 ⁻⁷	0.0231	0.004
1x10 ⁻⁶	0.0230	0.004
1x10 ⁻⁵	0.0227	0.004
1x10 ⁻⁴	0.0237	0.004
0.001	0.0186	0.014
0.05	0.0113	0.154
0.1	0.0134	0.074
0.5	0.0141	0.064

Sex and genetic principal components used as covariates. P values are corrected for multiple testing by Benjamini-Hochburg procedure (FDR=0.05). Positive associations in red text. Inverse associations in blue text.

Fig. 3.23A Age at AxSpA first symptom onset by PRS (HLA region)

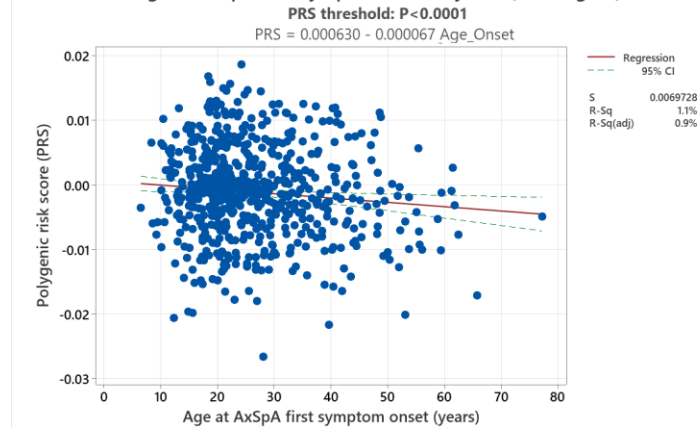


Fig. 3.23B Age at AxSpA first symptom onset by PRS (non-HLA region)

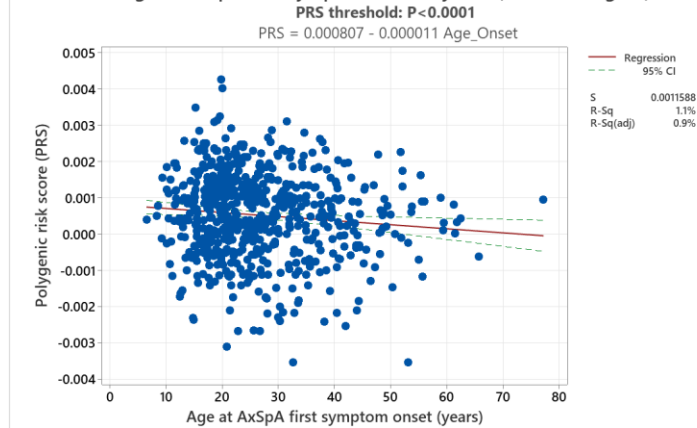


Fig. 3.23C Binary distribution of HLA-B27 status by PRS (HLA region)

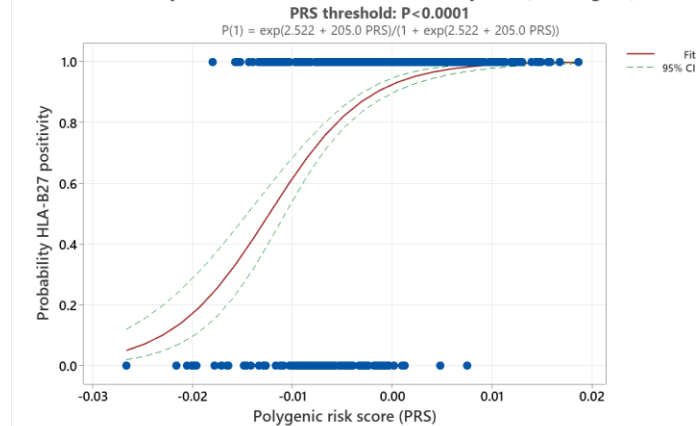


Fig. 3.23D Binary distribution of HLA-B27 status by PRS (non-HLA)

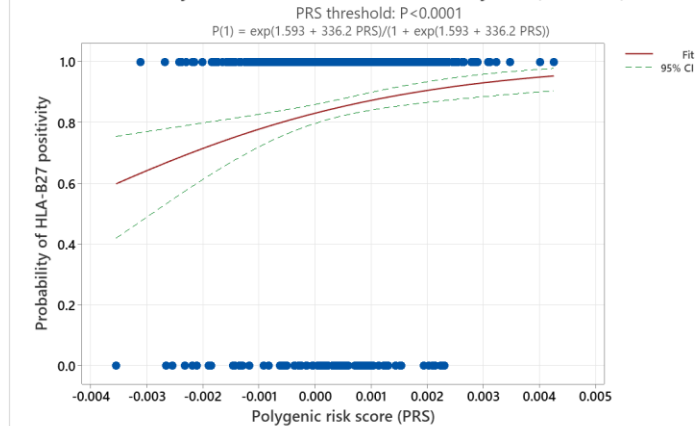


Fig. 3.23A: Relationship between PRS derived from the HLA region and age at AxSpA symptom onset.

Fig. 3.23B: Relationship between PRS derived from outside of the HLA region and age at AxSpA symptom onset.

Fig. 3.23C: Binary distribution of HLA-B27 status based on PRS score within the HLA region.

Fig. 3.23D: Binary distribution of HLA-B27 status based on PRS score outside of the HLA region.

Using a similar methodology, we examined whether patient reported outcomes and the risk of unemployment were associated with PRS. For each of the PROs (BASDAI, BASFI, BASMI, ASQoL, and HAQs) and current employment status, we conducted multivariate regression analysis separately, including gender, AxSpA disease duration, and genetic principal components in each analysis (Table 3.25). We found weak inverse correlations between PRS in the HLA region and all PROs, but these were not significant after adjusting for false discovery. The directionality of this possible association was consistent with our earlier findings regarding HLA-B27, although those associations were also not significant (Table 3.12). Curiously, we observed a weak positive association between SNPs of high statistical significance in the non-HLA region and gainful employment, but this was also not significant after adjusting for false discovery.

Table 3.25: Polygenic risk associated with disease activity, disability and gainful employment.

	PRS	All SNP variants			HLA region only			Non-HLA Region		
	Threshold	R ²	P	P*	R ²	P	P*	R ²	P	P*
BASDAI	1x10 ⁻⁸	0.0021	0.247	0.357	0.0075	0.028	0.107	0.0016	0.317	0.427
	1x10 ⁻⁷	0.0024	0.211	0.331	0.0075	0.028	0.107	0.0025	0.211	0.331
	1x10 ⁻⁶	0.0022	0.234	0.344	0.0079	0.025	0.100	0.0013	0.358	0.460
	1x10 ⁻⁵	0.0023	0.223	0.333	0.0078	0.026	0.102	0.0015	0.325	0.431
	1x10 ⁻⁴	0.0025	0.211	0.331	0.0080	0.025	0.100	0.0016	0.304	0.415
	0.001	0.0028	0.176	0.293	0.0079	0.025	0.100	0.0024	0.211	0.331
	0.05	0.0038	0.116	0.217	0.0085	0.019	0.092	0.0029	0.173	0.289
	0.1	0.0042	0.098	0.191	0.0084	0.020	0.095	0.0036	0.129	0.232
	0.5	0.0050	0.073	0.165	0.0085	0.019	0.092	0.0045	0.090	0.183
BASFI	1x10 ⁻⁸	0.0056	0.055	0.148	0.0080	0.023	0.100	0.0058	0.051	0.144
	1x10 ⁻⁷	0.0060	0.047	0.139	0.0080	0.023	0.100	0.0062	0.044	0.136
	1x10 ⁻⁶	0.0054	0.055	0.148	0.0078	0.025	0.100	0.0037	0.120	0.222
	1x10 ⁻⁵	0.0059	0.050	0.141	0.0081	0.022	0.099	0.0047	0.079	0.172
	1x10 ⁻⁴	0.0063	0.043	0.134	0.0085	0.019	0.092	0.0051	0.067	0.158
	0.001	0.0070	0.032	0.113	0.0087	0.018	0.089	0.0064	0.040	0.130
	0.05	0.0047	0.080	0.172	0.0092	0.015	0.078	0.0008	0.462	0.572
	0.1	0.0055	0.055	0.148	0.0090	0.015	0.081	0.0017	0.296	0.410
	0.5	0.0054	0.055	0.148	0.0092	0.015	0.078	0.0017	0.296	0.410
BASMI	1x10 ⁻⁸	0.0011	0.370	0.473	0.0016	0.272	0.385	0.0003	0.618	0.724
	1x10 ⁻⁷	0.0011	0.371	0.473	0.0016	0.272	0.385	0.0003	0.660	0.729
	1x10 ⁻⁶	0.0008	0.445	0.554	0.0015	0.296	0.410	<0.0001	0.939	0.959
	1x10 ⁻⁵	0.0008	0.439	0.549	0.0016	0.270	0.384	<0.0001	0.906	0.931
	1x10 ⁻⁴	0.0007	0.470	0.578	0.0017	0.266	0.382	<0.0001	0.989	0.990
	0.001	0.0008	0.448	0.556	0.0017	0.265	0.381	<0.0001	0.896	0.929
	0.05	0.0001	0.837	0.881	0.0013	0.319	0.427	0.0040	0.086	0.178
	0.1	<0.0001	0.930	0.952	0.0013	0.326	0.432	0.0026	0.166	0.284
	0.5	<0.0001	0.950	0.968	0.0013	0.323	0.430	0.0017	0.261	0.377
ASQoL	1x10 ⁻⁸	0.0049	0.247	0.165	0.0055	0.055	0.148	0.0016	0.302	0.413
	1x10 ⁻⁷	0.0052	0.247	0.154	0.0055	0.055	0.148	0.0021	0.247	0.357
	1x10 ⁻⁶	0.0048	0.247	0.167	0.0054	0.061	0.150	0.0013	0.348	0.450
	1x10 ⁻⁵	0.0054	0.247	0.148	0.0055	0.055	0.148	0.0026	0.188	0.310
	1x10 ⁻⁴	0.0054	0.247	0.148	0.0056	0.055	0.148	0.0025	0.202	0.328
	0.001	0.0065	0.247	0.129	0.0058	0.052	0.146	0.0042	0.095	0.187
	0.05	0.0059	0.247	0.141	0.0062	0.045	0.138	0.0018	0.279	0.391
	0.1	0.0057	0.247	0.146	0.0060	0.049	0.141	0.0017	0.296	0.410
	0.5	0.0063	0.247	0.133	0.0061	0.047	0.139	0.0024	0.211	0.331
HAQ-s	1x10 ⁻⁸	0.0095	0.013	0.074	0.0056	0.055	0.148	0.0053	0.062	0.152
	1x10 ⁻⁷	0.0098	0.011	0.071	0.0056	0.055	0.148	0.0048	0.076	0.169
	1x10 ⁻⁶	0.0092	0.014	0.078	0.0056	0.055	0.148	0.0033	0.145	0.253
	1x10 ⁻⁵	0.0098	0.011	0.071	0.0057	0.055	0.148	0.0046	0.085	0.177
	1x10 ⁻⁴	0.0096	0.013	0.074	0.0060	0.049	0.141	0.0036	0.126	0.227
	0.001	0.0109	0.008	0.058	0.0060	0.049	0.141	0.0056	0.055	0.148
	0.05	0.0088	0.016	0.085	0.0062	0.044	0.136	0.0017	0.296	0.410
	0.1	0.0079	0.023	0.100	0.0061	0.047	0.139	0.0011	0.396	0.499
	0.5	0.0097	0.012	0.073	0.0063	0.044	0.136	0.0027	0.188	0.310
Gainful employment	1x10 ⁻⁸	0.0087	0.031	0.112	0.0004	0.618	0.724	0.0132	0.008	0.058
	1x10 ⁻⁷	0.0084	0.034	0.117	0.0004	0.618	0.724	0.0123	0.010	0.066
	1x10 ⁻⁶	0.0078	0.040	0.130	0.0004	0.618	0.724	0.0098	0.021	0.098
	1x10 ⁻⁵	0.0074	0.045	0.138	0.0004	0.618	0.724	0.0082	0.035	0.121
	1x10 ⁻⁴	0.0071	0.050	0.142	0.0004	0.618	0.724	0.0064	0.064	0.155
	0.001	0.0087	0.031	0.112	0.0003	0.679	0.743	0.0102	0.019	0.092
	0.05	0.0082	0.036	0.121	0.0004	0.660	0.729	0.0048	0.107	0.204
	0.1	0.0056	0.082	0.175	0.0003	0.669	0.734	0.0016	0.348	0.450
	0.5	0.0050	0.100	0.194	0.0003	0.664	0.730	0.0013	0.400	0.503

Sex, AxSpA disease duration and genetic principal components used as covariates.

*P values corrected for multiple testing by Benjamini-Hochburg procedure (FDR=0.05).

Positive associations in red text. Inverse associations in blue text.

In our analysis of EAMs, we included both gender and AxSpA disease duration as covariates (Table 3.26). While we observed a weak inverse association with IBD outside of the HLA region, it was not statistically significant after accounting for false discovery. Interestingly, we found no association between PRS in the HLA region and AAU across all PRS thresholds, which was surprising, given our previous report of a strong association between HLA-B27 and AAU in AxSpA patients (Table 3.12). HLA-B27 itself, meanwhile was highly predicted by PRS in the HLA region, as discussed above. To examine this further, we generated additional PRS for the HLA region using only SNPs at statistical thresholds below $P < 1 \times 10^{-60}$, $P < 1 \times 10^{-120}$ and $P < 1 \times 10^{-180}$ in the base GWAS. We again found no relationship between PRS and AAU at each of these additional thresholds (data not shown). To explain these findings, we examined the allele frequencies in ASRI of all SNPs with statistical significance of $P < 1 \times 10^{-180}$ in the base GWAS study, comparing by HLA-B27 status and comorbid AAU (Table 3.27). We found that all SNPs captured within this threshold were predictive of HLA-B27 positivity in the ASRI cohort. However, SNPs within the HLA-B gene region showed weak capacity to predict AAU and the directionality of SNPs rs2524094 and rs2596501 with respect to AAU was inconsistent. However, although these two SNPs are both strongly associated with HLA-B27+ AxSpA, they are not in LD with each other ($r^2=0.068$, $D'=0.338$). Rs2524094 is an upstream variant affecting the HLA-C gene, for which it is an eQTL (Carithers and Moore, 2015). Meanwhile, SNPs near to other genes showed no association. This explained the lack of association between PRS and AAU in the AxSpA cohort. Interestingly, we observed an association between previously described protective SNPs in the CCHCR1 gene (Cortes et al., 2013b) and HLA-B27 negative AxSpA. These SNPs are not in significant LD with the HLA-B locus (for example LD between SNPs rs746647 and rs2524094, $r^2=0.85$, $D'=0.34$). A similar pattern less pronounced was observed with SNPs belonging to the BTNL2 region. This tentatively suggests that the genetic risk profile differs between HLA-B27 positive and HLA-B27 negative disease amongst SNPs of high effect size.

Table 3.26: Polygenic risk associated to extra-axial manifestations

	PRS	All variants			HLA region			Non-HLA Region		
	Threshold	R ²	P	P*	R ²	P	P*	R ²	P	P*
Acute anterior uveitis	1x10 ⁻⁸	0.0021	0.244	0.356	<0.0001	0.906	0.931	0.0040	0.105	0.202
	1x10 ⁻⁷	0.0022	0.230	0.339	<0.0001	0.906	0.931	0.0043	0.095	0.187
	1x10 ⁻⁶	0.0023	0.223	0.333	<0.0001	0.975	0.982	0.0048	0.077	0.171
	1x10 ⁻⁵	0.0023	0.223	0.333	<0.0001	0.975	0.982	0.0043	0.093	0.187
	1x10 ⁻⁴	0.0024	0.211	0.331	<0.0001	0.984	0.989	0.0047	0.080	0.172
	0.001	0.0026	0.192	0.313	<0.0001	0.989	0.990	0.0050	0.073	0.165
	0.05	0.0030	0.159	0.274	<0.0001	0.957	0.968	0.0038	0.116	0.217
	0.1	0.0036	0.126	0.227	<0.0001	0.957	0.968	0.0046	0.083	0.175
	0.5	0.0009	0.438	0.548	<0.0001	0.952	0.968	0.0001	0.769	0.818
Psoriasis	1x10 ⁻⁸	0.0042	0.103	0.200	0.0049	0.080	0.172	0.0007	0.498	0.612
	1x10 ⁻⁷	0.0046	0.088	0.182	0.0049	0.080	0.172	0.0013	0.374	0.475
	1x10 ⁻⁶	0.0051	0.073	0.165	0.0038	0.126	0.227	0.0034	0.145	0.253
	1x10 ⁻⁵	0.0049	0.080	0.172	0.0040	0.114	0.216	0.0026	0.211	0.331
	1x10 ⁻⁴	0.0049	0.079	0.172	0.0041	0.111	0.211	0.0024	0.223	0.333
	0.001	0.0045	0.091	0.185	0.0039	0.117	0.218	0.0016	0.317	0.427
	0.05	0.0031	0.160	0.275	0.0045	0.092	0.185	0.0001	0.825	0.873
	0.1	0.0030	0.168	0.286	0.0045	0.095	0.187	0.0001	0.829	0.875
	0.5	0.0027	0.196	0.320	0.0046	0.092	0.185	0.0001	0.853	0.894
Inflammatory bowel disease	1x10 ⁻⁸	0.0068	0.035	0.121	0.0052	0.067	0.158	0.0087	0.018	0.089
	1x10 ⁻⁷	0.0066	0.039	0.128	0.0052	0.067	0.158	0.0064	0.041	0.133
	1x10 ⁻⁶	0.0070	0.032	0.113	0.0049	0.077	0.171	0.0077	0.025	0.101
	1x10 ⁻⁵	0.0071	0.031	0.113	0.0057	0.055	0.148	0.0070	0.032	0.113
	1x10 ⁻⁴	0.0067	0.036	0.122	0.0053	0.064	0.155	0.0063	0.042	0.134
	0.001	0.0055	0.055	0.148	0.0055	0.055	0.148	0.0024	0.211	0.331
	0.05	0.0057	0.055	0.148	0.0051	0.073	0.165	0.0026	0.192	0.313
	0.1	0.0073	0.029	0.112	0.0049	0.075	0.168	0.0051	0.068	0.159
	0.5	0.0059	0.049	0.141	0.0050	0.073	0.165	0.0028	0.174	0.291

Sex, AxSpA disease duration and genetic principal components used as covariates.

*P values corrected for multiple testing by Benjamini-Hochburg procedure (FDR=0.05).

Positive associations in red text. Inverse associations in blue text.

Table 3.27: Frequency (%) of high impact SNPs in the HLA region stratified by HLA-B27 status and history of uveitis.

SNP ID	IGAS GWAS association (P)	B27+	B27-	B27+ vs B27-	B27+ AAU	B27+ no AAU	B27+ AAU vs no AAU	B27- AAU	B27- no AAU	B27- AAU vs no AAU
		n= 545	n= 97	P*	n= 209	n= 326	P*	n= 16	n= 80	P*
rs12192673	4.51x10 ⁻²⁰¹	14.04	3.09	<0.001	14.59	13.80	0.72	0.00	3.75	0.592
rs746647	1.77x10 ⁻²⁰¹	16.70	41.75	<0.001	16.51	16.56	0.44	31.25	43.75	0.240
rs1265114	7.42x10 ⁻²⁰²	16.70	41.75	<0.001	16.51	16.56	0.44	31.25	43.75	0.240
rs1265112	1.79x10 ⁻²⁰²	16.70	41.75	<0.001	16.51	16.56	0.44	31.25	43.75	0.240
rs2517985	7.5x10 ⁻²⁰¹	16.06	41.24	<0.001	16.27	15.64	0.44	31.25	43.13	0.242
rs130075	<1x10 ⁻²⁹⁹	84.31	98.97	<0.001	85.41	83.90	0.55	100.00	98.75	1.000
rs2524094	<1x10 ⁻²⁹⁹	76.42	49.48	<0.001	72.01	79.14	0.01	59.38	46.88	0.246
rs2596501	7.2x10 ⁻²⁸⁹	75.32	50.52	<0.001	78.71	73.01	0.04	40.63	52.50	0.249
rs2261033	8.8x10 ⁻²⁸⁷	64.50	39.18	<0.001	62.68	65.95	0.29	37.50	39.38	1.000
rs4248166	6.5x10 ⁻²¹⁰	68.62	78.87	0.004	69.86	67.64	0.46	81.25	78.13	0.816
rs2294884	8.8x10 ⁻²⁰⁸	68.44	78.87	0.004	69.86	67.48	0.42	81.25	78.13	0.816
rs2294882	2.5x10 ⁻¹⁹⁰	68.17	78.87	0.003	69.38	67.33	0.50	81.25	78.13	0.816
rs2294881	2.4x10 ⁻¹⁹⁰	68.17	78.87	0.003	69.38	67.33	0.50	81.25	78.13	0.816

Genes highlighted by gene region: In blue, protective SNPs near to the *CCHCR1* gene. In yellow: SNPs within the HLA-B locus. In green: protective SNPs near *BTNL2*.

*Fisher's exact.

Finally, due to the observed differences in clinical presentation between men and women with AxSpA, we performed a post-hoc analysis of polygenic risk between the genders, as presented in Table 3.28. We observed no difference in composite genetic risk after accounting for multiple testing.

Table 3.28: A post-hoc analysis of polygenic risk associated with gender in AxSpA patients.

Patient sex	PRS	All variants			HLA region			Non-HLA Region		
	Threshold	R ²	P	P*	R ²	P	P*	R ²	P	P*
	1x10 ⁻⁸	0.0007	0.499	0.499	0.0019	0.274	0.352	0.0039	0.117	0.306
	1x10 ⁻⁷	0.0008	0.483	0.499	0.0019	0.274	0.352	0.0037	0.125	0.306
	1x10 ⁻⁶	0.0009	0.451	0.487	0.0024	0.214	0.306	0.0036	0.131	0.306
	1x10 ⁻⁵	0.0010	0.432	0.486	0.0024	0.215	0.306	0.0042	0.101	0.306
	1x10 ⁻⁴	0.0010	0.426	0.486	0.0026	0.201	0.306	0.0039	0.117	0.306
	0.001	0.0012	0.388	0.476	0.0026	0.194	0.306	0.0042	0.102	0.306
	0.05	0.0027	0.188	0.306	0.0024	0.211	0.306	0.0068	0.037	0.306
	0.1	0.0029	0.173	0.306	0.0025	0.207	0.306	0.0067	0.039	0.306
	0.5	0.0028	0.184	0.306	0.0024	0.211	0.306	0.0053	0.066	0.306

AxSpA disease duration and genetic principal components used as covariates.

*P values corrected for multiple testing by Benjamini-Hochburg procedure (FDR=0.05).

Statistically significant results indicated in bold.

3.8 Discussion

Throughout this chapter, we have provided compelling evidence that genetic risk factors not only contribute to the development of AxSpA, but also impact patient outcomes. Both the HLA-B27 phenotype and *ERAP1* gene have been shown to play a significant role in disease onset and the likelihood of uveitis episodes. The importance of the *ERAP1* gene is further highlighted by fine mapping of the gene, which identified two additional SNPs with strong AxSpA association. In both cases, we found that genetic variation at these loci also impact on AxSpA disease severity. We have further demonstrated that aggregated genetic risk scores such as PRS also offer predictive value in the age of AxSpA disease onset. Moreover, PRS remain predictive even where the HLA region and *ERAP1* are excluded from the PRS model. This builds upon previously published research which has shown that PRS are more effective than the use of HLA-B27 alone in predicting the likelihood of AS (Li et al., 2021).

3.8.1 Genome-wide association analysis of clinical phenotypes

In our GWAS analysis, we did not identify any variants associated with the AS sub-phenotypes we investigated. However, our QQ-plots revealed distinct genetic risk associated with early disease onset (Fig. 3.14). Due to the small size of our AxSpA cohort, we could not detect SNPs with low effect sizes.

A notable finding was the reduced burden of HLA-Cw6 in AxSpA patients with concurrent psoriasis, as compared to studies based solely on psoriasis. This observation was evident in two separate approaches used to determine HLA-C type in the ASRI patient cohort. Although it is tempting to attribute this to the high burden of HLA-B27 in AxSpA patients, which may in itself confer psoriasis risk, we found that B27 was also less prevalent in AxSpA patients with psoriasis. Previous reports support this inverse association between HLA-B27 and psoriasis in AxSpA patients (Feldtkeller et al., 2003, Arévalo et al., 2018). The interplay between HLA-B27 and Cw6 in comorbid disease could be complex, as they are associated with different psoriatic presentations. HLA-Cw6 is linked with classical cutaneous psoriasis, while HLA-B27 is more strongly linked with psoriatic arthritis (PsA), which may also present with skin manifestations (Elkayam et al., 2000). Notably, HLA-B27 and HLA-Cw6 do not occur on a common haplotype, reducing their likelihood of concurrence. Despite low prevalence of Cw6 in AxSpA patients with comorbid psoriasis over an Irish healthy control group, Cw6 positive AxSpA patients are more likely to develop psoriasis, while even those without Cw6 have an increased likelihood of psoriasis over the general population (Arévalo et al., 2018). Thus, we suspect that there is genetic risk compensation by HLA-B27 in the AxSpA group with concurrent psoriasis, but due to the

persisting lack of clear clinical definition between PsA and AS, this will prove challenging to unravel (Kavanaugh et al., 2023).

Interestingly, the tag SNP rs73057527 on chromosome 3 which was identified as possibly associated with early disease onset was located near to a region which was subsequently implicated in AxSpA disease causation in our meta-analysis. This region, situated between the *CMC1* and *EOMES* genes contains several risk variants for autoinflammatory disease. However, we were unable to find supporting evidence, either through SNPs or eQTLs, of a link between this specific variant and AxSpA disease. While the evidence supporting this finding in this initial exploration is strictly limited, we are unable to out-rule the possibility of a novel association within this region.

3.8.2 The relationship between HLA-B27, ERAP1 and disease presentation

Previous studies have produced inconsistent results on the relationship between HLA-B27 and EAMs in ankylosing spondylitis (AS) cases. While Feldtkeller et al. reported an inverse correlation between HLA-B27 and uveitis, Huang et al. found a positive association, and Arévalo et al. identified no association (Arévalo et al., 2018, Huang et al., 2020, Feldtkeller et al., 2003). In our study, we found that HLA-B27 was associated with a higher likelihood of AAU in AxSpA patients. While our data was limited by lacking the date of AAU onset in our cohort, we found no difference in the ages of patients with reported AAU based on HLA-B27 carriage. This tentatively suggests that HLA-B27 is causative of AAU, but does not result in earlier onset in patients overall.

Consistent with previous studies, our analysis revealed that HLA-B27 is linked with earlier disease onset and a family history of AxSpA. However, we found no evidence of any effect of HLA-B27 on patient-reported outcomes for disease activity, physical disability, or quality of life. Additionally, we did not observe an association between HLA-B27 and radiographic spinal disease, a critical criterion for meeting the modified New York criteria for AS.

ERAP1 encodes an aminopeptidase which is known to functionally interact with HLA-B27. This epistasis was evidenced by the enrichment of the *ERAP1* risk allele, rs30187-T in HLA-B27 positive patients, when compared to their B27 negative counterparts. We found that *ERAP1* rs30187-T dosage compounded the effects of HLA-B27 in age at onset and frequency of AAU in AxSpA patients (Table 3.11). However, rs30187-T was not a useful biomarker in predicting AxSpA disease activity or axial radiographic change.

Exploration of the *ERAP1* gene region through haplotype analysis was challenging due to high D' observed across the region, which extended through several haplotype blocks. We computationally generated haplotypes across the region and scrutinized their disease association. The *ERAP1* coding variant, rs30187, was found to be associated with AxSpA within some haplotypes, but others containing the disease-associated variant showed no association.

Further investigation using generated tag variants revealed two major haplotypes associated with AxSpA in the Irish study cohort. One haplotype, containing rs149544-G, was disease-causing, while the other, rs7063-T, was protective. Interestingly, rs149544-G accounted for genetic risk at rs30187-T, which had been widely reported as a possible causative sequence variant. We found no evidence for residual genetic risk at rs30187 after rs149544 had been accounted for. While this may be a peculiarity specific to the Irish AxSpA cohort, there is general consensus that disease causing variants tend to remain consistent between populations, unlike tag SNPs (Lewis et al., 2021). Our findings are further supported by weak or absent rs30187-T association reported within Asian AxSpA cohorts (Chen and Zhang, 2015, Su et al., 2018). Collectively, these findings suggest that rs30187-T is unlikely to be a disease-causing SNP in AxSpA.

We noted that rs149544 is in strong LD with a previously reported sequence variant, rs26653 (Maksymowych et al., 2009a). Despite this strong LD, we found residual association of rs149544 after rs26653 had been accounted for. It is possible that additional disease relevant SNPs share haplotypes with these variants. This will require further investigation within a larger study cohort to fully elucidate.

We found that the protective SNP, rs7063-T, which has previously been reported as a splice altering variant (Hanson et al., 2018) and is also an eQTL strongly associated with *ERAP1* expression, is likely to represent the strongest protective association in the *ERAP1* gene. Importantly, rs7063 fully accounted for the association of previously reported protective SNPs, rs10050860 and rs17482078.

Although we found that rs149544-G is linked to both uveitis and spinal disease progression, we were unable to verify a relationship with patient age at disease onset. This was unexpected, given that rs30187-T is associated with earlier disease in patients and, as mentioned, rs149544 fully accounts for genetic risk at rs30187. However, we did note that rs7063-T, in its homozygous form, is markedly associated with late onset in patients. Interplay between rs7063 and rs30187 through shared haplotypes has been previously reported (Hanson et al., 2018), and it is possible that the observed relationship between rs30187 and age at AxSpA disease onset could be the result of haplotype cross-over with rs7063 and rs149544, given that rs30187 is located between these two SNPs.

In contrast to some earlier studies, we did not find any evidence of an association between AxSpA and the *ERAP2* gene region, although we examined SNPs that had been previously reported in this region (Robinson et al., 2015). However, we cannot rule out the possibility of associations of small effect size.

Our study underscores the significance of the *ERAP1* gene in both the aetiology and clinical manifestation of AxSpA. Although we focused on the most impactful variants in our investigation, it is likely that additional disease-associated variants with smaller effect sizes contribute to the complexity of *ERAP1*'s influence on disease onset and progression. Identifying causative SNPs within this region is challenging due to the interplay of associated SNPs on shared haplotypes. Nevertheless, given the crucial role of the *ERAP1* gene in AxSpA pathology, further exploration of genetic variation in this region is imperative.

3.8.4 The contribution of polygenic risk to disease phenotype and severity

To our knowledge, we are the first to examine the link between polygenic risk scores (PRS) and disease characteristics in AxSpA, which was made possible by the extensive phenotype information available on these patients. HLA-B27 was a highly determinant factor in the PRS within the MHC/HLA region, which was expected due to the high weighting of HLA-B27-associated SNPs. However, we also found that HLA-B27 was linked to increased polygenic risk outside of the HLA region, contrary to our expectation that compensatory genetic risk would present in B27-negative individuals. This may be explained by bias within our PRS model toward HLA-B27 positive disease, which comprised around 90% of the GWAS study upon which our PRS were based (Cortes et al., 2013b). This resulted in increased weighting of SNPs that cooperate with HLA-B27 in the PRS model, while the relative effect size of SNPs uniquely associated with HLA-B27 negative disease was, conversely diminished. PRS models, when trained on these GWAS will, therefore, under-perform in HLA-B27 negative AxSpA disease prediction. Nevertheless, our findings show that genetic risk outside of the HLA region in AxSpA is not identical between HLA-B27 positive and HLA-B27 negative patients. Furthermore, we found that high impact variants in the *CCHCR1* gene within the HLA region were also predictive of HLA-B27 status, despite not being in LD with the HLA-B gene. Our study, like others published to date, is insufficient in size to extensively explore the genetic dichotomies between B27+ and B27- disease.

Polygenic risk was robustly linked to earlier disease onset and diagnosis within both HLA and non-HLA compartments, indicating patients with earlier disease have higher genetic burdens than those who develop disease later. Moreover, our study shows that PRS can predict earlier disease onset even without including the HLA region and *ERAP1* gene. However, PRS did not prove

valuable in predicting whether a patient was likely to meet the modified New York criteria for AS and was a weak predictor of a family history of AxSpA. We tentatively found evidence that PRS based on the HLA region could predict disease severity, as measured by PROs such as BASDAI, BASFI, ASQoL, and HAQ-s. Notably, higher PRS in the HLA region was associated with better outcomes in all cases. While our own analysis of the clinical effects of HLA-B27 yielded no significant results in relation to these clinical instruments, directionality generally agreed with our PRS findings and with the larger Regisponser study by Arévalo et al. (2018). Our study is the first to observe these findings in the context of a polygenic risk score. However, it remains unclear why genetic risk in the HLA compartment, which is associated with earlier disease onset, increased spinal disease progression (Atagunduz et al., 2010), and comorbid disease, should paradoxically be linked to more favourable patient-reported outcomes in terms of disease activity and patient functioning.

In contrast to our findings with the other PROs, we observed no relationship between BASMI and PRS within our study. It's worth noting that BASMI is unique among these clinical indices as it is conducted entirely within the rigorous confines of metrology assessments in the clinic, rather than relying on self-reported outcomes. Therefore, it is a strong indicator for physical disability. It is known to correlate poorly with BASFI which was designed to assess the relative ease with which patients engage daily tasks (Zochling, 2011).

A weak inverse association between PRS and IBD was found in PRS outside of the HLA region. This may indicate some divergence of genetic risk in patients who develop comorbid colitis. However, this association was not significant after accounting for false discovery. AAU, while associated with HLA-B27 (Table 3.12) showed no association with PRS in the HLA region (Table 3.26). Interestingly, high impact SNPs correlated with HLA-B27 were not necessarily correlated with AAU. Some of these, even within the HLA-B gene region showed opposite directionality. While the evidence for this finding is limited, it is worth noting that one of these SNPs, rs2524094, is putatively linked by eQTL with HLA-C, a gene which was found to be down-regulated among several cell subsets in AxSpA patients from our single cell analysis (Chapter 5).

As a final note, the above example with AAU demonstrates how aggregated genetic risk can act to mask associations when applied to phenotypes beyond the model on which they were trained. Thus, while PRS have significant potential as biomarkers for specific disease, the development and application of combined models for inflammatory disease faces significant challenges.

3.9 Limitations of this study

While performing these studies, we encountered several limitations, some of which were discussed earlier. The size of our GWAS study was small, which impeded our ability to identify genetic traits linked to phenotypes. Additionally, the splitting of our cohort based on clinical characteristics further exacerbated this issue. Nevertheless, our QQ plots offered some early guidance on whether unique genetic risk is likely to play a role in these phenotypes on a genome-wide basis.

The cross-sectional design of our study limited our ability to retrospectively examine the onset of AAU in relation to AxSpA disease duration. We additionally lacked the diagnosis date of first AAU episode and data on recurrence. Differentiating between ulcerative colitis and Crohn's disease was not possible based on the data collected in the ASRI study. Furthermore, a significant number of patients involved in the ASRI study were actively receiving DMARD or NSAID therapy at the time of data collection. This is likely to have impacted most clinical instruments and markers of inflammation, which are particularly sensitive to therapeutic intervention.

The base GWAS data used for PRS calculation included an East Asian sub-cohort, comprising 3,117 of 25,764 participants. Moreover, the definition of the disease studied here (AxSpA) was slightly different from that of the base GWAS (AS). Furthermore, the utilisation of PRS within the context of this study involved application of polygenic risk to clinical phenotypes other than those on which the PRS are trained. Collectively, these factors are expected to have resulted in some reduction in power. We did not have access to an auxiliary AxSpA population to test PRS for over-fit which may be associated with selection of '*thresholds of best fit*'. To mitigate the likelihood of over-fit, we have therefore presented PRS regression data for all P-value thresholds, correcting outputs for false discovery. This will have resulted in a reduction in misleading associations at the cost of true positive detection.

Chapter 4

A meta-analysis of genetic risk in AxSpA

4. A meta-analysis of genetic risk in AxSpA

4.1 Introduction

Genome wide association studies (GWAS) have been instrumental in unravelling genetic risk in complex diseases. By identifying genetic variants associated with specific disease states, GWAS have contributed to our understanding of disease aetiology and facilitated the development of new therapies targeting specific biomolecular pathways. GWAS typically focus on common heritable traits with well-defined phenotypic characteristics and require large sample sizes to detect variants of small effect size. The largest GWAS performed to date for AS, which included a mixed European and Asian AS cohort, identified a total of 25 discrete genomic loci associated with the condition. This study was based on a population of 25,764 individuals, of whom 10,619 were AS cases (Cortes et al., 2013b). New loci identified in this study confirmed the polygenic nature of AS, with associated SNP variants located near to genes including *IL6R*, *FCGR2A*, *UBE2E3*, *GPR35*, *BACH2*, *ZMIZ1*, *NKX2-3*, *SH2B3*, *GPR65*, *IL27-SULT1A1*, *NOS2*, *TYK2* and *ICOSLG*. A later large meta-analysis that examined five related diseases, including AS, CD, UC, psoriasis and primary sclerosing cholangitis, quantified the genetic overlap between these conditions and identified a further 17 AS associated variants through their shared significance between these diseases (Rostami et al., 2020, Ellinghaus et al., 2016). This latter study revealed that AS shares approximately 50% of genetic heritability with each of Crohn's disease, ulcerative colitis and psoriasis (Ellinghaus et al., 2016).

Collectively, studies such as these have identified a total of 115 independent genetic associations for AS in 90 distinct genomic regions (Brown and Wordsworth, 2017). However, to identify further variants with smaller effect sizes, larger studies or meta-analyses that combine existing data from multiple GWAS will be required. Meta-analysis is a particularly powerful approach for smaller studies that may lack the statistical power to detect novel associations beyond those previously identified. The ASRI study cohort is comprised of 641 genotyped AxSpA patients of which 496 meet the modified New York criteria for AS. This study population has not previously been included in any AS or AxSpA GWAS study and thus represents a unique opportunity to identify additional genetic risk factors associated with the disease.

Due to recent changes in diagnostic criteria (Rudwaleit et al., 2009), which now include AS as a subtype of AxSpA, the physiological characteristics defining AS and AxSpA are becoming increasingly blurred. The clinical definition of AS using the modified New York (mNY) criteria, which were last updated in 1984, hinges on the presence of radiographic sacroiliitis. However, the emergence of MRI which can detect pre-structural inflammatory changes in patients with non-radiographic AxSpA (nrAxSpA) (Galloway and Machado, 2022, Robinson et al., 2021)

allows for earlier therapeutic intervention which may delay or prevent radiographic change altogether (Maguire et al., 2022a, Torgutalp et al., 2022). Defining subsets within modern patient cohorts is therefore difficult, which may lead to more genetically diverse study populations. This may subsequently impact the specific genetic markers or their relative effect sizes when compared to those observed in previous GWAS studies specific to AS (Kenyon et al., 2022). Despite this, our previous work in Chapter 2 Section 3.7 showed that the overall difference in genetic risk profiles, as measured by PRS, between nrAxSpA and AS patients within the ASRI cohort is negligible. We found no significant difference in PRS scores in HLA or non-HLA regions between AS and nrAxSpA patients after false discovery was accounted for. Therefore, we included both AS and AxSpA patients in meta-analysis between the ASRI study cohort and the largest AS GWAS study to date on the assumption that these two cohorts represent a similar disease phenotype (Cortes et al., 2013b).

4.2 Aims

- To perform a genome wide association analysis between the ASRI study cohort of 641 patients and 1,911 Irish healthy control individuals.
- To combine the findings of the ASRI AxSpA GWAS analysis with those of the largest GWAS study performed in AS patients to date.
- To identify nearby genes and expression quantitative trait loci (eQTL) associated with novel variants, where possible.
- To explore the relationships of novel variants with those previously identified in AxSpA and other auto-inflammatory diseases.

4.3 Methods

4.3.1 Acquisition of control samples and quality control

The available summary data from a previous AS GWAS (IGAS_GWAS) was used (Cortes et al., 2013a). We obtained genetic variant data from the Trinity Translational Medicine Institute Biobank containing 1,911 sequenced Irish healthy controls (TTMI controls). Variants matching SNPs reported in the summary data of the IGAS_GWAS (Cortes et al., 2013b) were extracted from the sequenced control data set by Dr. Deepak Bharti. Strands were aligned to match previously phased data from ASRI (Chapter 3 Section 3.3.5) by Dr. Deepak Bharti. ASRI data previously imputed (Chapter 3 Section 3.3.5) was merged with TTMI control data in Plink V1.9. SNPs were pruned where the genotyping rate was <98% or probability of Hardy-Weinberg equilibrium $P > 1 \times 10^{-10}$. Data was processed to structurally match that of the IGAS summary data: variants which were ambiguous or missing from either IGAS or ASRI data sets were pruned and variants on opposing strand were flipped.

4.3.2 Genome wide association analysis of ASRI data

A new GWAS was performed using ASRI AxSpA patients and Irish healthy controls (ASRI_GWAS). Briefly, Plink binary files were edited to assign ASRI patients as *cases* and Irish control data as *controls*. Association analysis (χ^2) was performed in Plink V1.9 (Chang et al., 2015) using the --assoc function.

To prevent spurious associations caused by variant call discrepancies between sequencing and chip genotyping data, we pruned genetic associations of unexpectedly high statistical significance in the ASRI_GWAS. We therefore compared the genome-wide significance of SNPs between the ASRI and IGAS_GWAS studies to identify statistical outliers within the ASRI case/control data set. GWAS summary data outputs for ASRI_GWAS and IGAS_GWAS were imported into Microsoft Excel together and P-values between the two data sets were compared. All variants with P-value <0.001 within the ASRI GWAS summary data set were cross-checked against the IGAS GWAS summary data and SNPs were pruned from the ASRI data set where $((P_{ASRI})/100) < (P_{IGAS})$.

Summary data outputs were then uploaded to the FUMA GWAS (FUMA V. 1.41, MAGMA V. 1.08) (Watanabe et al., 2017) online service for analysis. FUMA GWAS is an online service which integrates the features of several utilities, including a core GWAS analysis and annotation

module, MAGMA for gene-based analysis and GTEx for eQTL. SNP to gene association was set to 10Kb range.

4.3.3 Meta-analysis of GWAS studies

Unknown reference allele entries in the IGAS_GWAS summary data set (denoted by “?”) were replaced with their corresponding reference data from the ASRI data set in Microsoft Excel. Amended summary data were outputted in tab delimited format.

Meta-analysis of 2,552 individuals from the ASRI_GWAS data with 25,764 individuals from the IGAS study was performed in Metal (release August 2018) (Willer et al., 2010) using a sample size based, fixed-effects model as follows:

#ASRI_GWAS data:

MARKER SNP

ALLELE A2 A1

EFFECT log(OR)

PVALUE P

DEFAULTWEIGHT 2552

#IGAS_GWAS data:

MARKER rsid

ALLELE other_allele effect_allele

EFFECT beta

PVALUE p

DEFAULTWEIGHT 25764

4.3.4 Quality control: assessment of genetic association directionality between IGAS and ASRI

To assess the directionality of AxSpA association of variants between ASRI and IGAS summary data, we analyzed variants representing independent loci in all covered genomic regions. First, we clumped SNPs in the ASRI study using an r^2 cut-off of 0.1 within a 250Kb scanning window in Plink V1.9, resulting in a total estimate of 14,578 representative independent variants. Next, we performed genome-wide association analysis (χ^2) on these representative variants in Plink V1.9 using the --assoc function. We then compared the outputs by meta-analysis as detailed in section 4.3.3. The meta-analysis results were processed in Microsoft Excel for directional concordance at various P-value thresholds ($P < 5 \times 10^{-8}$, 1×10^{-7} , 1×10^{-6} , 1×10^{-5} , 1×10^{-4} , 1×10^{-3} , 0.01, 0.1, and 0.5). We calculated 95% confidence intervals for concordance using a binomial distribution model in Minitab 21.1.

4.3.5 Determination of P-value cut off for significance of markers

Initially, genetic association was examined at the widely accepted genomic threshold for statistical significance of 5×10^{-8} . However, due to considerable pruning of variants during quality control and the immune-related focus of the Illumina ImmunoChip 2, this threshold may have been unnecessarily strict for the remaining number of independent SNPs after accounting for LD. To correct for type 1 error, we therefore used previously developed methodologies that consider the total number of independent SNPs used in the study (Li et al., 2012). The number of independent loci in the ASRI data set, after QC and accounting for LD, was estimated to be 29,258 using the Genetic Type I Error Calculator (GEC) (Li et al., 2012). Therefore, we calculated the corresponding P-value threshold to be 1.71×10^{-6} after Bonferroni correction for $\alpha=0.05$. Based on this, variants with statistical significance $P < 3.42 \times 10^{-8}$ ($\alpha < 0.001$) were considered highly significant.

4.3.6 Evaluation of meta-analysis results

Prior to upload to FUMA-GWAS (Watanabe et al., 2017), both meta-analysis and summary statistics data of the IGAS summary data were pruned of the extended MHC region between Chr6:27,000,000 – 34,000,000 (Watanabe et al., 2017). However, FUMA-GWAS presents genome wide association thresholds only at $P < 5 \times 10^{-8}$. Therefore, GWAS plots for the Bonferroni corrected threshold of genome-wide significance at 1.71×10^{-6} were prepared in R 4.13 using the QQMan package (Turner, 2014). Novel lead SNPs within meta-analysis peaks were marked manually in Photoshop 24.0.1 and annotated manually using the Human Genome Browser

service (Kent et al., 2002) for genome build GRCh37/hg19 and Haploreg V4.1 (Ward and Kellis, 2012) for correlated SNPs ($r^2 > 0.8$). Possible regulatory roles of implicated variants were retrieved from expression quantitative trait loci (eQTL) data available on the GTEx Portal online database (Carithers and Moore, 2015), Release V8.

4.3.7 Calculation of LD between novel and previously reported variants

LD calculations were performed using LDpair Tool, part of the LDLink suite 5.5.1 Release (11/15/2022) (Machiela and Chanock, 2015) with reference data from the 1000 genomes GBR and CEU populations (380 individuals), mapped to GRCh37 (hg19) (Siva, 2008).

Haplotype block and association analyses were performed using Haploview 4.1 (Barrett et al., 2004). Briefly, SNPs within the region of interest were extracted from the ASRI data post-imputation in Plink V1.9 (see Chapter 3 Section 3.3.5) and randomly thinned to no more than 50 SNPs using Plink's `--thin` function. SNPs of specific interest were then merged back into the remaining SNP subset. Data was processed for LD in Haploview 4.1 with haplotypes constructed either using confidence intervals (95%) or single spine of LD. Haplotype association analysis was performed within the Haploview 4.1 environment.

To investigate linkage between ASRI SNPs and previously reported variants not present in the ASRI data set, the LDHap tool in LDLink suite 5.5.1 Release (11/15/2022) (Machiela and Chanock, 2015) was used with the 1000 genomes GRCh37 CEU and GBR reference populations (Siva, 2008).

4.4 Results

4.4.1 Investigation of meta-analysis data at genome wide significance

An overview of the ASRI GWAS data, based on 2552 (641 patients and 1911 controls) individuals is presented in Fig. 4.1, with a corresponding QQ plot in Fig. 4.2. Consistent with previous GWAS studies on AS (Cortes et al., 2013b), the strongest genetic association was observed in the HLA region of chromosome 6 (Cortes et al., 2013b), followed by the *ERAP1* gene locus in chromosome 5 and a gene desert region of 2p15.

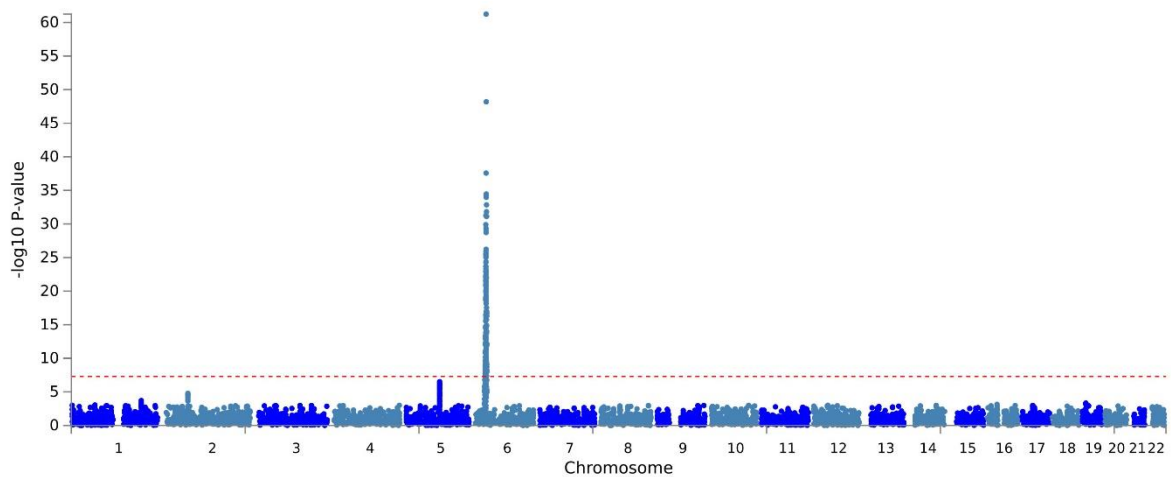


Fig. 4.1: Manhattan plot of genetic association to AS in the ASRI cohort, based on 69,636 SNPs in 641 patients and 1,911 controls. The dotted red line shows the P-value threshold for statistical significance (5×10^{-8}).

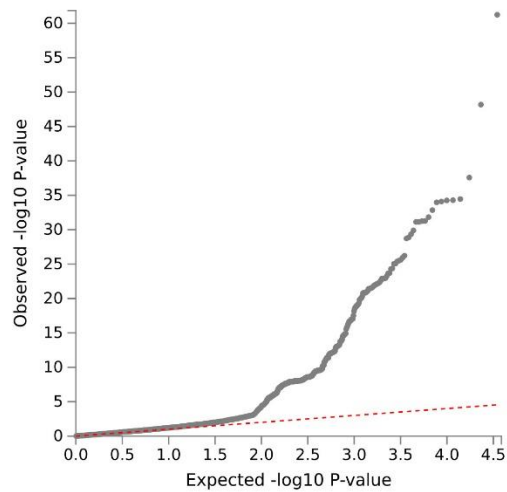


Fig. 4.2: QQ plot of statistical significance of genetic risk associated with AxSpA, based on 641 patients and 1,911 controls.

The directionality of association was compared between the ASRI and IGAS GWAS studies. After accounting for LD, 100% concordance was found between the ASRI and the reference (IGAS) GWAS for all loci statistically significant below a P-value threshold of 1×10^{-5} in the meta-analysis output. This was based on a total of 14,578 independent loci. The results are summarised in Table 4.1.

Table 4.1: Concordance of directionality of association stratified by P-value

Upper P-value threshold	SNP count	n	Concordance (%)	95% CI
1	14578		56.9	(56.1 - 57.8)
0.5	9714		66.8	(65.8 - 67.7)
0.1	3428		74.4	(73.0 - 75.9)
0.01	927		83.3	(80.7 - 85.6)
0.001	368		89.1	(85.5 - 92.1)
1×10^{-4}	192		96.9	(93.3 - 98.8)
1×10^{-5}	141		100	(97.9 - 100)
1×10^{-6}	117		100	(97.5 - 100)
1×10^{-7}	104		100	(97.2 - 100)
1×10^{-8}	99		100	(97.0 - 100)

The genome-wide statistical significance of our meta-analysis of 69,636 SNPs, compared between 11,260 patients and 17,056 controls, is presented in Fig. 4.3. The genome-wide threshold for statistical significance in this output is 5×10^{-8} . Our data correlated well with that of the IGAS GWAS study (Supplementary Fig. S4.1) (Cortes et al., 2013b).

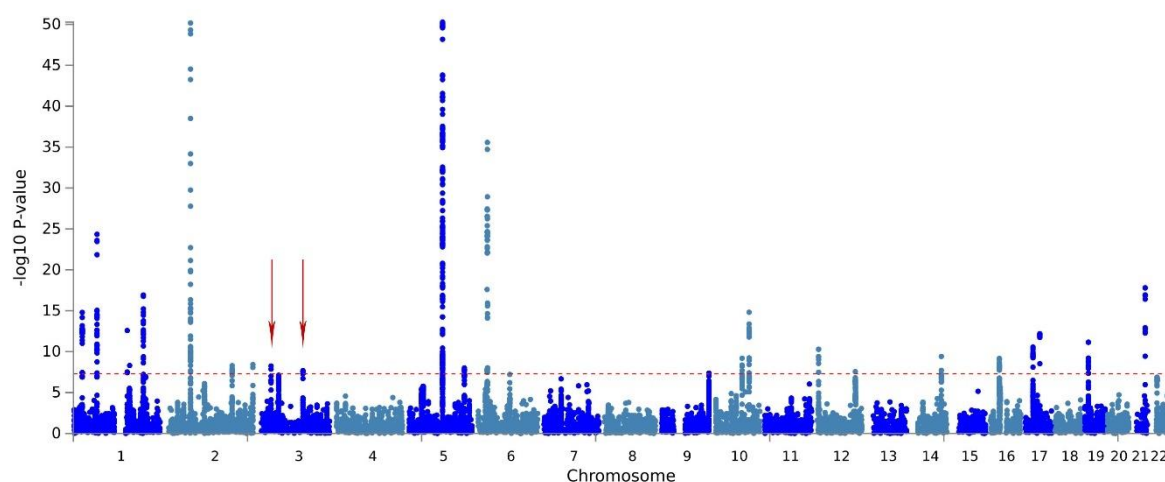


Fig. 4.3: Manhattan plot of genome wide significance, based on 69,636 SNPs in 11,260 patients and 17,056 controls. P-value threshold for statistical significance = 5×10^{-8} , as indicated by the dotted red line. Novel loci are indicated by red arrows. Note: SNPs within the MHC region, spanning Chr 6: 27,000,000 – 34,000,000, have been excluded from this graphical representation.

Two novel loci were identified in the regions of chromosome 3p and 3q. The localisation of these SNPs is presented in Table 4.2.

Table 4.2: Annotation of novel peak lead SNPs .

Lead SNP	Chromosome	Position	Nearby genes	Relationship	P
rs13093489	3	27,794,907	EOMES, LINC01981	Intergenic	5.725×10^{-9}
rs28653390	3	119,259,163	CD80	Intronic	2.079×10^{-8}

The lead variants of each of the two additional peaks detailed in Table 4.2, were cross-checked against published literature and online database services for evidence of prior implication in AxSpA.

The lead variant, rs13093489 which is located within an intergenic region of 3p (chr3:27,794,907) has previously been identified as a possible risk variant by the IGAS GWAS study (Cortes et al., 2013b), although it did not reach the cut-off for statistical significance in this study. Rs13093489 has also been linked to multiple sclerosis (International Multiple Sclerosis Genetics et al., 2011) through a correlated variant, rs11129295 ($r^2=0.872$, $D'=0.977$ (Machiela and Chanock, 2015)). Both rs13093489 and rs11129295 are eQTL for *EOMES* according to the NESDA NTR eQTL catalogue (Jansen et al., 2017, Westra et al., 2013). However, data available from the GTEx portal (Carithers and Moore, 2015) does not include data on any QTL effects of rs13093489, but does indicate that rs11129295 also affects expression of two further non-protein coding genes, *LINC02084* and *AC098614.2* in lung and arterial tissues, respectively (Fig. 4.4). Both *EOMES* (Fig. 4.5) and *LINC02084* (Fig 4.6) show enriched expression in whole blood and splenic tissues, tentatively suggesting an immunological role for these gene products. *AC098614.2*, however, is not enriched in immunologically relevant tissue types (Supplementary Fig. S4.2). Thus it seems possible by extension, that rs13093489 is a QTL for all three genes *EOMES*, *LINC02084* and *AC098614.2*.

The only other association within the same genetic region is through an intronic variant (rs10510607 – chr3:28,286,161) in the *CMC1* gene, located approximately 500Kb centromeric of *EOMES* which was associated through meta-analysis of AS and related diseases (Ellinghaus et al., 2016). However, there is no evidence that this more distant SNP is genetically correlated with rs13093489 ($r^2=0.0039$, $D'=0.096$ (Machiela and Chanock, 2015))

Fig. 4.4A

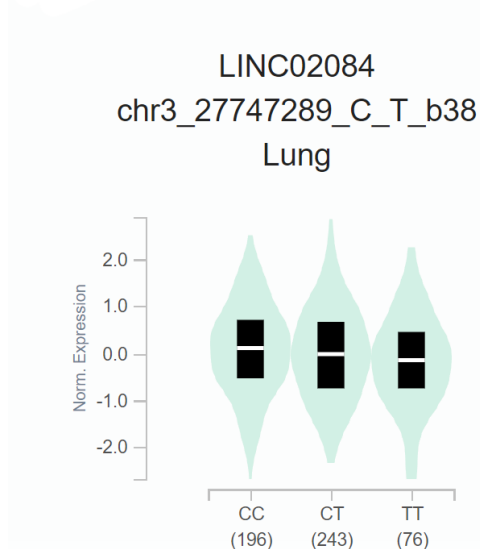


Fig. 4.4B

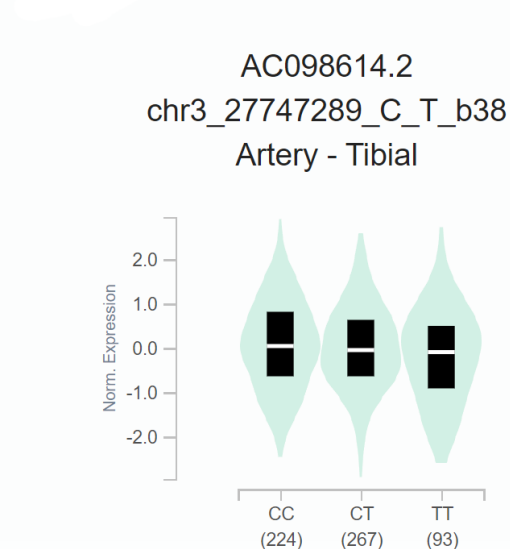


Fig. 4.4A: Expression of LINC02084 in lung tissue by rs11129295 genotype ($P=2.87 \times 10^{-5}$).

Fig. 4.4B: Expression of AC098614.2 in tibia arterial tissue by rs11129295 genotype ($P=7.89 \times 10^{-6}$).

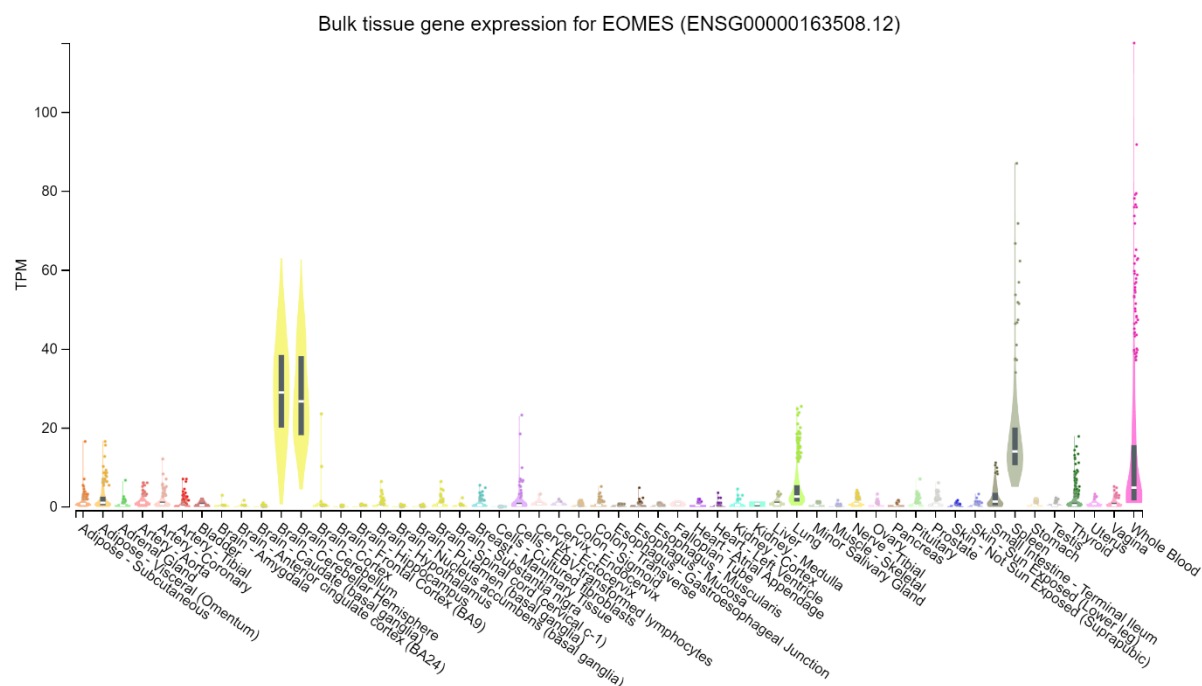


Fig. 4.5: Tissue expression of the EOMES gene from GTEx (Carithers and Moore, 2015).

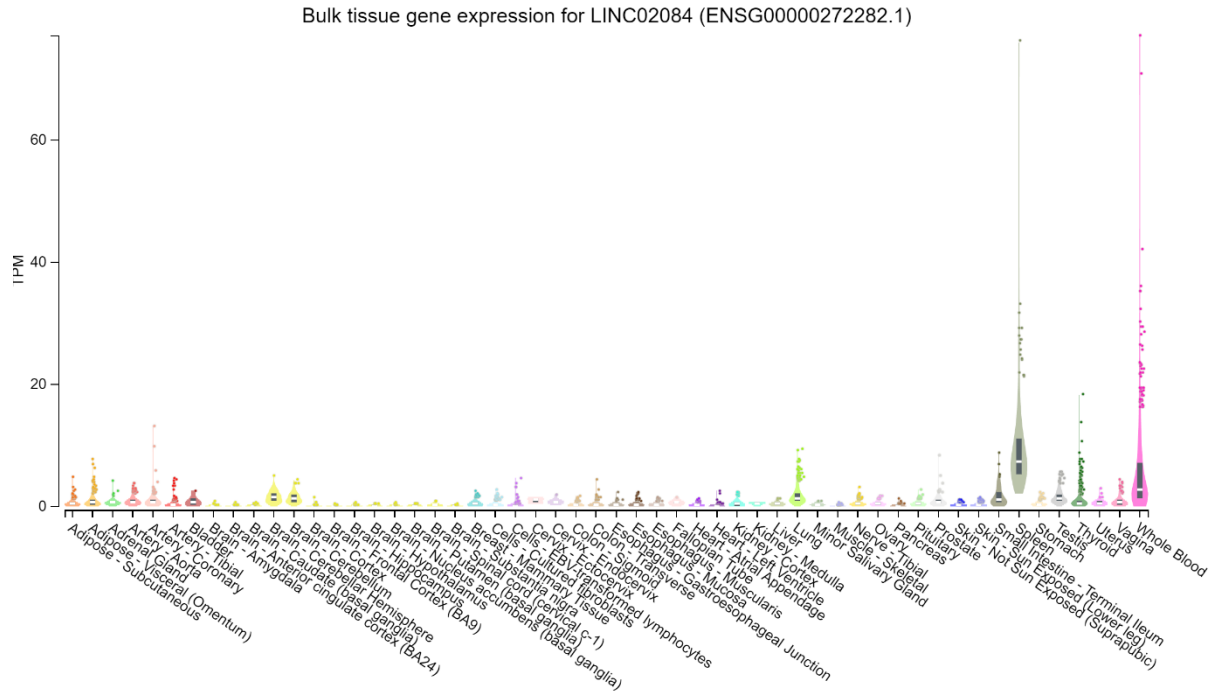


Fig. 4.6: Tissue expression of the LINC02084 gene from GTEx (Carithers and Moore, 2015).

The second novel lead variant, rs28653390, is situated within the non-coding region of the *CD80* gene and is not colocalised with any variants previously implicated in AxSpA. The minor allele, rs28653390-A was positively associated with AxSpA. Nearby variants, rs57271503 ($r^2=0.029$, $D'=0.895$ with rs28653390) and rs6804441 ($r^2=0.033$, $D'=1.0$ with rs28653390) (Machiela and Chanock, 2015), have previously been linked to primary biliary cholangitis (PBC) and systemic lupus erythematosus (SLE), respectively (Hitomi et al., 2019b, Yang et al., 2013b). Rs57271503 and rs6804441 are in partial LD ($r^2=0.788$, $D'=0.922$ (Machiela and Chanock, 2015), suggesting that these two SNPs may be tagging the same causative variant.

Given the proximity of rs57271503 and rs6804441 to rs28653390, we questioned whether they may also be associated with AxSpA. However, neither of these SNPs (nor any suitable proxies) were available in our meta-analysis. Within the ASRI data set, only rs57271503 was available

for analysis. We therefore investigated whether this SNP may also be associated with AxSpA within our study cohort of 641 AxSpA patients and 1,911 healthy Irish controls. We found that rs57271503 ($P=0.0006$) was likely to be protective against AxSpA, thus in opposite directionality of our meta-analysis lead variant, rs28653390 (Table 4.3). Notably, the directionality of rs57271503 in AxSpA matches that reported in PBC, where a protective association was also observed (Hitomi et al., 2019b).

Table 4.3: Association analysis of rs57271503 and rs28653390 in 641 AxSpA patients and 1,911 Irish controls

CHR	SNP	Base position	Freq. AxSpA	Freq. Control	P	OR
3	rs57271503	119244593	0.1718	0.2166	0.000566	0.7502
3	rs28653390	119259163	0.2082	0.1714	0.002936	1.271

To verify the independent associations of rs28653390 and rs57271503, we constructed haplotypes that included both SNPs as well as representative SNPs from the region between them, selected at random (see Fig. 4.7). We then performed AxSpA association analysis for these haplotypes, as shown in Table 4.4. We identified two significant haplotypes: one containing the possibly protective minor allele rs57271503-A, and the other containing the causative minor allele of rs28653390-A. All other haplotypes were unrelated to AxSpA. These findings indicate that rs28653390 is not correlated with rs57271503 which was previously associated to PBC, and suggest that rs57271503 is likely to be independently associated with AxSpA.

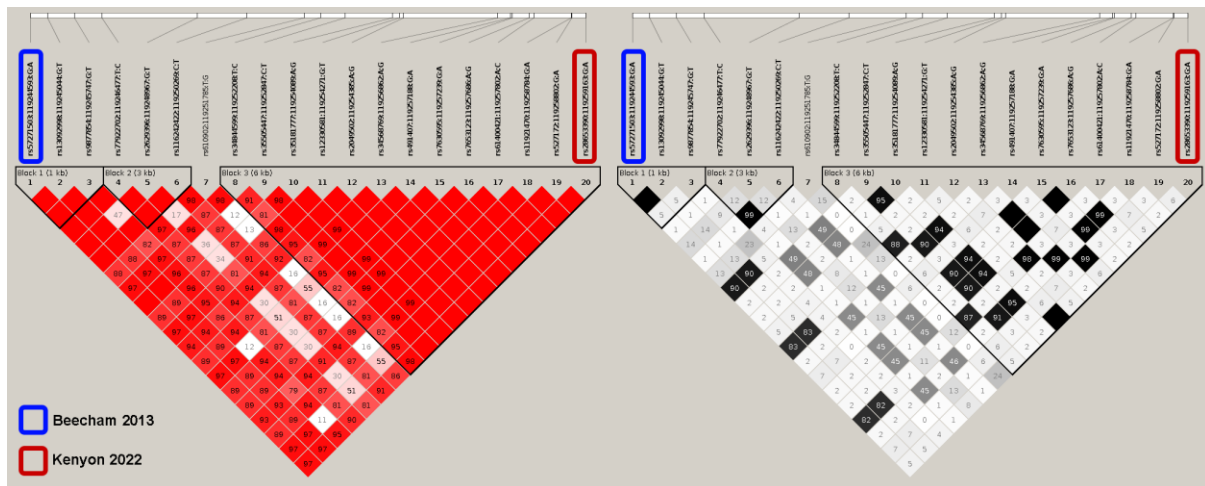


Fig. 4.7: LD block analysis of 641 AxSpA patients and 1,911 Irish controls between rs57271503 and rs28653390. Left: D-prime, right: r-squared. Haplotypes constructed using 95% confidence intervals.

Table 4.4: AxSpA associations of expanded haplotypes at the CD80 locus between rs57271503 and rs28653390.

Haplotype	rs57271503 Allele	rs28653390 Allele	Haplotype Freq.	Case Ratio Counts	Control Ratio Counts	Case,Control Frequencies	P Value*
ATTTTCGCCAGGAAGACGGG	A	G	0.183	196.9 : 1080.7	730.8 : 3007.5	0.154, 0.195	0.006
GGTTTCTTCATAAAGAAGGA	G	A	0.165	245.3 : 1032.3	588.9 : 3149.4	0.192, 0.158	0.0126
GGTTGCGTCAGAAGGAAGAG	G	G	0.154	190.7 : 1086.9	588.9 : 3149.4	0.149, 0.158	0.4803
GGGTGCTTCAGAAAGAAGGG	G	G	0.115	157.1 : 1120.5	423.9 : 3314.4	0.123, 0.113	0.4803
GGTTTCGTCAGAAAGAAGGG	G	G	0.104	127.3 : 1150.3	401.0 : 3337.2	0.100, 0.107	0.4803
GGTCGTGTTGAGAAGAAGG	G	G	0.061	89.6 : 1188.0	216.7 : 3521.5	0.070, 0.058	0.2332

*Corrected for false discovery by Benjamini-Hochburg method.

4.4.2 Investigation of further variants of interest

In genome-wide association studies, a threshold cut-off for statistical significance of 5×10^{-8} has been widely adopted, based on the Bonferroni correction of an expected 1 million independent loci across the human genome. However, this threshold is not universally applicable for an expected false discovery rate (α) of 5%. For instance, studies with deep sequencing may capture 1.5 million or more SNPs and rare variants, especially where highly diverse African populations are studied, leading to an unreported inflation of the false discovery rate where this widely adopted threshold is used (Sveinbjornsson et al., 2016, Fadista et al., 2016). Conversely, for studies which include smaller numbers of independent loci, the 5×10^{-8} threshold may be overly

stringent, resulting in a high type 2 error rate. To address this, some studies use Bonferroni correction based on the total number of available loci (Kim et al., 2014, Aslam et al., 2020), or correct the number of independent loci for LD to determine the effective number of independent SNPs (M_e) (Hébert et al., 2015, Melroy-Greif et al., 2017).

In our data set, we estimated M_e to include 29,258 independent SNPs using Genetic Type I Error Calculator (see Methods section 4.3.5) and calculated the threshold for genome-wide significance after Bonferroni correction to be 1.71×10^{-6} at $\alpha=0.05$. This revealed additional loci at this threshold of statistical significance which were not previously reported within the IGAS GWAS summary data, as indicated in Fig. 4.8. These were located at chromosomal locations in 1q, 2p, 3p, 5p, 7p, 7q, 11q and 22p.

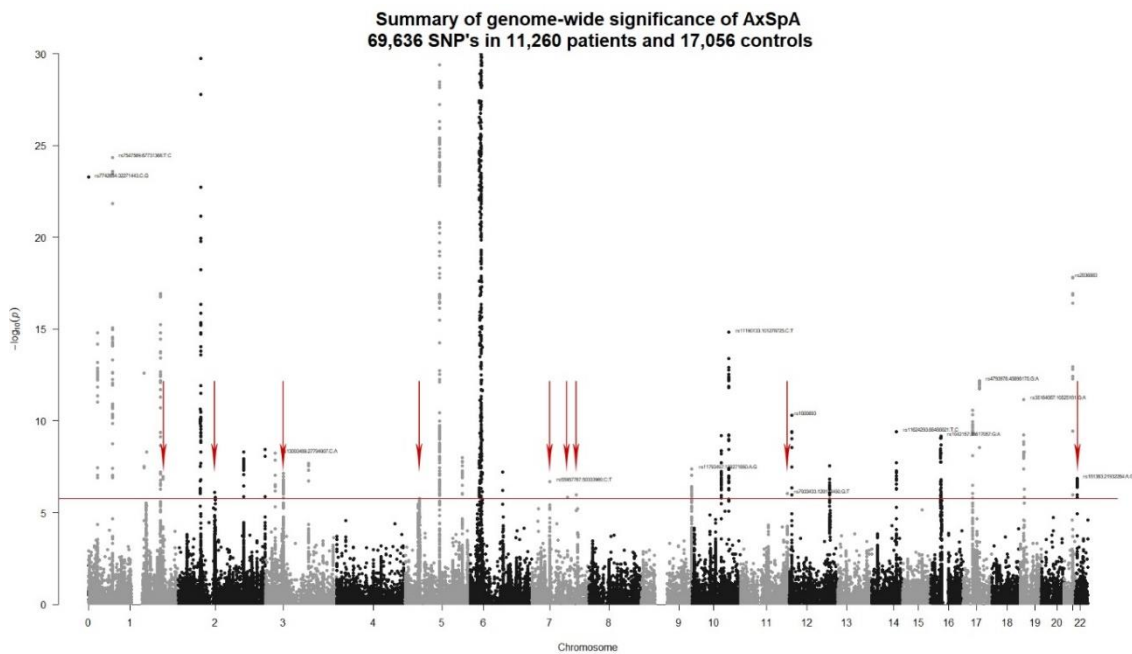


Fig. 4.8: Manhattan plot of genome wide significance, based on 69,636 SNPs in 11,260 patients and 17,056 controls with P-value threshold for statistical significance corrected for 29,258 independent variants (1.71×10^{-6}). Note: $-\log_{10}P$ has been cut-off at 30 in this graphical representation.

Table 4.5: Additional novel variants of significance after Bonferroni correction for multiple testing.

Chr	BP	Lead SNP	P (lead SNP)	Nearby gene(s)	Potential eQTL targets*	Previously reported associations	Outcome
1	206,972,072	rs4845141	1.11×10^{-7}	Upstream of <i>IL19</i> TSS	None reported.	Previously reported variants near <i>IL19</i> not in significant LD.	Novel
2	102,649,208	rs6707803	7.94×10^{-7}	Downstream of <i>IL1R2</i>	None reported.	Not previously reported.	Novel
3	49,785,133	rs7374183	7.69×10^{-8}	Intronic variant in <i>IP6K1</i>	<i>UBA7</i> , <i>RBM6</i>	Not previously reported. Nearby SNPs associated with UC and Crohn's	Novel
5	40,438,097	rs73090828	1.69×10^{-6}	Upstream of <i>PTGER4</i>	<i>PTGER4</i> , <i>CARD6</i>	Previously reported in meta-analysis under lead SNP rs1992661.	Previously reported
7	50,333,960	rs55957767	2.13×10^{-7}	Upstream of <i>IKZF1</i>	<i>IKZF1</i>	Previously reported variants near <i>IKZF1</i> not in significant LD.	Novel
7	100,315,517	rs1734907	1.52×10^{-6}	Upstream of <i>EPO</i>	<i>EPHB4</i>	Not previously reported.	Novel
7	124,439,599	rs2402752	1.13×10^{-6}	Intergenic <i>C7orf77/POT1</i>	<i>POT1-AS1</i> , <i>GPR37</i>	Not previously reported.	Novel
11	128,194,450	rs7933433	9.19×10^{-7}	Upstream of <i>ETS1</i>	None reported.	Previously reported in meta-analysis under same lead SNP.	Previously reported
22	21,932,264	rs181363	1.35×10^{-7}	<i>UBE2L3</i>	<i>CCDC116</i> , <i>UBE2L3</i>	Previously reported in meta-analysis under lead SNP rs2266961.	Previously reported

*Potential eQTL targets sourced from GTEx Release V8 (Carithers and Moore, 2015).

We investigated each of the additional lead SNPs presented in Fig. 4.7 for colocalised genes and previously reported AxSpA associations. Table 4.5 above summarises the results of our investigations.

The lead SNP rs4845141 (chr1:206,971,972-206,972,172, GRCh37/hg19) is located at a site intronic to transcript variants 1 and 4 of *IL19* and upstream of the transcription start sites (TSS) of other *IL19* transcript variants and is linked with expression of *IL24* (Westra et al., 2013). This SNP has not previously been reported in association to AS. Although a nearby variant, rs3024493, was previously associated with AS by Ellinghaus et al (Ellinghaus et al., 2016), it was found not to be in significant LD with rs4845141 ($r^2=0.083$, $D'=0.608$ (Machiela and Chanock, 2015)).

Rs6707803 (Chr2:102,649,108-102,649,308) is situated approximately 5KB downstream of the *IL1R2* gene and is an eQTL for *IL1R1* (Carithers and Moore, 2015). This region has been previously suspected in AS through the nearby variant, rs10185424 (Cortes et al., 2013b). Both rs10185424 and a second colocalised SNP, rs13022757 have been reported in association with IBD (Liu et al., 2015, de Lange et al., 2017). However, these variants were not available within our data set for haplotype analysis. Nevertheless, examination of haplotypes within the 1000 genomes CEU and GBR (GRCh37) cohorts suggests these variants have no specific pattern of concurrence (Clarke et al., 2012, Machiela and Chanock, 2015) (Fig. 4.9). Another SNP, rs917997, which has been functionally linked to both coeliac disease and Crohn's disease (Dubois et al., 2010) was not in significant LD ($r^2=0.018$, $D'=0.317$ (Machiela and Chanock, 2015)). Investigation of proxy SNPs for these previously published variants from within the ASRI meta-analysis suggests that all are weakly associated with AxSpA (Supplementary Table S4.2).

RS Number	Position (GRCh37)	Allele Frequencies	Haplotypes				
rs13022757	chr2:102613546	C=0.516, G =0.484	C	G	G	G	C
<u>rs6707803</u>	chr2:102649208	A=0.95, G =0.05	A	A	A	G	A
rs10185424	chr2:102662888	G=0.539, T =0.461	G	T	G	T	T
Haplotype Count			181	143	23	18	14
Haplotype Frequency			0.4763	0.3763	0.0605	0.0474	0.0368

Fig 4.9: Haplotypes of significant tag SNPs derived from 1000 genomes data for CEU and GBR populations. Analysis performed in LDHap (Machiela and Chanock, 2015, Clarke et al., 2012). Disease associated alleles highlighted in red. Underlined: meta-analysis lead SNP. AxSpA disease associated allele, rs6707803-G is not strongly correlated with either of the IBD-associated variants, rs13022757-G or rs10185424-T.

Rs7374183 (Chr3: 49,784,883-49,785,383, hg19) is located within an intronic region of the *IP6K1* gene (Sherry et al., 2001). There have been no reports of association of this gene to AS, but multiple GWAS have linked *IP6K1* to UC and Crohn's disease through several nearby variants (Anderson et al., 2011, Jostins et al., 2012, Liu et al., 2015, de Lange et al., 2017). These variants were not available for analysis within our data set. However, examination of haplotypes constructed from these variants and rs7374183 within the 1000 genomes CEU and GBR (GRCh37) cohorts reveals close correlation of minor alleles, indicating overlap in tagged eQTL(s) (Fig. 4.10). eQTL data (Carithers and Moore, 2015) indicates that the lead SNP, rs7374183, regulates expression of *UBA7* and *RBM6*, but not *IP6K1* in whole blood (Fig. 4.11). It is interesting to note that none of the aforementioned SNPs which were previously implicated in UC or Crohn's were linked to expression of *IP6K1*. Several of these SNPs showed differential expression for *RBM6* and for a third gene, *MST1* (data not shown).

UBA7 is a known *STAT1* gene target (Przanowski et al., 2014) that has a defined immunological function, including a role in CD8 cytotoxic T-cell coordination (Fan et al., 2020, Manry et al., 2017). *RBM6*, on the other hand, is a known RNA binding protein and tumour suppressor which has been linked to several types of cancer (Wang et al., 2019, Zabarovsky et al., 2002). Although there have been no reports of association of *IP6K1* with AxSpA, the eQTL data points to a potential regulatory role for SNPs in this region on *UBA7* and *RBM6* in pathogenesis of the disease.

RS Number	Position (GRCh37)	Allele Frequencies	Haplotypes				
rs9836291	chr3:49697459	G=0.732, A =0.268	G	A	G	G	A
rs4855884	chr3:49704077	G=0.542, A =0.458	G	A	G	A	A
rs1131095	chr3:49714225	T=0.732, C =0.268	T	C	T	T	C
rs9822268	chr3:49719729	G=0.732, A =0.268	G	A	G	G	A
rs3197999	chr3:49721532	G=0.742, A =0.258	G	A	G	G	G
<u>rs7374183</u>	chr3:49785133	G=0.521, A =0.479	G	A	A	G	A
Haplotype Count			126	98	80	72	4
Haplotype Frequency			0.3316	0.2579	0.2105	0.1895	0.0105

Fig. 4.10: Haplotypes of variants near to *IP6K1* previously associated to IBD together with novel AxSpA associated variant rs7374183. Analysis performed in LDHap (Machiela and Chanock, 2015). Disease associated alleles are highlighted in red. Underlined: meta-analysis lead SNP. A common haplotype at 25.79% frequency was found to contain minor (disease associated) alleles for all SNPs investigated, including ASRI meta-analysis lead SNP, rs7374183.

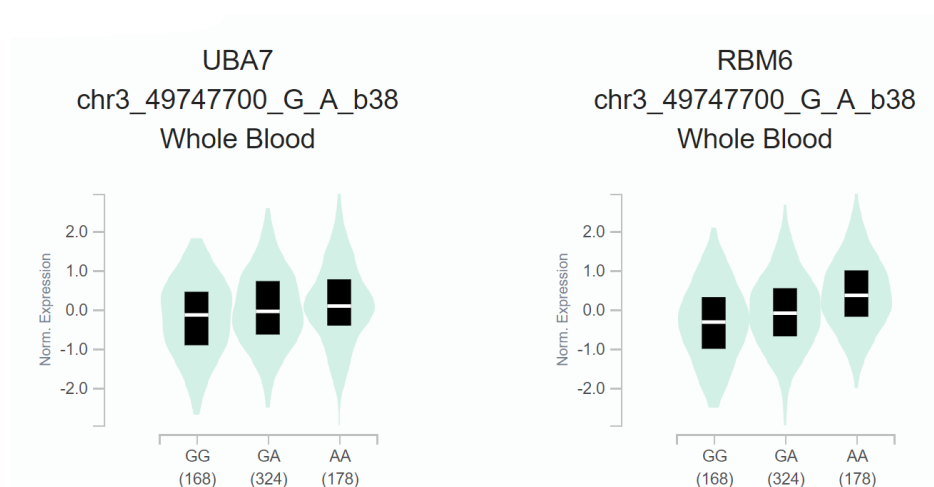


Fig. 4.11: Expression of UBA7 ($P=2.4 \times 10^{-58}$) and RBM6 ($P=2.0 \times 10^{-53}$) by lead SNP rs7374183 genotype.

Rs73090828 (Chr5: 40437997-40438197) is located approximately 250Kb upstream of *PTGER4* and has no documented associations to AS. However, *PTGER4* has previously been implicated in AS through a variant, rs12186979, which is not in LD with rs73090828 ($r=0.049$, $D'=0.552$)

[1]. Another SNP associated with AS, rs1992661 - which was identified through meta-analysis of inflammatory diseases (Ellinghaus et al., 2016), is in LD with the lead variant, rs73090828 ($r^2=0.106$, $D'=1.0$ (Machiela and Chanock, 2015)). Rs1992661 (O/R=1.24, $P=0.001$) was also positively associated to AxSpA in our cohort but with a lower effect size than rs73090828 (O/R=1.30, $P=0.003$).

We suspected that our novel SNP, rs73090828, and previously reported SNP, rs1992661, may be tagging the same causative haplotype, given their high D' (Fig. 4.12). To confirm this, we used the ASRI cohort and Irish control data to interrogate the simple haplotypes of rs1992661_G:A and rs73090828_C:A (Table 4.6). We found that the haplotype containing the minor alleles of both tag variants was positively associated to disease, while the converse was also true. This implied overlap between these alleles in tagging a local eQTL. Notably, the haplotype containing the minor allele rs1992661-A and major allele, rs73090828-C showed no significant association. We confirmed this by using haplotype maps which included additional SNPs situated between these two variants. We found that positive disease association was observed only where rs73090828-A was present. Rs1992661-A, meanwhile was associated with AxSpA only in haplotypes where rs73090828 was also present (Table 4.7). All remaining haplotypes containing rs1992661-A showed no disease association after accounting for false discovery.

Our data, therefore, verifies both that rs73090828-A is a more efficient tag of disease causation at this locus and that rs1992661-A is likely to be a passenger on this haplotype.

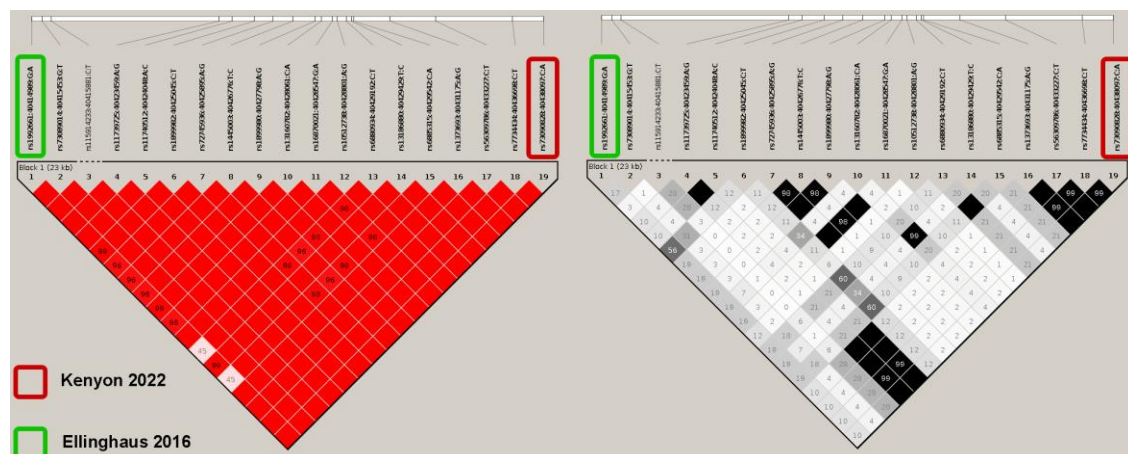


Fig. 4.12: LD block analysis of 641 AxSpA patients and 1,911 Irish controls between rs73090828 and rs1992661. Left: D-prime, right: r-squared. Haplotypes constructed using 95% confidence intervals.

Table 4.6: AxSpA association of simple haplotypes of rs1992661 and rs73090828.

Haplotype		Freq.	Case, Control		Directionality	P
rs1992661	rs73090828		Ratio	Counts		
A	C	0.455	0.469	0.450	1.447	N/A
G	C	0.409	0.371	0.423	10.719	Protective
A	A	0.136	0.160	0.128	8.704	Causation

Table 4.7: AxSpA association of derived expanded haplotypes between rs1992661 and rs73090828.

Block	Haplotype		Frequency	Case, Control		ChiSq	Directionality	P*
	rs1992661	rs73090828		Frequencies				
AGAATACAAGATCAACCC	A	C	0.222	0.237	0.217	2.244	N/A	0.201
GGAACACACGACTCACCC	G	C	0.205	0.198	0.208	0.55	N/A	0.55
ATAATACACGATTAACCC	A	C	0.205	0.202	0.206	0.088	N/A	0.767
AGGCCACACGACTCGTTA	A	A	0.136	0.160	0.128	8.225	Causation	0.025
GGAACGTGCAATTAACCC	G	C	0.124	0.108	0.130	4.258	N/A	0.114
GGAACACACGGCTCACCC	G	C	0.078	0.065	0.082	3.62	N/A	0.114

*Corrected for false discovery by Benjamini-Hochburg method ($\alpha=0.05$).

Three lead variants were identified in chromosome 7. The first of these, rs55957767 (Chr7: 50,333,860-50,334,060) is located upstream of the *IKZF1* gene, which was previously implicated in a meta-analysis (Ellinghaus et al., 2016). However, the SNP associated in that study, rs4917129 (Chr7: 50323174), is not in LD with the SNP observed here, rs55957767 ($r^2=0.025$, $D'=0.256$), which thus appears to be a novel association.

The second variant, rs1734907 (Chr7: 100,315,267-100,315,767) is an intergenic SNP positioned 2kb upstream of *EPO* and 11kb downstream of *POP7*. It has been associated with IBD and is an eQTL affecting the nearby receptor tyrosine kinase gene, *EPHB4* (Pouget et al., 2019), which is involved in vasculature morphogenesis and homeostasis (Erber et al., 2006, Luxán et al., 2019). *EPHB4* has also previously been implicated in neoplasia (Noren and Pasquale, 2007).

The third variant, rs2402752 (Chr7: 124,439,349-124,439,849), is another intergenic SNP positioned 9kb upstream of *C7orf77* and 25KB downstream of *POT1*. The locus is an eQTL for the poorly characterised *POT1-AS1* (130KB upstream) and *GPR37* (33KB downstream) genes (Fig. 4.13). While *GPR37* has not previously been linked to auto-inflammatory disease through GWAS it has a putative immunological role in promoting macrophage phagocytosis and resolution of inflammatory pain, and further linked to survival of sepsis in mice (Bang et al., 2018, Bang et al., 2021). *POT1-AS1*, meanwhile, is suspected to regulate an enzyme of glucose metabolism, PDK1 (Chen et al., 2021c). It has been linked to chromosomal instability and malignancy through indirect regulation of pyruvate dehydrogenase kinase 3 (Went et al., 2018, Aksenova et al., 2021, Chen et al., 2021c).

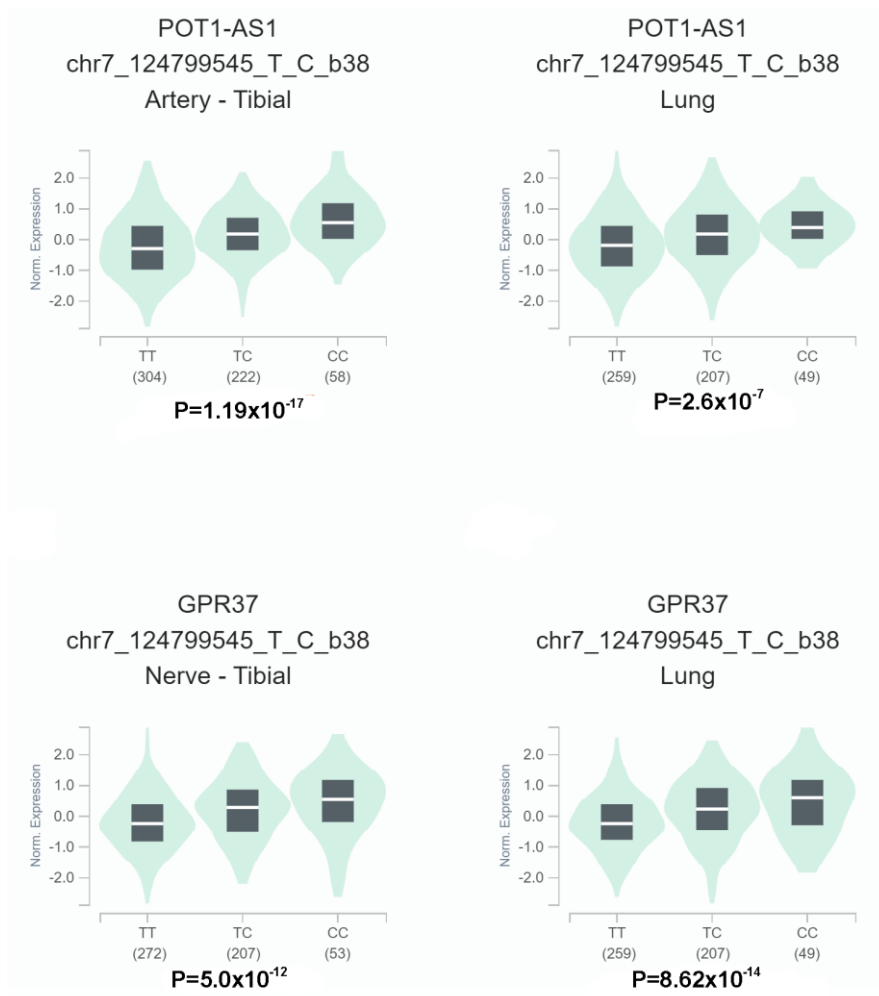


Fig. 4.13: Relationship between rs2371108 (chr3:27715527 GRCh38/hg38) genotype and bulk tissue gene expression of POT-AS1 in tibial artery and lung and GPR37 in tibial nerve and lung. Sourced from GTEx Portal (Carithers and Moore, 2015).

Finally, variants in chromosomes 11q (rs7933433) and 22p (rs181363) near the *ETS1* and *UBE2L3* genes, respectively, were found to be previously reported in association to AxSpA (Ellinghaus et al., 2016) either through the same or closely correlated lead SNPs (Table 4.5).

4.5 Discussion

In this meta-analysis we identified eight previously unreported loci associated with AxSpA. Of these, two reached the genome-wide significance level, while the remainder were significant below a false discovery threshold of 5%. Notably, several of the novel variants we report here are located within or near to genes with established roles in lymphocyte function. Of particular interest is the implication of *CD80*, which functions within antigen presentation, a key pathway linked to AxSpA disease. CD80 is highly expressed on antigen presenting cells, where it acts as part of the co-stimulatory signal complex in the antigen presentation event by interacting with its ligands, CD28 or CTLA4. These ligands are expressed on T-cells and respectively upregulate or inhibit T-cell activation. CD80 is furthermore involved in a carefully balanced system where it competes with CD86 for CD28 binding sites (Kennedy et al., 2022). The expression of CD80 and CD86 are tightly regulated by cell type, developmental stage and activity. In addition, CD80 appears to be capable of disabling CTLA-4 molecules by triggering targeted ubiquitination (Kennedy et al., 2022). CD80 has previously been linked to autoimmune disease, including SLE, MS, primary biliary cirrhosis, and autoimmune thyroid disease (Yang et al., 2013b, Liu et al., 2012, Saevarsdottir et al., 2020). Therapeutic intervention of this pathway using abatacept has been successful in treating TNF inhibitor (TNFi) refractory rheumatoid arthritis and PsA (Blair and Deeks, 2017, Noisette and Hochberg, 2018). However, trials in TNFi refractory AS have returned disappointing results (Lekpa et al., 2012). The lead variant reported in this study is situated close to a variant previously reported as associated with PBC. Through haplotype analysis we also identified a variant within the *CD80* 3' UTR, rs57271503, that has previously been reported in PBC and which may be independently associated with AxSpA with same directionality ($P=6 \times 10^{-4}$). However, further investigation within a larger study size is still needed to confirm this association.

The second lead variant reaching genome-wide significance, rs13093489 is situated within an intergenic region centromeric of *EOMES* and has previously been highlighted as probably AS associated (Cortes et al., 2013b). A variant in high LD with it, rs11129295 ($r^2=0.91$), has also been linked to MS, with eQTL evidence suggesting that it affects *EOMES* gene expression (Westra et al., 2013). This variant has additionally been linked to expression of the long non-coding RNA gene, *LINC02084* (Carithers and Moore, 2015). The tissue expression profiles of *EOMES* and *LINC02084* follow a similar pattern, with enrichment observed in both whole blood and splenic tissues (Figs 4.5-4.6). Despite this correlation, no published study has yet investigated whether *LINC02084* regulates the nearby *EOMES* gene. Recent studies have identified a role for *EOMES* in suppressing Th17 differentiation of CD4⁺ T-cells (Chen et al., 2021b, Ichiyama et al., 2011). However, there remains no direct evidence from any study implicating *EOMES* as a functionally relevant gene in disease aetiology. *LINC02084* remains

poorly characterised and has not specifically been implicated in autoimmune disease either. It has been identified as a prognostic indicator in several neoplasms (Zhang et al., 2022b). Interestingly, reduced expression of *LINC02084* has been tentatively linked to SARS-CoV-2 disease severity (Shaath and Alajez, 2021). Although no expression data was available for the variant previously reported by Ellinghaus et al, rs10510607, its intronic location within the *CMC1* gene could indicate this is the most likely target (Ellinghaus et al., 2016). Interestingly, Hi-C data indicates interaction of both rs10510607 and rs13093489 with the long non-coding RNA gene, *LINC01981*, but not with *EOMES* (Schmitt et al., 2016). It is important to note that expression profiles reported in the GTEx and Hi-C databases do not include stimulated immune cells. More comprehensive eQTL databases will be required to establish how immunologically relevant SNPs in this region interact in specific activated immune cell subtypes.

The novel association near the *IL1R2* gene, rs6707803, reported here is an eQTL for *IL1R1* (Carithers and Moore, 2015). This is the second association within the *IL1R2* gene region which has been reported in AxSpA. Genetic association of the *IL1R2* region has previously been shown in a wide range of other conditions, including IBD, SLE, Parkinson's disease and rheumatoid arthritis (Lin et al., 2012, Fang et al., 2022a, Liu et al., 2021, Cortes et al., 2013b, Liu et al., 2015). This supports the notion that common underlying mechanisms contribute to each of these various conditions. Although we were unable to confirm that the tag variant identified in this study was common to IBD, we did find tentative evidence that SNPs implicated in coeliac disease are also linked to AxSpA. Interestingly, a nearby SNP, rs917997, which is associated with coeliac disease and likely also AxSpA (Supplementary Table S4.2) has been found to affect expression of *IL1R1* and a number of other local genes including *IL18RAP*, by modulating expression of a nearby long non-coding RNA with regulatory effects (Castellanos-Rubio et al., 2016).

The lack of eQTL association between the novel SNP identified in this study, rs7374183, and the *IP6K1* gene in which it resides was unexpected. However, eQTL data linking this SNP to the nearby *UBA7* gene in relevant tissues (Fig 4.11) suggests that *UBA7* may be a significant gene in AxSpA. This is supported by eQTL data which linked other imputed GWAS variants to differential expression of *UBA7* in IBD (Hulur et al., 2015). We found no evidence to support involvement of the *IP6K1* gene in inflammatory disease, based on eQTL evidence for the SNPs in this region we investigated. However, this work is again limited by the lack of available data on cell subtypes and activity. *IP6K1* may yet play a role under specific circumstances as it plays roles in neutrophil extracellular traps and cell migration (Madhi et al., 2019, Jadav et al., 2016). *UBA7* seems a rational candidate in AxSpA, given its known roles in CD8 cytotoxic T-cell coordination. Further work is needed to unravel the precise mechanisms underlying *UBA7*'s involvement in inflammatory disease.

Throughout this chapter we have observed a recurrent pattern of loci associated, either physically or through eQTL, with genes which were previously implicated in IBD. This is somewhat at odds with our findings in Chapter 3, Section 3.7 (Table 3.26), where we tentatively found an inverse relationship between polygenic risk and concomitant IBD in AxSpA patients. This could be explained by design of the PRS model, which was based on known AxSpA risk, and not for identifying associations with comorbid conditions. Furthermore, our polygenic risk did not examine AxSpA and IBD in isolation, which may have caused a perceived inversion effect where the relative effect sizes of variants differ between the two discrete conditions. Finally, it is important to consider that although some of the SNPs identified in this chapter affect the same genes as previously implicated in IBD, the SNPs reported here were often not in LD with those variants previously reported in IBD. Our data therefore supports previous reports of genetic commonality between AxSpA and IBD, but also indicates that each condition has a unique component of heritable risk which remains evident even where the two conditions occur comorbidly. Thus, the overall genetic risk profile for AxSpA with comorbid IBD is slightly different from that of isolated AxSpA.

While the use of tag variants by GWAS offer a broad-scale approach to genetic association analysis, they often fail to capture causative SNPs across diverse populations, resulting in underrepresentation in GWAS studies (Amariuta et al., 2020). For instance, our novel SNP, rs73090828, near the *PTGER4* gene was previously indirectly reported by Ellinghaus *et al* using SNP rs1992661, despite low LD between these SNPs in terms of r^2 . Our fine mapping work through haplotypes suggests that rs73090828 is likely to be an improved tag variant for this association within our population, although this may not be true for other populations. This type of work helps to focus in towards the likely locations of disease-causing SNPs, which is an essential next step for GWAS analysis. Furthermore, our data strongly suggests that multiple causal variants may flank genes of interest in AxSpA. This effect, which was particularly evident in our previous work on *ERAP1* gene region (Chapter 3, Section 3.6.1), is supported here by our findings around *CD80*, *IL19*, *IL1R2* and *IP6K1*. Identifying these exact variants would enhance the application of predictive tools such as PRS for use within the clinical setting and inform the design of future SNP based arrays as a low cost alternative to sequencing (Amariuta et al., 2020).

4.6 Limitations of this study

The phenotypes examined in the IGAS and our novel data sets were not exactly matched, which may have resulted in some loss of power due to population stratification caused by underlying differences in nationality and disease classification. Although our previous work suggests that the difference between AS and AxSpA patients within the ASRI data set is expected to be low, further population stratification is anticipated as the IGAS data set included both European and Asian AS patients. As a result, some loci may have experienced a reduction in power, while others may have been strengthened by the meta-analysis, particularly where tag SNPs mark functional variants (Amariuta et al., 2020).

In the ASRI GWAS study, the control population was not specifically selected, potentially including a proportion of individuals with AxSpA or related autoinflammatory conditions. While this may have slightly reduced the study's power, the low prevalence of AxSpA makes this effect negligible.

The ASRI data set used genetic data obtained from both sequencing (controls) and Chip-based (patients) sources, which may have led to occasional residual genotyping errors or strand flips that could not be resolved through quality control. We therefore filtered variants of unexpectedly high statistical significance within the ASRI GWAS when compared to the larger IGAS GWAS. Although this measure minimized the possibility of genotyping irregularities in the ASRI study which may lead to false findings, some residual P-value inflation cannot be ruled out.

Chapter 5

Analysis of peripheral blood mononuclear cell gene expression profiles using a single cell quantitative RNA sequencing approach

5. Analysis of peripheral blood mononuclear cell gene expression profiles using a single cell quantitative RNA sequencing approach.

5.1 Introduction

In Chapters 3 and 4, we introduced and comprehensively examined genetic risk across the human genome, including a meta-analysis of disease associated variants in AxSpA and the more nuanced heritable risk linked to specific presentations of the disease. However, as we observed when examining eQTL associations of the novel SNPs reported in our meta-analysis study, many genetic variants do not functionally affect their neighbouring genes but may have longer range interactions. This is a phenomenon which has previously been reported in numerous other studies, where it has been suggested that over half of pathologically relevant SNPs may not affect their nearest genes (Zhu et al., 2016, Cortes et al., 2013b).

Thus, evidence is required in order to extrapolate from findings of associated variants to identifying which genes they influence and how these disease-causing variants impact downstream molecular pathways. To bridge this gap, the next logical step is to investigate the expression patterns of genes within disease relevant cell types. The objective of this type of exploration is to establish which genes and molecular pathways exhibit differential behaviour between healthy and diseased states and provide insights into how these differences may be related to the genetic variability observed in GWAS.

Some advances in this approach have recently been made using bulk RNA quantitative sequencing, with one study reporting on gene expression in several important cell types in AxSpA disease, including CD4+ T cells, CD8+ T cells and monocytes (Brown et al., 2023). Amongst their findings, they identified a total of 35 differentially expressed genes which were located within 500kb of previously reported GWAS loci. Many more differentially expressed genes were not directly linked with GWAS variants, such as the immunoglobulin superfamily member CD83, which was up-regulated in the monocytes of AxSpA patients. Although the function of CD83 has not yet been decisively established, it is putatively associated with inflammatory disorders and its inhibition has proven an effective treatment in mouse autoimmune disease models (Li et al., 2019). Findings such as these crystallise the importance of gene expression analysis as a complementary methodology to GWAS.

Although bulk RNA sequencing offers robust data for well-defined cell types, its limitations become evident when trying to capture the granularity of cellular diversity and analyse cryptic

cell subsets which can be identified only through their unique gene expression profiles. A breakthrough in addressing this limitation comes from novel single-cell approaches, which enable the comprehensive characterization of individual cells. This innovation empowers researchers to delve into the intricate heterogeneity of immune cells in complex diseases (Ginhoux et al., 2022). Thus, single-cell solutions represent the natural progression in the exploration of molecular pathophysiology.

Some work on single cell RNA sequencing analysis in AxSpA patients has already been reported. In one recent study, AxSpA patients showed increased expression of several genes in CD8⁺ T cells, including *FOS*, *JUNB*, *TNF* and *NFKB*, all of which are activated by stimulatory pathways, including TNF (Xu et al., 2021, Koo et al., 1992, Boyce et al., 2005). Another study found reduced numbers of NK cells in AxSpA patients, with these cells expressing lower levels of cytotoxic molecules, including granzymes (*GZMA*, *GZMB* and *GZMM*) and NK cell receptors, such as *KLRB1* and *NCR3* (Ren et al., 2022). These seemingly contradictory findings between T cell and NK cell activity suggest that their roles may be opposed in AxSpA disease, with NK cells in some way involved in moderating the pathological immune response which allows AxSpA to perpetuate.

The use of single cell approaches are at an early stage in AxSpA. More work needs to be done to address questions such as the relationship between clinical presentation and gene expression profiles, the potential existence of different populations of cells and overall activation status of immune cells in patients *ex-vivo*. Therefore, we performed single cell RNA sequencing analysis on the peripheral blood mononuclear cells of biological DMARD (bDMARD) naïve patients both with active disease and with long standing quiescent AxSpA, together with healthy controls. Our aim was to profile their immune systems at the cellular level and examine whether cell subsets are differentially active between these patients and controls. Furthermore, since activated cells are more representative of the disease state, we stimulated T cells harvested from peripheral blood using CD3 and CD28 to identify how activated cells may behave differently between these groups.

Here, we present our preliminary findings on the genomic expression profiles of activated and resting peripheral blood mononuclear cells by quantitative single cell RNA (qscRNA) analysis and compare these between AxSpA patients and healthy controls. We further explore the differences in cellular activities between older untreated disease and younger AxSpA patients who are bDMARD naïve.

5.2 Aims

- To identify and compare distinct cell populations in AxSpA patients and healthy controls using qscRNA analysis.
- To perform differential expression analysis between AxSpA patients and healthy controls across multiple cell subsets.
- To investigate whether differential gene expression analysis highlights molecular pathways which are disease associated in specific annotated cell subsets.

5.3 Methods

5.3.1 Patient recruitment

In this study, we enrolled two groups of patients, comprising a total of 8 biologic therapy-naïve patients with AxSpA and 4 healthy controls. Patient recruitment was conducted through two sources: individuals attending the Rheumatology Clinic at St. James's Hospital Dublin and participants registered in the ASRI study.

Eligibility criteria for patient inclusion necessitated individuals to be aged 18 or older and to have received a diagnosis of AxSpA. Diagnoses were established based on either the modified New York (mNY) criteria or the Assessment of Spondyloarthritis International Society (ASAS) criteria for AxSpA. Ethical approval for this study was granted by the Joint Research Ethics Committee overseeing St. James' and Tallaght University Hospital.

5.3.2 Patient recruitment and data collection

Data from patients who were registered with ASRI was collected during a structured interview process previously described. The data for other patients was obtained from clinical records.

The collected data included age, sex and clinical characteristics, including self-reported age of disease onset, age at diagnosis, satisfaction of the mNY criteria, and compliance with the ASAS criteria for AxSpA.

High-sensitivity C-reactive protein (hsCRP) levels were quantified using the Abbott Architect c8000 system, which is housed at Eurofins Biomnis Ireland - an ISO 15189 accredited medical

diagnostics laboratory facility. Calibration of assays and the application of controls followed established standard procedures at Eurofins Biomnis Ireland.

5.3.3 PBMC isolation and lymphocyte preparation

Samples were acquired on two separate days, with the process detailed below performed on the day of collection. The two batches of samples collected will hereafter be referred to as *Batch 1* and *Batch 2*. Both batches of samples were treated similarly throughout the experimental process.

The following was performed in the TCD Omics core facility by Makala Carroll and Dr. Conor Finlay:

Fresh EDTA anti-coagulated whole blood was mixed in equal volume with 2% dextran solution, inverted until homogenous and allowed to sediment at room temperature for 30 mins. The supernatant was transferred to a separate tube, filled to 25ml with complete RPMI (cRPMI) and centrifuged at 200 RCF for 7 mins with centrifuge brake disabled. Pellets were resuspended in 55% Percoll and layered onto 65% Percoll in a separate tube, then centrifuged at 1500 RCF for 30 mins with centrifuge brake disabled. PBMCs were transferred to a separate tube and washed twice with cRPMI, then counted. Cells were then resuspended to $2 \times 10^6/\text{ml}$.

Cells were split, with half transferred directly to the single cell isolation and sequencing pipeline and the remaining half stimulated for 36 hours using CD3/CD28 agonists.

24 well culture plates were prepared with half of the wells coated for 2 hours with anti-CD3 (200 μl of 10 $\mu\text{g}/\text{ml}$ solution) and half untreated. 1ml of each cell suspension was applied to two anti-CD3 coated and two untreated wells. 25 μl of anti-CD28 was then added to each of the anti-CD3 coated wells and plates were incubated for 36 hours at 37C, 4.5% CO₂.

Cells were harvested at 36 hours into 1ml of cRPMI or FBS, then washed twice in PBS before being resuspended in 500 μl FBS.

For FACS purification, pooled samples were resuspended in FACS buffer (2% FBS) and incubated with 10 $\mu\text{g}/\text{ml}$ of propidium iodide and immediately sorted 200,000 live single cells using a BD FACS melody. Cells were then transferred to the single cell processing pipeline, below.

5.3.4 Single cell isolation and sequencing

The following was performed by Dr. Conor Finlay of the TCD Omics core facility:

scRNA-sequencing was performed using the BD rhapsody single cell analysis system (BD Biosciences) according to the manufacturer's instructions. Briefly, cells were incubated with Human FcX (BioLegend), before labelling cells with antibodies from the Human Single Cell Sample Multiplexing Kit (BD Biosciences). Samples were washed to remove unbound antibody, pooled and stained with a 67-plex panel of AbSeq (BD Biosciences) oligo-tagged antibodies directed against cell surface antigens. After washing, cells were resuspended in BD Sample Buffer and counted using the BD Rhapsody Scanner. A total of 34,757 cells *ex-vivo* and 34,699 stimulated cells were available for single cell capture, which was performed using the BD Rhapsody Express instrument, BD Rhapsody cartridges and the BD Cartridge Reagent Kit (original version). 30,805 *ex-vivo* and 30,936 stimulated cells with single cell capture beads were captured on the cartridges prior to cell lysis. After cell lysis, an on-bead cDNA reaction was performed using the BD Rhapsody cDNA kit. cDNA libraries were prepared using the BD Rhapsody Whole Transcriptome Analysis library preparation kit. Final sequencing libraries were assayed using a Tapestation 4200 (Agilent) with D100 screen tape and quantified using a QuBit dsDNA HS Kit (ThermoFisher). cDNA libraries were diluted to equal molarity and pooled at the following ratio RNA (62.5%), AbSeq (31.25%) and Sample Tag (6.25%). Sequencing was performed using an S2 100 cycle Kit on a Novaseq 6000 (with a final sequencing concentration of 300 pM with 20% PhiX spike in (Illumina)). Paired end sequencing was performed with the following cycle allotment: R1 64, R2 58, i7 8, resulting in 3.4 billion quality reads aligning to cells. Bioinformatics was performed using the BD Rhapsody WTA Analysis pipeline (Revision 11). Alignment was performed to human genomic sequence and annotation (GRCh38, Release 29) and a whitelist of AbSeq oligo sequences. Amongst the *ex-vivo* cell sample, a total of 30,805 cells were identified, of which 224 had an undetermined sample tag call and 5,286 cells were identified as doublet. In our cultured cells which were stimulated with CD3 and CD28, a total of 30,805 cells were identified, of which 5,003 were doublet and 238 were undetermined. All doublet and undetermined cells were removed from analysis. The remaining 25,295 *ex-vivo* and 25,564 stimulated cells were included in the study.

5.3.4 Single cell sequencing data processing

The following was performed by Dr. Conor Finlay of the TCD Omics core facility:

The counts matrix and sample tag calling table were imported into R (v4.3.1) and analysis performed in the Seurat (Satija et al., 2015) package environment (v4.3.0.1) or using Bioconductor (duVerle et al., 2016) packages (v3.17). We filtered out genes that appeared in less than 3 cells, retaining 28475 genes. We filtered out cells that had less than 100 or more than 2000 unique genes, less than 7% or more than 30% of reads aligning to mitochondrial genes, retaining 24,487 *ex-vivo* and 25,432 stimulated cells. Data was normalized using the NormalizeData function in Seurat. 2,000 Variable features were identified using the Seurat function FindVariableFeatures. Data scaled using the ScaleData function was used to produce a PCA. The first 30 PCs were used to generate a UMAP (Uniform Manifold Approximation and Projection, a k-nearest neighbours graph using the FindNeighbours function. Clusters were identified using the FindClusters (arguments: resolution = 0.5, graph.name = "RNA_snn"). Clusters were annotated manually or automatically using the CIPR R package. Differential gene expression was performed using the FindAllMarkers() function. For differential gene expression of AxSpA patients vs healthy controls we performed pseudo-bulk analysis using the EdgeR (Robinson et al., 2010) package, first removing sex chromosome genes to avoid sex bias. Trajectory analysis was performed using the slingshot R package, which generated a pseudotime variable.

5.3.5 Differential expression and geneset enrichment analysis

The following was performed by Dr. Deepak Bharti:

The DEG (differentially expressed genes) analysis was conducted over each cell type cluster. DEG analysis was performed using DESeq2 v1.38.0 and R v.4.2.2. We selected SNPs with Log₂ fold change less than -0.25 or greater than 0.25 with statistical significance $P < 0.05$, after accounting for false discovery using Benjamini-Hochberg methodology. The DEG were plotted using EnhancedVolcano v.1.16.0 in R stats v. 4.1.3.

Geneset Enrichment Analysis was performed using differentially expressed genes obtained from DESeq2, in each cell type. GSEA performed using ClusterProfiler v.4.6.2. (gseKEGG function using default parameters). The enrichment plot was created using enrichplot v.1.18.4 (function dotplot)

Differentially expressed genes were cross referenced with AxSpA SNP associations using Haploreg v4.2 (Ward and Kellis, 2012), based on an associated SNP reference set (Brown and Wordsworth, 2017). Genes were considered to be associated where linkage disequilibrium (LD) $r^2 \geq 0.4$.

5.4 Results

A total of 12 individuals (8 patients and 4 healthy controls) were included in the study. A summary of the demographics of these participants is presented in Table 5.1.

Table 5.1: Patients included in study.

ID	Age, years	Sex	mNY ¹	Age at onset, years	Age at diagnosis, years	CRP, mg/L
Patients						
AS1	74	M	yes	23	35	7.5
AS2	68	M	yes	31	41	0.4
AS3	68	M	yes	28	45	0.9
AS4	38	M	no	24	24	1.6
AS5	33	M	no	19	25	1.3
AS6	29	M	yes	21	27	5.9
AS7	21	M	no	19	19	1.2
AS8	37	M	yes	27	36	6.1
Controls						
C1	42	M	--	--	--	0.4
C2	24	M	--	--	--	0.3
C3	40	M	--	--	--	4.6
C4	60	M	--	--	--	0.7

¹Modified New York criteria for AS.

A report of cell viability from FACS flow is presented in Table 5.2. Notably, cell viability was lower in Batch 1 (median 52%, IQR 45.1 - 64.5%) than in Batch 2 (median 72.9%, IQR 65.3 – 78.1%) for cultured cells stimulated with CD3 and CD28 (P=0.013).

Table 5.2: PBMC viability ex-vivo and cultured cell viability following 36 hour stimulation with CD3 and CD28.

ID	Batch ^a	Ex-vivo total conc. (cells/ml)	Ex-vivo viable conc. (cells/ml)	Ex-vivo viability (%)	Stimulated total conc. (cells/ml) ^b	Stimulated viable conc. (cells/ml) ^b	Stimulated viability (%) ^b
Patients							
AS1	1	8,400,000	7,300,000	86.9	4,500,000	2,063,000	45.8
AS2	1	4,400,000	3,600,000	81.8	3,444,000	1,778,000	51.6
AS3	1	1,833,000	1,633,000	89.1	1,444,000	1,000,000	69.3
AS4	1	5,700,000	5,000,000	87.7	4,667,000	2,444,000	52.4
AS5	1	3,000,000	3,000,000	100.0	3,000,000	1,889,000	63.0
AS6	2	2,667,000	2,467,000	92.5	3,778,000	3,222,000	85.3
AS7	2	4,700,000	4,100,000	87.2	5,778,000	3,778,000	65.4
AS8	2	4,600,000	4,000,000	87.0	6,667,000	4,333,000	65.0
Controls							
C1	1	4,400,000	3,700,000	84.1	7,778,000	3,333,000	42.9
C2	2	3,300,000	3,000,000	90.9	8,556,000	6,333,000	74.0
C3	2	4,500,000	4,200,000	93.3	8,222,000	6,222,000	75.7
C4	2	5,300,000	5,100,000	96.2	5,111,000	3,667,000	71.7

^aBatch 1 processed 07/12/2022, Batch 2 processed 01/03/2023.

^bFACS sorted T-cells following 36 hour stimulation.

In total, 78.74% of *ex-vivo* and 82.24% of stimulated cell sequence reads were uniquely aligned to the reference human genome (GRCh 38/hg38). A summary of sequence alignment events is presented in Table 5.3.

Table 5.3: Sequence read quality summary

General	Ex-vivo	Stimulated
Total Reads in FASTQ	1,662,765,744	1,742,668,124
Reads passing quality filter, %	96.44	96.93
Cellular reads uniquely aligned, %	78.74	82.24
mRNA reads per cell (mean)	21,543	24,732
AbSeq reads per cell (mean)	11,556	12,819
Sample tag reads per cell (mean)	694	1,404
Sequence quality		
Read pair overlap, %	0.54	0.21
Reads too short (<40bp), %	0.51	0.19
Reads with low base quality, %	0.43	0.26
Reads with high single nucleotide frequency, %	3.24	2.81
Total fraction of reads filtered out, %	3.56	3.07
Library quality		
Total filtered reads, n	1,603,617,277	1,689,081,945
R2 bases with quality score >30, %	92.08	92.13
Read pairs containing valid cell label, %	88.59	90.56
Total fraction of reads uniquely aligned, %	78.74	82.84
Alignment categories		
Aligned cellular reads, n	1,420,286,248	1,529,355,570
Annotated transcriptome, %	48.3	52.62
Introns, %	5.48	2.23
Intergenic regions, %	3.8	2.65
Antisense, %	1.55	1.01
Non-unique, %	1.88	1.64
Ambiguous, %	1.37	1.66
AbSeq, %	32.38	31.54
Sample tag, %	2.85	4.44
Unaligned, %	2.5	2.21

We performed quality control analysis of mapped RNA reads to verify reliability of cellular gene counts across individuals and to assess the average read counts per gene. Additionally, we investigated the proportion of cellular RNA attributed to mitochondrial reads in each individual (read counts above 35% are suggestive of cell death). The results of these analyses are graphically depicted in Fig. 5.1.

We found consistent read counts across individuals, with median mitochondrial reads remaining below 25% for all subjects. Notably, we observed that cell populations stimulated with CD3 and CD28 exhibited a higher average gene expression. This heightened gene expression was associated with a reduction in the proportion of mitochondrial reads relative to total mapped reads in CD3 and CD28 stimulated cells.

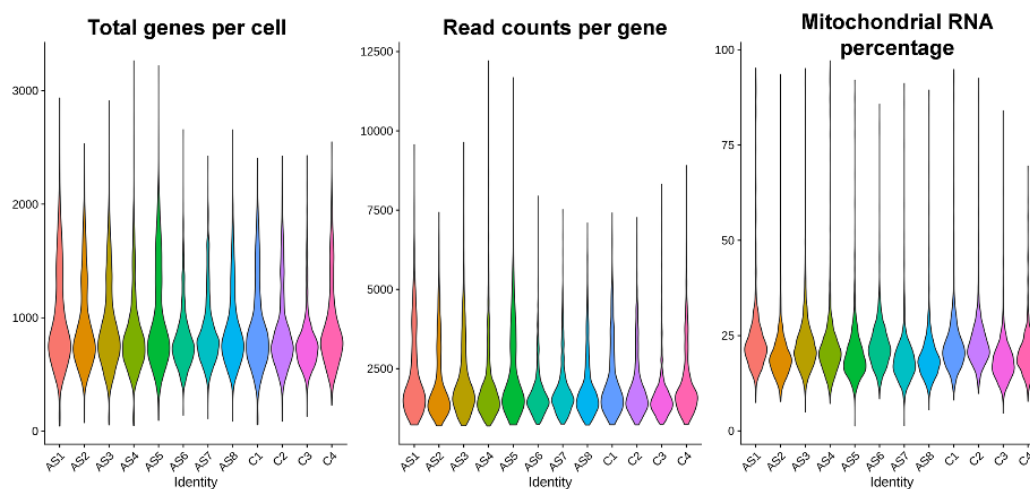
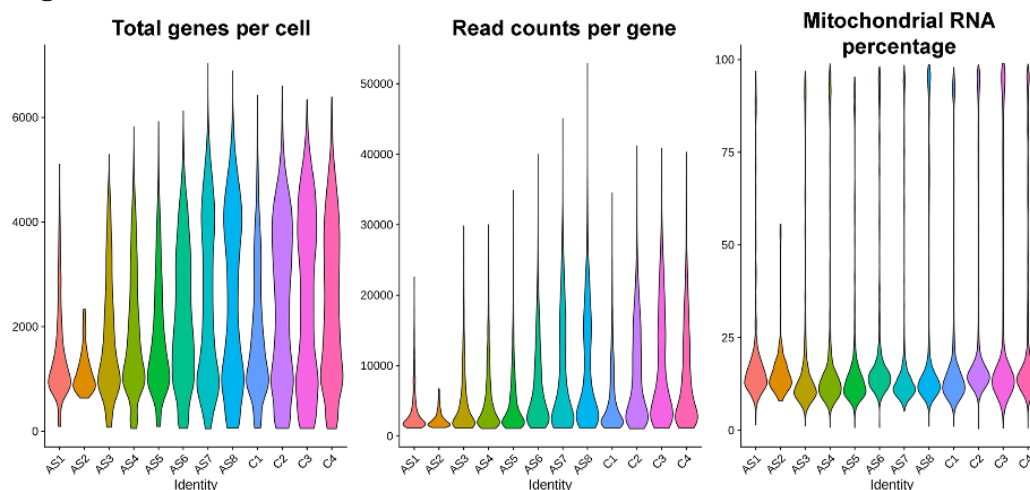
Fig. 5.1A**Ex-vivo Sample Quality****Fig. 5.1B****CD3/CD28 Stimulated Sample Quality****Fig. 5.1:** Quality characteristics of RNA expression in cells *ex-vivo*.**Fig. 5.1A:** *Ex-vivo* cells**Fig 5.1B:** Cells stimulated with CD3 and CD28.

Table 5.4 provides a summary of putative cell calls for both *ex-vivo* and stimulated cells. Notably, within *Batch 1*, samples AS1, AS2 and AS3 exhibited lower coverage per cell and lower call rates specifically within the stimulated cells. This was not reflected amongst *Batch 2* samples. These findings strongly suggest the presence of a pre-analytical batch effect, most likely stemming from cell death during processing of the first batch of stimulated cells.

Table 5.4: Putative cells called by sample tag.

Ex vivo						Stimulated			
Sample ID	Batch	Raw Reads, n	Cells called, n	Raw reads in called cells, n	Mean reads per called cell, n	Raw Reads, n	Cells called, n	Raw reads in called cells, n	Mean reads per called cell, n
Patients									
AS1	1	4,292,878	1,203	2,129,819	1,770.42	1233637	691	566,510	819.84
AS2	1	4,275,971	1,415	1,721,173	1,216.38	202,111	132	84,143	637.45
AS3	1	5,020,709	4,749	2,746,756	578.39	1,290,887	1,293	690,365	533.92
AS4	1	3,907,319	3,136	2,332,773	743.87	7,406,827	3,622	4,288,302	1183.96
AS5	1	3,846,509	1,547	2,250,078	1,454.48	2,700,288	1,417	1,525,406	1076.50
AS6	2	2,783,367	2,159	1,589,538	736.24	7,387,533	2,590	4,794,791	1851.27
AS7	2	968,062	1,269	508,842	400.98	2,266,651	1,629	1,458,521	895.35
AS8	2	4,521,387	2,547	2,091,647	821.22	8,134,149	2,337	5,413,163	2316.29
Controls									
C1	1	2,567,554	1,566	1,361,724	869.56	4,719,414	2,699	2,634,218	976.00
C2	2	4,375,639	2,398	2,687,202	1,120.60	14,754,704	3,733	9,881,999	2647.20
C3	2	2,720,708	2,397	1,404,967	586.14	10,640,980	2,372	7,232,838	3049.26
C4	2	1,181,920	909	543,291	597.68	7,180,661	3,049	4,688,859	1537.84

5.4.1 Annotation of *ex-vivo* cell subsets

A summary of immune cell subtypes as reported by the BD Rhapsody Sequence Analysis Pipeline V1.12.1 is presented in Fig. 5.2. The total cell counts per individual between AxSpA and healthy controls were approximately equal (Fig. 5.3, $P=0.457$). A detailed breakdown of cell counts and the proportions of each cell type is presented in Table 5.5.

We identified a total of 14 distinct cell clusters in the total cell population, which were annotated manually by Dr. Conor Finlay. Cell subset proportions in the *ex-vivo* samples were approximately equal to expected reference ranges (Yi et al., 2019). Amongst the cell clusters, we identified an unexpected population of cells clustered near to CD4 and CD8 positive lymphocytes. These cells exhibited positive surface expression for CD4 and CD5 (see Supplementary Material Fig. S5.2), implying the presence of a T-lymphocyte subpopulation with an unknown differentiation pattern. Intriguingly, no other antibody binding was uniquely associated with this population. In light of these findings, we have labelled this subset '*aberrant T cells*'.

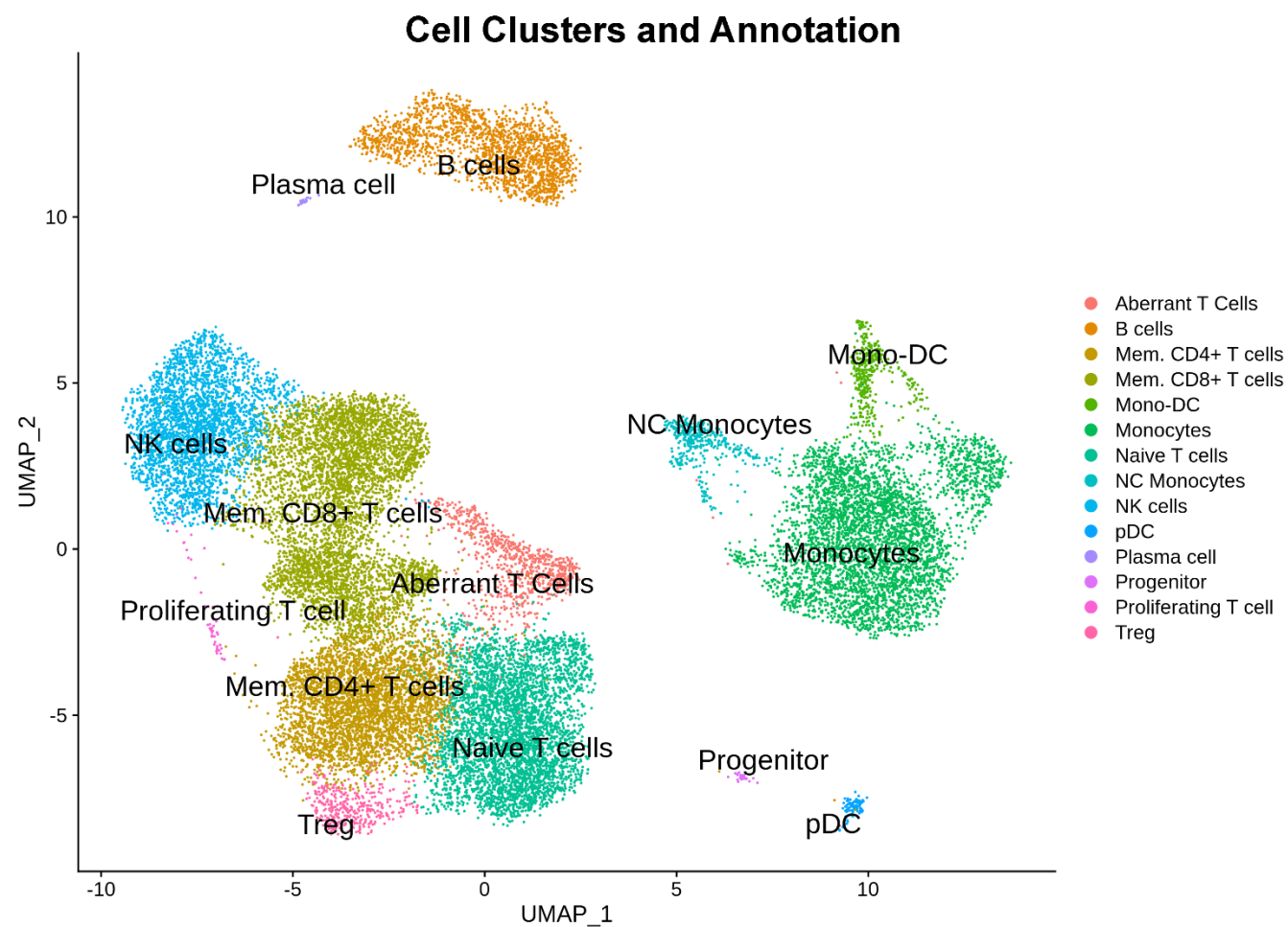


Fig. 5.2: *Ex-vivo* cells from combined patients and controls, clustered by gene expression and annotated.

Table 5.5: Cell types determined by PCA of UMAP by sample ex-vivo.

Sample ID	Total cells	Naive T cells	Memory CD4+ T cells	Memory CD8+ T cells	T-regulatory cells	Proliferating T cell	Aberrant T cells	B cells	Plasma cells	NK cells	Monocytes	Monocytes (non-classical)	Monocyte dendritic cells	Plasmacytoid dendritic cells	Progenitor cells
Patients															
AS1	1142	109 (9.5)	116 (10.2)	198 (17.3)	4 (0.4)	1 (0.1)	39 (3.4)	60 (5.3)	1 (0.1)	230 (20.1)	299 (26.2)	36 (3.2)	31 (2.7)	16 (1.4)	2 (0.2)
AS2	1384	132 (9.5)	61 (4.4)	555 (40.1)	9 (0.7)	0 (0.0)	37 (2.7)	55 (4.0)	1 (0.1)	101 (7.3)	390 (28.2)	15 (1.1)	22 (1.6)	6 (0.4)	0 (0.0)
AS3	4503	430 (9.5)	1136 (25.2)	574 (12.7)	79 (1.8)	8 (0.2)	194 (4.3)	209 (4.6)	6 (0.1)	577 (12.8)	1093 (24.3)	115 (2.6)	51 (1.1)	27 (0.6)	4 (0.1)
AS4	3001	422 (14.1)	484 (16.1)	790 (26.3)	30 (1.0)	9 (0.3)	107 (3.6)	464 (15.5)	4 (0.1)	272 (9.1)	297 (9.9)	33 (1.1)	65 (2.2)	20 (0.7)	4 (0.1)
AS5	1504	399 (26.5)	91 (6.1)	180 (12.0)	16 (1.1)	4 (0.3)	58 (3.9)	109 (7.2)	1 (0.1)	110 (7.3)	432 (28.7)	40 (2.7)	50 (3.3)	11 (0.7)	3 (0.2)
AS6	2129	224 (10.5)	303 (14.2)	973 (45.7)	26 (1.2)	6 (0.3)	73 (3.4)	152 (7.1)	0 (0.0)	128 (6.0)	189 (8.9)	33 (1.6)	16 (0.8)	4 (0.2)	2 (0.1)
AS7	1251	424 (33.9)	167 (13.3)	178 (14.2)	27 (2.2)	0 (0.0)	28 (2.2)	73 (5.8)	0 (0.0)	154 (12.3)	187 (14.9)	7 (0.6)	4 (0.3)	2 (0.2)	0 (0.0)
AS8	2498	548 (21.9)	238 (9.5)	274 (11.0)	28 (1.1)	3 (0.1)	103 (4.1)	249 (10.0)	2 (0.1)	533 (21.3)	450 (18.0)	25 (1.0)	34 (1.4)	8 (0.3)	3 (0.1)
Controls															
C1	1482	164 (11.1)	298 (20.1)	167 (11.3)	31 (2.1)	7 (0.5)	38 (2.6)	108 (7.3)	4 (0.3)	300 (20.2)	266 (17.9)	56 (3.8)	34 (2.3)	9 (0.6)	0 (0.0)
C2	2326	699 (30.1)	328 (14.1)	355 (15.3)	57 (2.5)	2 (0.1)	51 (2.2)	157 (6.7)	5 (0.2)	214 (9.2)	390 (16.8)	19 (0.8)	43 (1.8)	3 (0.1)	3 (0.1)
C3	2377	592 (24.9)	391 (16.4)	771 (32.4)	24 (1.0)	2 (0.1)	60 (2.5)	141 (5.9)	1 (0.0)	236 (9.9)	117 (4.9)	7 (0.3)	22 (0.9)	2 (0.1)	11 (0.5)
C4	890	137 (15.4)	193 (21.7)	99 (11.1)	49 (5.5)	5 (0.6)	23 (2.6)	91 (10.2)	0 (0.0)	114 (12.8)	158 (17.8)	8 (0.9)	11 (1.2)	0 (0.0)	2 (0.2)

Total cell counts per individual

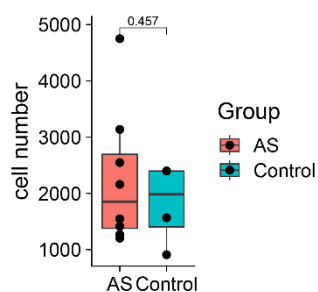


Fig. 5.3: Total cell counts by group.

We compared the proportions of each of the cell subsets detailed in Fig. 5.2 between AxSpA patients and healthy controls and also between batches, as presented in Fig. 5.4. Several cell subsets were found to be different between *Batch 1* and *Batch 2*, with these including naïve T cells and myeloid cell types. Interestingly, the most divergent subset were the *aberrant T cells*, which were increased in AxSpA patients when compared to healthy controls, although this did not reach statistical significance after accounting for false discovery ($P_{adj}=0.0714$). This specific cell type displayed elevated levels in AxSpA patients across both sample batches, suggesting a consistent trend warranting further investigation.

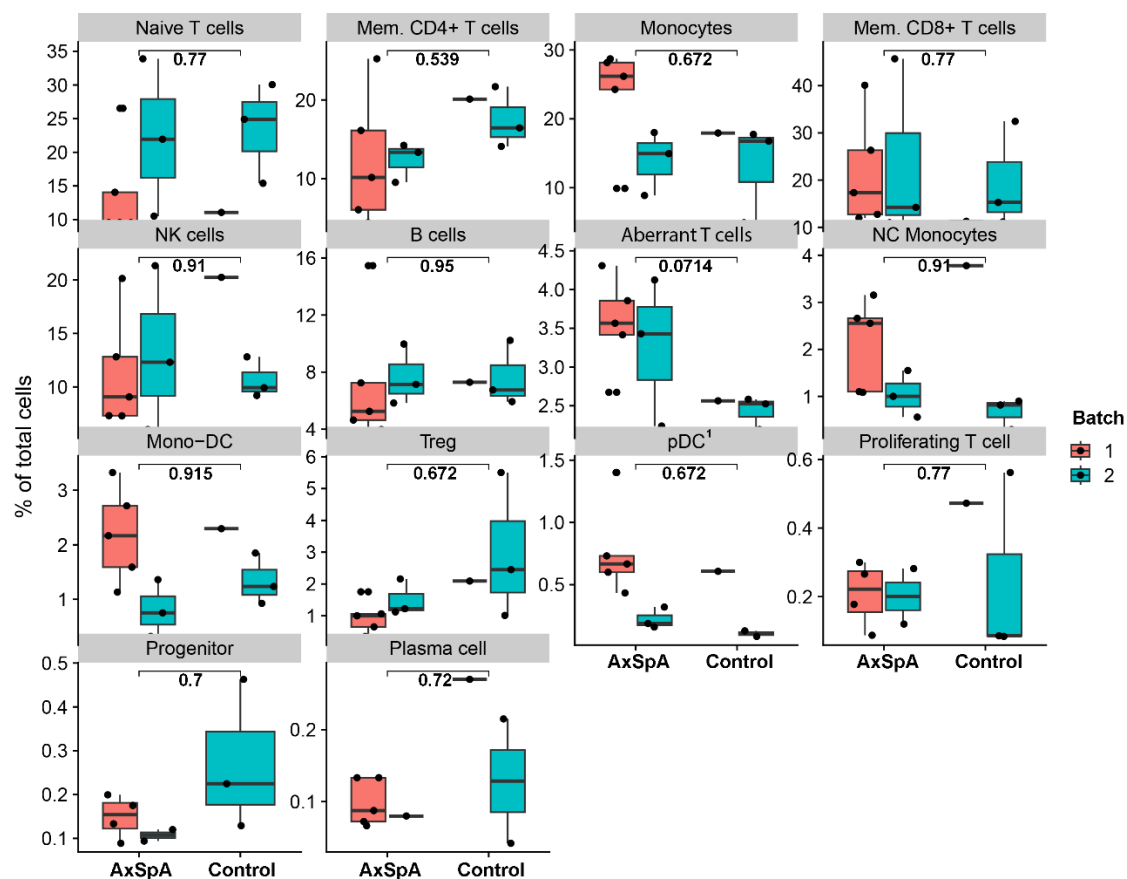


Fig. 5.4: Cell proportions subdivided by group (AxSpA and control) and batch. Statistical significance determined between AxSpA cases and controls and corrected for false discovery using Benjamini-Hochberg methodology. ¹Plasmacytoid dendritic cells.

We next performed gene expression analysis across all cell clusters, combining both patients and controls, to profile the gene expression of the major cell populations. The results of this analysis are presented in Fig. 5.5.

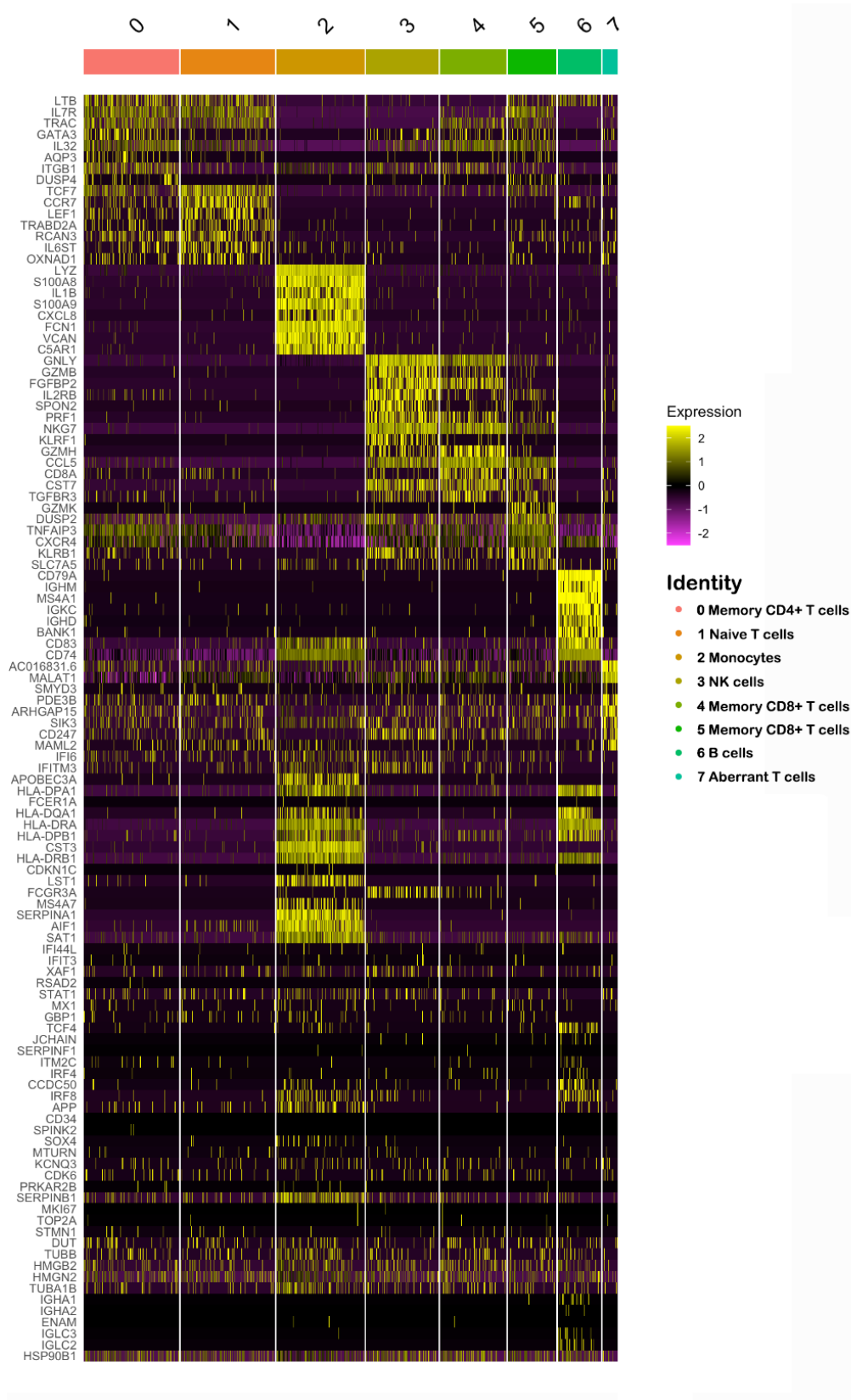


Fig. 5.5: Heatmap showing the top 110 most differentially expressed genes across the seven most populous cell clusters, and the *aberrant T cells* identified in *ex-vivo* cells. This figure represents combined patient and control data.

Amongst CD4+ T cell subset, the *aberrant T cell* cluster showed markedly increased expression of the genes *AC016831.6* and *MALAT1*, while overall expression of *PDE3B*, *MAML2* and *SIK3* and *CD247* were also all increased. Several of these genes are known to play roles in diabetes or its management. These include *MALAT1*, a lncRNA molecule associated with hyperglycaemia and some types of carcinoma (Safran et al., 2021); *PDE3B*, encoding a phosphodiesterase regulator of the ubiquitous secondary signalling molecules cAMP/cGMP which is linked with the insulin receptor signalling cascade and regulatory T cell function (Degerman et al., 2011, Wang et al., 2022b); and *AC016831.6*, which encodes a lncRNA that is up-regulated in response to metformin treatment (Conceição et al., 2022).

The other genes, namely *CD247* and *SIK3*, are associated with T cell differentiation in the thymus and regulation of mTOR signalling, respectively (Safran et al., 2021). Notably, *SIK3* is also involved in chondrocyte differentiation during skeletogenesis and has previously been associated with Spondyloepimetaphyseal Dysplasia, Krakow Type (Csukasi et al., 2018).

5.4.2 Differential expression analysis in ex-vivo cells of AxSpA patients and healthy controls

We examined the differential gene expression between AxSpA patients and healthy controls across key immune cell types, including B cells, T cells, NK cells and monocytes, all of which have previously been implicated in AxSpA (Brown et al., 2023, Ding et al., 2023). The ten most statistically significant up-regulated and down-regulated genes for each cell cluster are detailed in Tables 5.6 and 5.7. A full list of all differentially expressed genes by cell type is also available in Supplementary Material Table S5.2. Our analysis revealed differentially expressed (DE) genes within all cell clusters, with the largest number of DE genes (247) observed in monocytes, as described in Tables 5.6 and 5.7.

A consistent pattern of increased gene expression in AxSpA patients across all investigated cell types was a prominent finding (Fig. 5.6). This effect was most notably pronounced within the subset of naïve T-cells (Fig. 5.6B).

Within the HLA associated genes, we noted increased expression of *HLA-B* in B cells and T cells of AxSpA patients (Table 5.6). Conversely, a reduction in expression of *HLA-DQB1* and *HLA-C* was observed in monocytes and lymphocyte clusters, respectively.

Other genes of notable interest included *JUND*, which was enriched in all cell clusters and *MYADM*, which was enriched in all lymphocyte clusters. *JUND* is component of the activating protein 1 (AP-1) transcription factor. Intriguingly, several other up-regulated genes strongly

represented in our analysis were also components of AP-1. These included *JUNB*, *FOSB* and *ATF3*. *MYADM*, meanwhile has an established immunological role in cell adhesion and regulation of protein kinase C (Safran et al., 2021).

In contrast, only a handful of genes displayed reduced expression in AxSpA, with the ribosomal protein *RPS26* exhibiting this pattern across all cell clusters. *RPS26* is a component of the 40S ribosomal subunit and has roles in T cell survival (Safran et al., 2021, Chen et al., 2021a). A SNP in *RPS26* (rs1131017) is associated with several autoimmune diseases including Type 1 Diabetes and is associated with *RPS26* expression levels. This genomic region controls the expression of over hundreds of genes in CD4 and CD8+ T cells in trans and it has been suggested that *RPS26* abundance mediates this by altering the translational efficiency of affected mRNAs (Li et al., 2013, Kasela et al., 2017). Thus, our finding of reduced *RPS26* expression in all major cell subsets could have broad implications for gene expression in these cells.

Our findings are broadly consistent with prior research involving cell type-specific bulk RNA sequencing analysis as documented by Brown et al. in their 2023 study. In alignment with their work, we observed a significant up-regulation of *CD83* expression in monocytes ($P_{\text{adj}}=2 \times 10^{-9}$), which we additionally found in B cells ($P_{\text{adj}}=1 \times 10^{-7}$). We also corroborate increased expression of *CD69* in naïve T cells ($P_{\text{adj}}=3.2 \times 10^{-16}$), while additionally finding a similar pattern in B cells ($P_{\text{adj}}=3.9 \times 10^{-6}$) and NK cells ($P_{\text{adj}}=5.5 \times 10^{-58}$).

However, some of our findings contradict previously published single cell RNA expression data, particularly amongst components of the AP-1 transcription factor. These included *JUNB*, which was previously found to be up-regulated in CD8+ T cells of AxSpA patients was down-regulated in our study ($P_{\text{adj}}=8 \times 10^{-9}$) and was not significantly different in other cell types (Xu et al., 2021). We also did not find evidence for up-regulation of *FOS* in CD8+ T cells ($P_{\text{adj}}=0.23$). However, *FOS* expression was increased in B cells ($P_{\text{adj}}=6 \times 10^{-4}$), naïve T cells ($P_{\text{adj}}=1.1 \times 10^{-7}$), memory CD4+ T cells ($P_{\text{adj}}=2.4 \times 10^{-5}$) and NK cells ($P_{\text{adj}}=3.6 \times 10^{-26}$). This may indicate functional redundancy between different members of these transcription factor families; for example, we see up-regulation of *JUND* in our memory CD8+ T cells.

Finally, the *aberrant T cell* cluster previously introduced showed increased expression of several genes in AxSpA patients which were not observed in other cell clusters. These included *AC16831.6* ($P_{\text{adj}}=1.1 \times 10^{-5}$), *GNG2* ($P_{\text{adj}}=5.0 \times 10^{-4}$), *PPP1R16B* ($P_{\text{adj}}=1.5 \times 10^{-3}$), *SIK3* ($P_{\text{adj}}=0.0064$), *TGFB1* ($P_{\text{adj}}=0.019$) and *SYTL3* ($P_{\text{adj}}=0.015$). *RIPOR2* ($P_{\text{adj}}=0.0042$) is down regulated. This protein is a negative regulator of chemokine induced responses such as T cell migration (Rougerie et al., 2013).

Table 5.6: Top up-regulated genes by statistical significance and cell type.

Cell type	Up-regulated genes, n	Most statistically significant differentially expressed genes, gene code {log2 fold change} (P value adjusted)									
		HLA-B	FOSB	JUND	CD83	YPEL5	CD69	NFKB1	REL	EIF1	PPP1R15A
B cells	42	{0.46} (1.1E-21)	{0.89} (1.5E-16)	{0.37} (4.5E-13)	{0.47} (1.0E-07)	{0.49} (2.0E-07)	{0.50} (3.9E-06)	{0.80} (2.8E-05)	{0.50} (3.7E-05)	{0.50} (3.7E-05)	{0.48} (9.0E-05)
Naïve T cells	114	TNFAIP3 {0.84} (7.9E-55)	JUND {0.62} (4.1E-51)	KLF6 {1.06} (6.9E-46)	HLA-B {0.39} (5.8E-43)	YPEL5 {0.73} (1.5E-29)	RGCC {0.87} (2.0E-29)	RPL9 {0.28} (1.3E-27)	EIF1 {0.38} (6.3E-27)	NFKBIA {0.79} (5.4E-26)	IRF1 {0.77} (3.1E-24)
Memory CD4+ T cells	28	KLF6 {0.75} (2.5E-36)	TNFAIP3 {0.48} (9.5E-29)	JUND {0.40} (2.9E-26)	HLA-B {0.27} (9.8E-21)	MYADM {0.81} (3.8E-19)	RGCC {0.61} (7.6E-19)	YPEL5 {0.47} (2.2E-12)	HSPA5 {0.43} (1.4E-09)	TSC22D3 {0.39} (7.5E-09)	IRF1 {0.50} (2.8E-08)
Memory CD8+ T cells	20	MYADM {0.76} (6.7E-24)	JUND {0.37} (3.7E-19)	KLF6 {0.44} (1.5E-15)	NFKBIA {0.47} (1.3E-13)	RGCC {0.50} (6.1E-11)	IRF1 {0.41} (7.7E-09)	YPEL5 {0.34} (4.5E-08)	HSPA5 {0.36} (6.2E-08)	MCL1 {0.41} (1.9E-07)	TSC22D3 {0.33} (2.8E-07)
T regulatory cells	3	HLA-B {0.53} (8.4E-05)	JUND {0.66} (9.3E-04)	EIF1 {0.58} (9.3E-04)							
Aberrant T cells	13	AC016831.6 {0.75} (1.1E-05)	GNG2 {1.09} (5.0E-04)	PPP1R16B {0.93} (1.5E-03)	TNFAIP3 {0.93} (1.5E-03)	ATP2B1 {0.87} (5.2E-03)	SIK3 {0.69} (6.4E-03)	SYTL3 {0.79} (1.5E-02)	JUND {0.79} (1.5E-02)	HLA-B {0.44} (1.5E-02)	TGFB1 {0.79} (1.9E-02)
NK cells	56	NFKBIA {1.37} (8.2E-67)	JUND {0.87} (7.3E-60)	CD69 {1.45} (5.5E-58)	TNFAIP3 {0.92} (2.6E-43)	FOSB {1.37} (3.0E-39)	PPP1R15A {1.32} (3.7E-38)	TSC22D3 {0.99} (8.4E-32)	MYADM {1.18} (3.2E-31)	JUN {1.16} (3.6E-28)	KLF6 {0.84} (4.8E-28)
Monocytes	174	FOSB {0.72} (1.1E-59)	RGCC {0.93} (5.6E-45)	MAP3K8 {0.98} (2.2E-43)	IER3 {0.98} (1.1E-40)	GADD45B {1.15} (1.6E-32)	BTG1 {0.60} (1.6E-31)	HSPA5 {0.92} (5.4E-31)	PPIF {0.71} (1.1E-30)	PNP {0.96} (7.8E-30)	PHACTR1 {0.93} (8.4E-30)

P values corrected for false discovery using Benjamini-Hochberg methodology.

Table 5.7: Top down-regulated genes by statistical significance and cell type.

Cell type	Down-regulated genes, n	Most statistically significant differentially expressed genes, gene code {Log2 fold change} (P value adjusted)									
		RPS26	SELL	RCSL1	IGHM	ACTB	UBC	ZFP36L1	BCL11A	SEPT6	RHOB
B cells	10	{-0.52} (4.8E-09)	{-0.58} (1.6E-05)	{-0.47} (7.3E-04)	{-0.32} (1.0E-03)	{-0.25} (2.1E-03)	{0.33} (8.1E-03)	{-0.35} (1.2E-02)	{-0.38} (3.4E-02)	{-0.33} (3.9E-02)	{-0.40} (3.9E-02)
Naïve T cells	1	RPS26 {-0.69} (1.2E-44)									
Memory CD4+ T cells	18	MT-ND3 {-0.29} (7.2E-90)	RPS26 {-1.01} (5.4E-62)	HLA-C {-0.29} (3.0E-16)	ACTB {-0.34} (2.6E-14)	PTPRC {-0.30} (6.3E-07)	ETS1 {-0.30} (1.1E-05)	TCF7 {-0.29} (6.2E-05)	SELL {-0.38} (1.3E-04)	TXNIP {-0.29} (1.9E-04)	YWHAB {-0.26} (3.1E-04)
Memory CD8+ T cells	27	RPS26 {-1.19} (5.2E-111)	HLA-C {-0.54} (4.9E-76)	IL32 {-0.31} (1.2E-11)	JUNB {-0.35} (8.0E-09)	CAST {-0.44} (4.6E-08)	NSA2 {-0.39} (8.3E-08)	RPS16 {-0.18} (3.2E-07)	IL7R {-0.39} (7.1E-06)	LCP1 {-0.28} (1.2E-05)	TRBC2 {-0.27} (1.3E-05)
T regulatory cells	1	RPS26 {-0.82} (9.3E-04)									
Aberrant T cells	3	MT-ND3 {-0.26} (1.1E-06)	HLA-C {-0.50} (2.6E-02)	RIPOR2 {-0.54} (4.2E-02)							
NK cells	51	MT-ND3 {-0.29} (3.3E-51)	HLA-C {-0.50} (9.2E-45)	FOSB {1.37} (3.0E-39)	ACTB {-0.42} (1.8E-19)	GNLY {-0.34} (9.2E-14)	PRF1 {-0.41} (8.6E-10)	CAP1 {-0.55} (4.7E-09)	LCP1 {-0.38} (3.0E-08)	CAST {-0.47} (4.4E-08)	YWHAB {-0.42} (6.4E-08)
Monocytes	73	TXNIP {-1.15} (1.0E-59)	LYZ {-0.40} (1.1E-41)	HLA-DQB1 {-0.70} (1.1E-27)	HLA-A {-0.30} (6.5E-26)	RPS26 {-0.57} (2.3E-25)	HLA-C {-0.33} (2.6E-25)	SMAP2 {-0.66} (2.7E-23)	NCF1 {-0.80} (1.2E-19)	PLEK {-0.41} (4.5E-18)	CYBB {-0.45} (4.3E-17)

P values corrected for false discovery using Benjamini-Hochberg methodology.

Fig. 5.6A

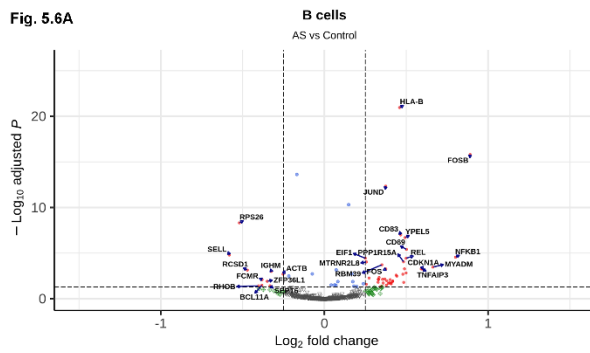


Fig. 5.6B

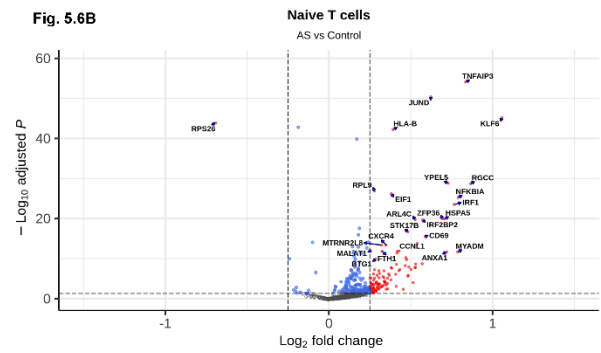


Fig. 5.6C

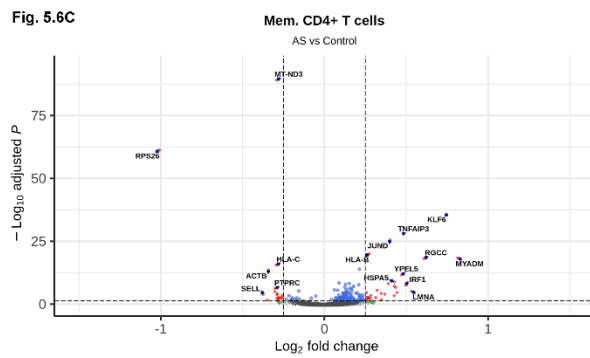


Fig. 5.6D

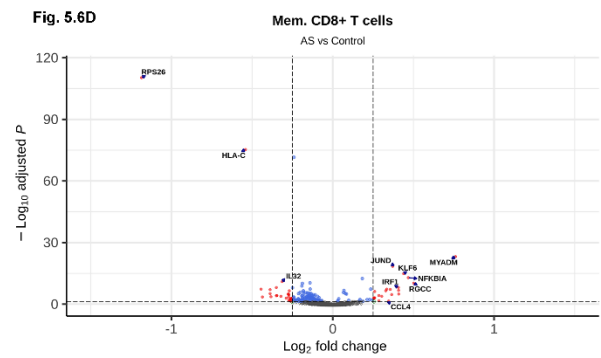


Fig. 5.6E

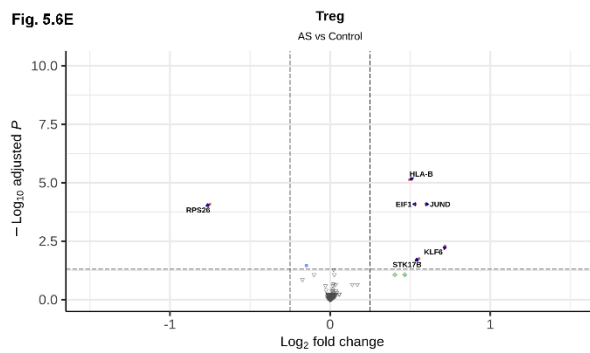


Fig. 5.6F

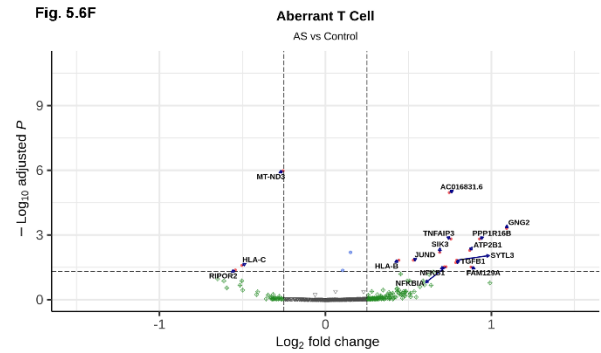


Fig. 5.6G

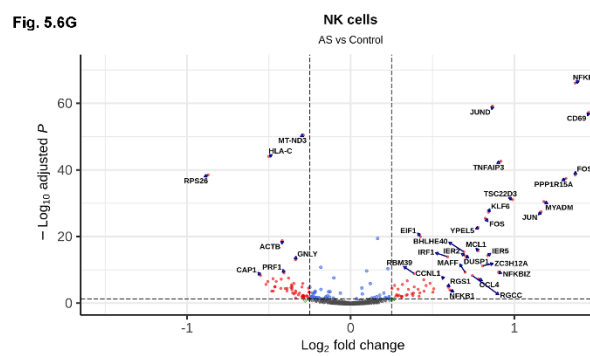
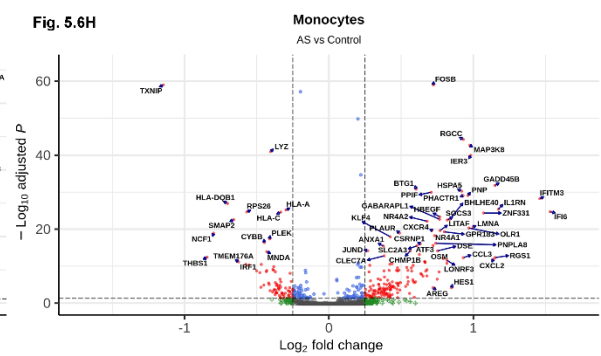


Fig. 5.6H



▽ NS ◆ Log₂ FC ● padj-value ● padj-value and log₂ FC

log₂ FC cutoff: 0.25; p-adj-value cutoff: 0.05 Total = 12435variables

Fig. 5.6: Differential expression of genes by annotated cell cluster.

Finally, we examined differentially expressed genes for overlap with previously reported AxSpA disease associated SNPs which were in LD of $r^2 \geq 0.4$ (Brown and Wordsworth, 2017). We identified a total of 12 genes correlated between DE analysis and GWAS using this methodology.

Amongst the correlated genes, we noted several genes related to the TNF- α /NF- κ B pathway which were up-regulated across multiple cell clusters. These genes included *TNFAIP3*, *NFKBIA* and *NFKB1*. All of these genes play important roles in the regulation of gene expression, particularly in inflammatory processes.

TNFAIP3 encodes a member of the zinc finger family of proteins (aka A20), which is rapidly induced by TNF. A20 is an E3 ubiquitin ligase which acts as a negative regulator of NF- κ B and has been implicated in the control of apoptosis, inflammation and autoimmunity (Malynn and Ma, 2019). Deficiency leads to spontaneous activation of many immune cell types and perinatal death in mice through hyper immune responses in multiple organs (Lee et al., 2000). SNPs in *TNFAIP3* are associated with multiple conditions of autoimmune aetiology, including SLE, RA, IBD, psoriasis, coeliac disease, Sjogren syndrome and Behçet's disease (Huang et al., 2017).

NFKB1 encodes the NF- κ B transcription factor P50;P105 subunit, which plays critical roles across all nucleated cell types. Dysregulation of *NFKB1* expression has been linked with a broad range of auto-immunity including, for example Crohn's disease and primary biliary cholangitis (PBC) (Sudhakar et al., 2022, Hitomi et al., 2019a), while mutations in the gene are associated with congenital immune deficiency (Mandola et al., 2021).

NFKBIA encodes an inhibitor molecule of NF- κ B/REL complexes and is activated through the TNF- α /TNFR1 signalling pathway. Strong genetic evidence links *NFKBIA* to multiple autoimmune diseases, including Behçet's disease, Crohn's disease (Klein et al., 2004, Yenmis et al., 2015). This is also supported by functional evidence demonstrating increased *NFKBIA* expression in juvenile inflammatory arthritis (Zhang et al., 2023). Interestingly, *NFKBIA* has also been linked with intervertebral disc degeneration, together with other related molecules, which include *TNFAIP3* (Hong et al., 2021).

We also noted correlation with genes associated with T cell differentiation and function, including *RUNX3*, which was up-regulated in four cell clusters, including monocytes and T cell subsets, excluding CD8⁺ T cells. At least two SNPs nearby to *RUNX3* have been associated with AS (Wordsworth et al., 2021), and functional analysis has found that at least one of these SNPs, rs4648889, situated within the enhancer region of *RUNX3* affects the binding of multiple transcription factors to this site (Vecellio et al., 2021). Furthermore, *RUNX3* expression is increased in patients with psoriasis, while over-expression stimulates increased production of Th17 related cytokines, including IL-17 in healthy control individuals (Fu et al., 2016).

Increased expression of FOS was also noted across multiple lymphocyte subsets, as previously discussed above.

Table 5.8: Up-regulated and down-regulated genes in LD ($r^2 \geq 0.4$) of GWAS SNPs¹

Gene_ID	Monocytes	B cells	Naïve T	Memory CD4+	Memory CD8+	Treg	Aberrant T	NK
TNFAIP3	--	0.60 (5.6×10^{-4})	0.84 (7.9×10^{-55})	0.48 (9.5×10^{-29})	0.26 (7.0×10^{-7})	--	0.76 (1.5×10^{-3})	0.92 (2.6×10^{-43})
NFKBIA	--	0.33 (7.5×10^{-3})	0.79 (5.4×10^{-26})	0.43 (7.1×10^{-8})	0.47 (1.3×10^{-13})	--	0.73 (3.0×10^{-2})	1.37 (8.2×10^{-67})
RUNX3	0.64 (1.7×10^{-9})	--	0.27 (3.3×10^{-3})	0.26 (2.1×10^{-3})	--	--	--	--
FOS	--	0.38 (5.6×10^{-4})	0.34 (1.1×10^{-7})	0.26 (2.4×10^{-5})	--	--	--	0.82 (3.6×10^{-26})
ETS1	--	--	--	-0.30 (1.1×10^{-5})	--	--	--	--
NFKB1	0.54 (8.4×10^{-6})	0.80 (2.8×10^{-5})	--	--	--	--	0.72 (3.0×10^{-2})	0.61 (1.1×10^{-4})
PTGER4	--	--	0.53 (1.6×10^{-8})	--	--	--	--	0.45 (2.7×10^{-4})
IKZF1	--	--	--	--	--	--	--	-0.29 (4.8×10^{-3})
NR3C1	0.43 (2.0×10^{-4})	--	--	--	--	--	--	--
AHR	0.32 (3.7×10^{-4})	--	--	--	--	--	--	--
ITGAL	-0.41 (4.3×10^{-4})	--	--	--	--	--	--	--
ZMIZ1	0.41 (9.0×10^{-5})	--	--	--	--	--	--	--

¹Reference AxSpA associated SNPs obtained from Brown and Wordsworth, 2017.
Data presented as log2 fold expression change (P-value FDR corrected).

5.4.3 Molecular pathway enrichment analysis by cell type

We used cell type-specific pathway enrichment analysis (Fig. 5.7) to identify differentially activated molecular pathways in AxSpA. Differential pathway enrichment was observed between AxSpA patients and controls across all cell subsets, except for T regulatory cells (Fig. 5.7E), for which the total available cell count was particularly low (349 cells). Thus, we excluded T regulatory cells from further analysis.

Viral response pathways, such as those associated with Covid 19, EBV and CMV, were up-regulated in AxSpA patients across all cell subsets. T cell clonotypes specific to viruses including EBV and CMV have previously been observed in AS, particularly CD8+ T cells (Wordsworth et al., 2021). Interestingly, while the general trend was consistent, the specific pathways enriched varied among cell clusters. Additionally, specific bacterial response pathways related to *Y. pestis* and *Shigella* demonstrated enrichment in memory CD4+ T cells. The fact that these pathways

are highlighted may relate to the proposed involvement of dysbiosis and and leaky gut syndrome in AS where a significant proportion of patients have been reported to suffer from sub-clinical intestinal inflammation and 10-15% with overt IBD (De Vos et al., 1989, Rizzo et al., 2017).

Conversely, the Fc gamma pathway of phagocytosis was down-regulated in multiple cell types, most notably in monocytes, a key phagocytic cell type, but also in NK cells and CD4⁺ memory T cells. Notably this did not include down-regulation of the FC γ gene itself. Similarly, pathways linked with intra-cellular pathogen response were also down-regulated in NK cells and CD4⁺ T cells. Pathways associated with cellular adhesion and extravasation were also down-regulated in monocytes and NK cells.

Notably, the IL-17 pathway displayed up-regulation in monocytes, NK cells and non-naïve T cell clusters of AxSpA patients, including both CD4⁺ and CD8⁺ memory cells, as well as the *aberrant T cell* cluster (Fig. 5.2). In CD4 and CD8 memory cells, enrichment was observed in a lesser proportion (<20%) of IL-17 pathway genes, while monocytes and the aberrant T cell cluster exhibited a more substantial proportion (>30%) of up-regulated genes from within this pathway.

Specifically, within the *aberrant T cell* cluster, B cells and NK cells of AxSpA patients, we observed enrichment of genes involved in osteoclast differentiation, while the *aberrant T cell* cluster also exhibited enrichment of chemokine gene expression.

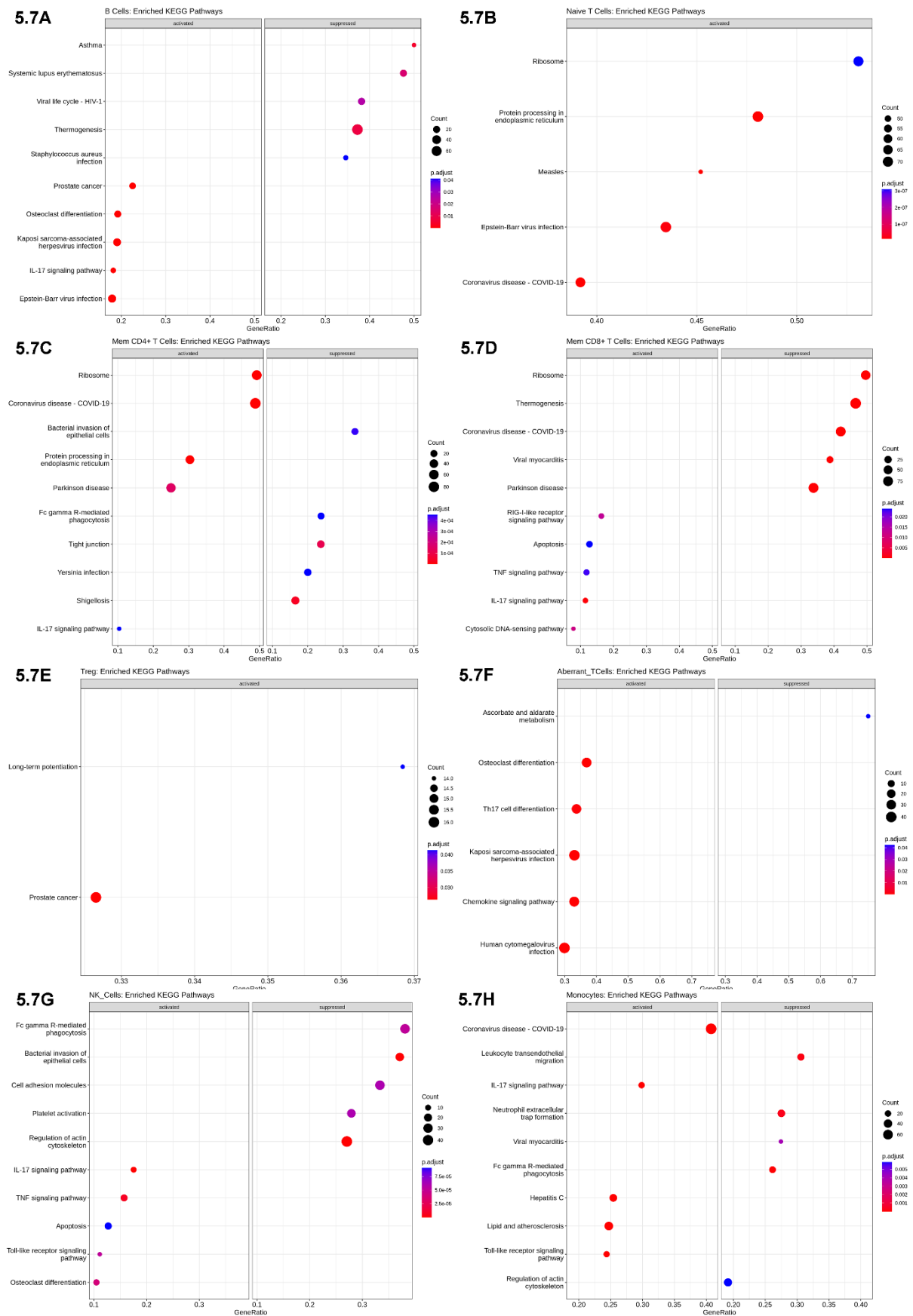


Fig. 5.7: Molecular pathway enrichment by cell type.

5.4.4 Investigation of cells stimulated with CD3 & CD28

Annotation of stimulated cells was attempted using automated methodologies trained using a combination of Blueprint Encode Data and Database Immune Cell Expression Data using the CIP-R package (performed by Deepak Bharti). However, we found that these methodologies were excessively counting dendritic cells, while simultaneously under counting other cell subsets when compared to our *ex-vivo* results (data not shown). We suspect that the activated cell expression profiles are not compatible with these reference databases. We, therefore, intend to perform differentiation manually using an integrated method, combining surface antigen and RNA expression to more accurately differentiate these cell subsets. This work has not yet been completed.

5.5 Discussion

We report here a preliminary investigation of gene expression of *ex-vivo* (unstimulated) cells from AxSpA patients and healthy controls at the single cell level. We have found supporting evidence for previous work performed using bulk RNA quantitative sequencing and further revealed differential expression (DE) within specific cell subsets. We have also identified the involvement of viral response pathways and provided further evidence for the IL-17 and antigen presentation pathways in disease pathogenesis.

Notably, we have isolated an unusual cell cluster exhibiting increased expression of disease-relevant molecular pathways, and which appear to be more numerous in AxSpA patients compared to healthy controls, although this did not reach statistical significance after FDR was accounted for ($P_{\text{adj}}=0.0714$).

This is a distinct sub-population of CD4⁺ T cells that clustered separately from the main T cell populations in our analysis, which we labelled *aberrant T cells*. Thus this cell subset merits attention as it is likely to represent an activated T cell subset which is increased in patients. Preliminary investigation through pathway enrichment revealed an increase in Th17 related genes within this subset.

As shown in Fig. 5.5, expression of a cluster of genes is diagnostic of the *aberrant T* cell subset. One of these is *SIK3*, which is involved in skeletal development. Interestingly, this gene was also differentially up-regulated in AxSpA patients when compared to healthy controls within this cell subset, as was *TGFB1*, which is another important gene involved in cellular differentiation and tissue modelling (Fig. 5.6F). These findings were supported by our pathway enrichment analysis, in which osteoclast differentiation was increased in both the *aberrant T cell* cluster and in B cells of AxSpA patients.

Through DE analysis, a prevailing trend of up-regulated gene expression was consistently observed in AxSpA patients across all cell clusters when compared to healthy controls. Notably, among the genes that exhibited down-regulation in AxSpA patients, one gene, the ribosomal protein *RPS26*, stands out with its pronounced decrease in expression across all cell types. *RPS26* is believed to play roles in the regulation of mRNA translation initiation and in T cell activation. Reduced expression of the gene has been reported in a recent transcriptome wide association study (TWAS) performed on rheumatoid arthritis (Li et al., 2023, Ni et al., 2022). Intriguingly, a smaller-scale single cell RNA sequencing study involving a single pair of monozygotic twins who developed radiographic AxSpA (rAxSpA) following SARS-CoV-2 infection highlighted a significant reduction in *RPS26* expression compared to a 45-year-old patient presenting with typical rAxSpA (Oh et al., 2022).

As discussed in Section 5.4.2 above, the *RPS26* locus has previously been associated with Type 1 diabetes (T1D), vitiligo and asthma, amongst other immune-related conditions and *RPS26* has been reported to regulate hundreds of unlinked genes in CD4+ and CD8+ immune cells (Kasela et al., 2017). The connection with T1D is interesting, given our observation of increased expression levels of multiple genes involved in diabetes and the response to diabetic therapy. None of the patients involved in the study had received a diagnosis of diabetes. Although we cannot out-rule underlying undiagnosed diabetes among possible causes, the dysregulation of *RPS26* expression provides a possible lead in explaining this finding.

While our DE analysis generally correlated with previous bulk and single cell RNA expression studies, we frequently observed distinctions in the specific genes associated with these patterns. For instance, while previous investigations identified increased expression of *JUNB* in CD8+ T cells of AS patients, we conversely found that *JUNB* expression was reduced in ASRI AxSpA patients ($P_{\text{adj}}=8 \times 10^{-9}$). However, we also found up-regulation of the closely related molecule, *JUND*. Similarly, we did not find evidence for up-regulation of *FOS* in CD8+ T cells ($P_{\text{adj}}=0.23$), yet we observed *FOS* elevation in B cells ($P_{\text{adj}}=6 \times 10^{-4}$) and NK cells ($P_{\text{adj}}=3.6 \times 10^{-26}$). Notably, a parallel pattern emerged as we identified strong up-regulation of a closely related molecule, *FOSB* across several cell types, including monocytes ($P_{\text{adj}}=1.0 \times 10^{-59}$), CD4+ T cells ($P_{\text{adj}}=2.7 \times 10^{-3}$), CD8+ T cells ($P_{\text{adj}}=8.4 \times 10^{-3}$), B cells ($P_{\text{adj}}=1.5 \times 10^{-16}$) and NK cells ($P_{\text{adj}}=3 \times 10^{-39}$). Although we have yet to explain these differences, we await results from our activated T cell data, which may help illuminate these findings. It may be that there is significant functional redundancy between these molecules. We also cannot out-rule differences in the pre-analytical treatment of cells between these studies as a cause.

Overall, a notable feature of our differential gene expression analysis was augmented expression of several transcription factors, a number of which were components of the AP-1 transcription factor. AP-1 is formed of homodimers or heterodimers of the *JUN*, *FOS* and *ATF* families, with the precise regulatory function of AP-1 dependent on its specific components (Hernandez et al., 2008). All of these genes are heavily involved in the regulation of cellular proliferation and activation (De Bosscher et al., 2003). These findings tie in with our earlier report of significant numbers of transcription factors linked with AxSpA aetiology and with recent epigenomic profiling which also identified up-regulation of AP-1 associated genes (Kenyon et al., 2022, Brown et al., 2023). There is also increased expression of both activating and regulating constituents of the potent transcription factor, NF- κ B across most cell subsets in our analysis and increased expression of another proliferative transcription factor (and proto-oncogene), *ETS1*, is seen in CD4+ T cells.

In combination with activating genes, regulatory genes such as *TNFAIP3* and *NFKBIA* are also up-regulated. While upregulation of negative feedback loops such as provided by these genes is to be expected in activated cells, these specific genes are interesting particularly since they are genetically associated with AxSpA. Both *TNFAIP3* and *NFKBIA* showed increased expression in all patient subsets apart from monocytes; whereas inflammatory disease is usually associated with reduced expression in the case of *TNFAIP3* (Malynn and Ma, 2019, Martens and van Loo, 2020). However the action of its protein, A20, is likely cell and stage specific, and dependent on the activation status of the cell. For example, while increased A20 expression increases cell death in B cells, it can have the opposite effect in T cells (Malynn and Ma, 2019). Interestingly, we are not the only group to observe these findings in *TNFAIP3* – recently published findings in bulk RNA sequencing analysis also revealed increased *TNFAIP* expression in CD4⁺ and CD8⁺ T cells of AS patients *ex-vivo* (Brown et al., 2023). This is again likely to be a complicated system and we await the analysis of the activated cells with interest, since it is possible that disease associated variants affect stimulation of the gene through nearby enhancers and a different picture may emerge in these cells (Sokhi et al., 2018). However the number of associations involving components of the NF- κ B pathway seems to indicate it plays a key role in dysregulation of the AxSpA immune environment.

Caution is advisable in making simplistic deductions about the actions of these regulatory mechanisms given their documented complexity. However, in our context, the consistent up-regulation of strongly proliferative transcription factors in some combination across practically all subsets analysed here, points to cells in an activated state in AxSpA samples compared to control samples. We may interpret this as suggesting that active disease is detectable in peripheral blood cells in these patients who are DMARD naïve, and while a very preliminary finding, this could be instructive to pursue in terms of diagnosis and improved understanding of the clinical picture.

We were interested to find up-regulation of *HLA-B* across multiple cell types. Although the association of HLA-B with AxSpA is well known, our findings present a novel dimension by demonstrating this phenomenon in the context of gene expression. The import of this observation is not clear, however increased expression of HLA-B27 has been suggested as a potential disease mechanism. Interestingly, there is variation in the expression levels of other Class I and II molecules, for example HLA-C and HLA-DQ. We aim to investigate in future whether haplotypic effects may underlie changes in HLA gene expression across the MHC. We are not aware of this observation having been reported by previous gene expression studies based on bulk sequencing of specific immune cell types. Due to the preliminary stage of the work reported here, further examination will be required to identify if this finding is affected by variables such as patient age or batch effects, and ultimately whether it is replicated in other studies.

An aim of this work was to examine whether genes encoded near GWAS hits were over-represented in our DE genes across the various subsets. We found only a small overlap as described in Section 5.4.2, including NF- κ B family members, *NFKB1*, *FOS*, *ETS1*, *PTGER4* among others. This was not unexpected since a previous study of the transcriptomics of coeliac disease by this lab showed some, but very little, activity in ex-vivo cells, a situation which changed dramatically when the cells were stimulated (Quinn et al., 2015). However, we do see some overlap with GWAS genes and DE genes between cases and controls, as discussed. We anticipate that the stimulated cells will show a much expanded overlap and will be more informative regarding the effects of GWAS associated variants, many of which we would not a priori expect to be active in unstimulated cells (Fairfax et al., 2014). Again our evidence pointing to partial activation of peripheral immune cells in patients will need to be further examined in the context of both ex-vivo and stimulated datasets.

Another gene which is genetically associated with disease is *RUNX3* and we find this is DE in naïve T cells, CD4⁺ memory T cells and monocytes. *RUNX3* is also associated with Crohn's disease, multiple sclerosis, asthma and psoriatic arthritis (Korinskaya et al., 2021). RUNX proteins are transcription factors that are active in many cell types. RUNX1 and RUNX3 may interact with each other (and many other transcription factors) and are important players in hematopoiesis, with RUNX1 having a 'master regulator' role and *RUNX3* particularly involved in the proliferation and differentiation of lymphocyte and myeloid lineages (Mevel et al., 2019). Interestingly, RUNX2 is essential for osteoblast and chondrocyte development and RUNX1 and 3 interact with it to deliver skeletal development, although their roles are minor (Bauer et al., 2015). The relationship of *RUNX3* with AS is complex with at least two independent upstream genomic variants showing evidence of regulatory involvement separately in monocytes and CD8⁺ T cells (Wordsworth et al., 2021). Interestingly, while we do see a DE effect with patient monocytes showing higher *RUNX3* expression compared to controls, we do not see this in CD8⁺ T cells, but rather in CD4⁺ and naïve T cells.

Finally, the prominent associations of viral response pathways is interesting given the recent discoveries of centrality of EBV infection to multiple sclerosis development (Bjornevik et al., 2022) and triggering effect of reoviruses in coeliac disease (Bouziat et al., 2017) among other diseases. As discussed above, the reactivity of AS patients' immune systems to various viruses has been reported previously in AS, however it has not been shown that there is a causative link between viral exposure and disease development or progression (Wordsworth et al., 2021).

Overall, we continue to find evidence for T cell involvement in AxSpA disease. Our differential pathway analysis, in particular, identified T cell functions in the form of viral response and the IL-17 pathway as key gene groups dysregulated in AxSpA patients. Interestingly, up-regulation

of the IL-17 pathway was not restricted to T cells, but also observed in monocytes, NK cells and B cells. In addition, we specifically observed aberrant regulation of the *RUNX3* gene which is a key driver of differentiation in T cells, including Th17. Several lymphocyte cell subsets also showed up-regulation of pathways which promote osteoclast differentiation, possibly linking these cells to the destructive tissue remodelling and osteoporosis which is observed in AxSpA patients.

While our preliminary findings here look promising, we have much further work to perform, including the annotation and analysis of the stimulated cell population and examining the differences between cells of younger patients with highly active disease and older patients with a more quiescent presentation. We must also consider the possible covariate effects that older patients within the study group may have on the preliminary results presented here. Nevertheless, we are hopeful that our activated T cells will confirm many of these findings and possibly reveal new insights into how these cells might behave differently in an inflammatory environment. This may prove valuable to understanding why AxSpA continues to perpetuate following its initial onset.

5.6 Limitations of this study

Due to the limitations of patient availability, particularly those who were biologic treatment naïve with established but mild AxSpA disease, several of our patients were considerably older than our control group. It is therefore possible that some of the immunological differentiation observed between patients and controls may be explained by this overall age difference. We intend to further analyse this possibility using PCA to establish what degree of differential expression can be attributed to the age of participants. However, this older patient sub-group also presents a rare opportunity to study the differences between younger participants with active AxSpA disease against older patients with overall more indolent disease.

We have not yet successfully annotated our CD3 and CD28 stimulated cell populations due to the limitation of time. These activated cell types will require integration of gene expression and antibody binding data to accurately identify, due to their unusual gene expression characteristics.

Chapter 6

Conclusion

6. Conclusion

In this thesis we aimed to investigate the relationships between genetic traits and phenotype in AxSpA disease. In our initial plan, we expected to recruit up to 2,000 patients for genotyping. However, this goal could not be reached due to the onset of the Covid 19 pandemic, which led to a complete stoppage of patient recruitment for a period of over two years, and which has not yet been re-established in much of the participating hospital network. The patients which have since been recruited have not yet been made available for genotyping. This led us to have a significantly smaller cohort than expected for genetic analysis, compared to those available for clinical comparison. The smaller size of this genotyped cohort impaired our statistical power in performing genome-wide association analysis for the various phenotypes which we investigated. We, therefore, maximised our potential to detect genetic risk associated with both disease aetiology and clinical phenotype using methodologies, such as PRS and meta-analysis.

To achieve our aim of relating clinical presentation to genotype, we explored a broad range of aspects of AxSpA, starting with the diversity of clinical presentation within the ASRI AxSpA cohort. We then advanced to investigate how genetic risk factors correlate with clinical presentation of the disease. We also used the opportunity presented by the newly genotyped cohort to identify previously undiscovered genetic associations with disease through meta-analysis with a previously published AS GWAS. Finally, we investigated how gene expression within disease-relevant immune cells derived from the peripheral blood of AxSpA patients differs from that of healthy control individuals using single cell RNA sequencing analysis and if dysregulated genes overlapped with genetically associated genes.

Throughout each of the chapters in this thesis, we confirmed published findings made on the basis of similar studies and further expanded upon their work in novel areas. We also identified unique traits within the ASRI cohort which diverge from previously published research.

6.1 The ASRI AxSpA Cohort is similar to, but not the same as other AxSpA cohorts

In our initial investigation of the clinical features of ASRI patients, we found that many of the characteristics were similar to those from other previously published studies. The cohort matched the typical male-to-female ratio of approximately 3 to 1 and approximately 85% of patients were HLA-B27 positive, very closely tracking the average reported in nearby UK (Wordsworth et al., 2021). Similarly, the proportions of patients presenting with various peripheral manifestations, such as uveitis (34.4%) and psoriasis (16.8%) were very similar to previous studies. However,

we also found distinct differences between ASRI and other AxSpA and AS cohorts. For example, women in ASRI were younger than men and, uniquely, women were more likely to have a clinical history of acute uveitis, a manifestation classically associated with male predominance in AxSpA. Interestingly, we found that the HLA-B27 allele, which was generally less frequent in women than in men in ASRI, was higher in women than in men with a history of uveitis. This could indicate that HLA-B27 may be an effective predictive marker for concomitant uveitis, particularly in women.

The under-diagnosis of women has been well documented in AS, and this could be related to their reduced axial involvement. However, ASRI is one of the few cohorts to include both radiographic and non-radiographic axial disease, in accordance with the newer ASAS criteria for AxSpA. These newer criteria may, therefore, favour the inclusion of younger women within modern AxSpA studies. Indeed, within the ASRI cohort we did not observe a significant difference in the typical delay between disease onset and diagnosis between men and women. These results suggest that the ASAS criteria may be providing a much-needed improvement in the early diagnosis of women with AxSpA. It will be interesting to learn whether these findings will be reflected in reports from other modern cohorts.

6.2 Dactylitis is an outlier amongst peripheral manifestations

Although we lacked the specific dates at which patients developed various peripheral manifestations, such as uveitis and IBD, we successfully constructed models using logistic regression to assess the likelihood of these events with AxSpA disease duration. A discernible pattern emerged, revealing that certain manifestations tend to manifest at a steady rate over time *after* the onset of initial AxSpA symptoms, while others seem to *precede* the disease or occur shortly after its onset but tend not to occur later on in disease. One of the most interesting findings was that dactylitis, which appears to manifest early in AxSpA, is robustly linked with other clinical features, including HLA-B27 negative disease, reduced axial involvement, increased disease activity and worse quality of life scores. Dactylitis was also strongly correlated with all other peripheral manifestations which we assessed. It is well documented that dactylitis bears strong links with psoriatic arthritis (PsA) (Kaeley et al., 2018). This raises the question of whether a proportion of AxSpA patients presenting with dactylitis could actually be misdiagnosed PsA with axial involvement. The close presentation of these two diseases makes differential diagnosis difficult. However, if dactylitis in suspected AxSpA cases is predictive of PsA, this may be an important finding, since the treatment of these two conditions may differ (Anne Constanze et al., 2023). To our knowledge, we are the first group to extensively investigate dactylitis in association with AxSpA and make this novel observation which we are

in the process of submitting a publication. We hope that other groups will also be encouraged to report on this under-studied manifestation. The outcome of such studies may have ramifications for the therapeutic options available to this sub-group, particularly if misdiagnosis of PsA as AxSpA proves to be an underlying issue. Regardless, the increased incidence of peripheral manifestations in this sub-group suggests that patients presenting with dactylitis may benefit from close clinical monitoring to facilitate earlier therapeutic intervention.

6.3 HLA-B27 and *ERAP1* contribute to disease phenotype

The impact of HLA-B27 on AxSpA disease has been extensively studied and our findings corroborate previous research showing that HLA-B27 is linked with earlier disease onset, increased incidence of uveitis and is inversely related to psoriasis. Contributing to the debate over whether HLA-B27 is associated with radiographic disease progression, we found no discernible relationship between HLA-B27 and radiographic positivity or in spinal mobility, as measured by BASMI.

One of the genes most robustly associated with AxSpA aside from HLA-B27 is the *ERAP1* gene. Although *ERAP1* is well established in AxSpA aetiology, we are the first group to investigate whether common genetic variants within this gene region affect disease characteristics. We initially found correlation between the well documented risk variant, rs30187 and disease characteristics, including age at disease onset and concomitant uveitis, with these associations evident only in HLA-B27 positive individuals. By investigating the gene further using haplotypes to differentiate risk alleles, we identified two additional strongly associated SNPs, rs149544 and rs7063. Rs149544 is a novel risk variant which was also associated with earlier AxSpA disease onset and episodes of uveitis in our patients. Rs7063, conversely, was strongly protective of AxSpA in homozygous individuals. This was evidenced not only through the markedly increased age at first disease onset and reduced uveitis in patients who were homozygous for the protective minor allele, but also in the Hardy-Weinburg disequilibrium which showed skewed allele frequency toward the major allele in AxSpA patients, again only in HLA-B27 positive individuals. Interestingly, a recent paper has conversely reported strong association between rs7063 and psoriasis (Marusina et al., 2023). The researchers postulate that rs7063 increases the gene expression of *ERAP2*, which acts to modulate the peptidome in favour of psoriasis onset. This is plausible since reduced expression of *ERAP1* due to the presence of rs7063 (Hanson et al., 2018) could, conceivably, result in compensatory increased expression of *ERAP2*, although this has not been comprehensively investigated.

6.4 HLA-B27 negative AxSpA remains unresolved

The strong epistasis between HLA-B27 and *ERAP1* points strongly to the importance of antigen presentation in the disease aetiology of HLA-B27 positive individuals. *ERAP1* epistasis is also found with HLA-Cw6 in psoriasis patients (Tsoi et al., 2012). However, there is no evidence for *ERAP1* association in HLA-B27 negative AxSpA. This raises the question of whether the aetiology of HLA-B27 negative AxSpA is truly similar to that in HLA-B27 positive disease. From the perspective of genetics, there is no substitute risk allele in HLA-B27 negative disease to replace the dominant effect of HLA-B27. Furthermore, GWAS to date, although limited in power, have not been able to determine any major differences in the genetic risk profile between HLA-B27 positive and negative disease beyond HLA-B27 (Cortes et al., 2013b). We also did not find evidence for a mechanism of ‘genetic risk compensation’ for HLA-B27 in B27 negative individuals using polygenic risk scores (PRS). Conversely, we found tentative evidence suggesting that additional genetic risk variants outside of the HLA region and *ERAP1* are likely to be co-opted by HLA-B27, leading to increased non-HLA genetic risk burden in HLA-B27 positive patients. It is possible that genome-wide coverage of the original Immunochip design, which was used to perform these studies, is simply missing regions of the genome where this novel genetic risk resides. However, even if this is true, we may rationally expect that if a key mechanism of disease causation (i.e. antigen presentation in the HLA) is not involved in the disease aetiology, that the genetic associations within the genomic regions already explored would differ substantially between HLA-B27 positive and negative patients, but this is not the case. Perhaps alternative avenues of investigation, such as single cell RNA sequencing may yield more answers as to what mechanism underlies HLA-B27 negative AxSpA.

6.5 Possible involvement of infectious disease and the gut microbiome in AxSpA aetiology

Despite the improved mapping of genetic risk variants in AxSpA disease, the event which triggers disease onset in genetically susceptible patients remains elusive. However, several recent studies have highlighted associations between various viral pathogens and autoimmunity. Among the most prominent of these is the association of EBV with multiple sclerosis and systemic lupus erythematosus (Robinson and Steinman, 2022, Bogers et al., 2023). Although the mechanisms of these associations are not well understood, it is thought that the viral gene, *EBNA2* hijacks risk genes which are associated with B-cell related auto-immunity (Bogers et al., 2023). During our investigation into the underlying reasons for the lack of PRS association with uveitis in AxSpA patients, we found that variants within *CCHCR1*, a gene with high disease impact located within the MHC region, exhibited increased prevalence among patients lacking the HLA-B27 allele.

CCHCR1 is of particular interest due to its known association with the immunological response to viral pathogens, notably Human Papillomavirus (HPV) (Muller and Demeret, 2014). It is also linked with other diseases, including psoriasis and type 2 diabetes (Tervaniemi et al., 2012, Brenner et al., 2020). This finding prompts the question of whether HLA-B27-negative AxSpA may have some connection to viral response, possibly even to HPV itself. Interestingly, within our pathway enrichment analysis, we identified differential expression of specific pathways, including EBV, CMV and Kaposi sarcoma associated herpesvirus within lymphocyte subsets, supporting the notion that viral infection may be an important mechanism in AxSpA pathogenesis. Of the 8 patients included in our single cell study, at least 4 were B27+, while 3 were unknown. It would therefore imply that the phenomenon is not limited to HLA-B27 negative patients given the likelihood that, at best, a low proportion of this patient sample are B27 negative. This is an avenue we are considering exploring further within the ASRI cohort.

Evidence pointing to microbial involvement is not limited to viral pathogens. In a recent systematic review, SpA patients were found to have a restricted microbiome profile, somewhat resembling that observed in IBD (Wang et al., 2022a). This ties in with the fact that a substantial proportion of AxSpA patients present with subclinical gut inflammation, while around 10% develop overt IBD (Rizzo et al., 2017). Interestingly, one of the candidate microbes identified, *Collinsella aerofaciens*, which was associated with IBS in AxSpA patients, is known to increase gut permeability and to stimulate IL-17 response by intestinal epithelial cells (Wang et al., 2022a, Chen et al., 2016). In our differential expression analysis, we found that pathways involving phagocytosis and the response to intracellular bacteria were all down-regulated in AxSpA patients within relevant cell types, including monocytes and NK cells. Also down-regulated were response genes to the bacterial invasion of epithelial cells and genes associated with leukocyte trans-endothelial migration (Chapter 5, Fig. 5.7). Though study of the microbiome still seems at an early stage, these findings collectively suggest that AxSpA patients may have gut dysbiosis and an attendant impaired immunological response in the gut with systemic effects. This represents one plausible mechanism by which AxSpA disease could be triggered in otherwise healthy individuals.

6.6 Evidence for Th17 cells in AxSpA continues to mount, but the relationship is complicated

Throughout our research, we continued to find further evidence for the involvement of Th17 cells in the pathogenesis of AxSpA. This included the discovery of novel genetic risk variants with putative roles in Th17 cell differentiation and function, including *IKZF1* and a second SNP association near to *PTGER4* (Ueno et al., 2018, Napolitani et al., 2009). This was further

supported in our single cell RNA sequencing analysis, in which we observed up-regulation of Th17 or IL-17 pathways in multiple lymphocyte subsets and in monocytes.

However, it is difficult to reconcile a role for Th17 cells with the failure of IL-23 blockers in AxSpA treatment (Baeten et al., 2018). IL-23 is a key upstream effector of Th17 cell differentiation and function, making it a rational target for IL-17 pathway inhibition. Hence, it came as a surprise that IL-23 inhibitors were ineffective in treating AxSpA although they were highly effective in the treatment of psoriasis (Reich et al., 2019), while anti-IL-17 therapeutics are effective in both. Perhaps this points to some functional disconnect between IL-23 and IL-17 in AxSpA, or that some other signalling mechanism substitutes for IL-23 in Th17 differentiation in AxSpA. However, this notion is at variance with the observation that a coding variant in the IL-23 receptor gene is one of the strongest AxSpA genetic associations identified by GWAS, meaning that IL-23 must have an important role at least in a proportion of patients.

In terms of understanding this divergence in response to treatment, it may be relevant that the documented breadth of activities of the two cytokines are different. While IL-23 is known for a quite restricted role in Th17 cell development, IL-17 has a broader range of action and is produced by a variety of lymphoid and myeloid cells, making inhibition of the latter a more likely strategy (Floss et al., 2015, Mills, 2023).

Clearly however, the molecular mechanisms that lead to IL-17 involvement in these diseases differ at the level of IL-23, and perhaps further downstream, calling into question whether Th17 cells are involved in AxSpA. This is a topic that was broached recently in a published review in which the authors concluded that Th17 cells are not the principal inflammatory cells in AxSpA pathology, but that IL-17 is still important (Zhang et al., 2022a).

Nonetheless, while it is yet uncertain as to why IL-23 inhibition has failed in AxSpA treatment, our data continues to point to the involvement of Th17 cells. Given the observations above, we suspect that the role of these cells in some way diverges between AxSpA and psoriasis. We hope that our single cell analysis of activated T cells may help shed some light on why IL-23 is not a suitable therapeutic target in AxSpA. It would be interesting to compare the expression characteristics of these activated lymphocytes acquired from AxSpA patients with similarly treated cells from psoriatic arthritis patients to establish in which subsets differences that may tie in with IL-23 effector functions are observed.

6.7 Future directions

While it may be tempting to attribute some limitations of our study to the relatively modest sample size of ASRI within the context of genetic research, it is important to recognise the unique challenge of assembling the extensive phenotypic information that has made this study possible. This endeavor is arduous to replicate on a larger scale, rendering rapid expansion of the study size an unattainable objective. Therefore, our intention is to focus our future research efforts on the current available cohort while updating our work as the sample size grows over time.

The PRS models we applied in this study were limited by the fact that they were trained on a restrictive model of AS patients only. This resulted in limited ability to accurately predict patients who would develop peripheral manifestations, based on their personalised score. The reasons behind this were exemplified by our findings in uveitis, where genetic risk variants of high impact were associated in opposite directionality between AxSpA with and without uveitis, causing PRS to offer no value in predicting uveitis whatsoever. However, some novel solutions to this using mixed model approaches may yield interesting results. Additionally, while PRS have powerful predictive capability, they have been criticised for being a blunt tool - although PRS are a good, amalgamated score to assess the genetic burden contributing to disease development, they are difficult to reverse engineer to identify causative variants in any individual or group. This is particularly the case in diseases with many identified disease associated loci contributing low levels of risk, only a subset of which are required to confer susceptibility in a given individual and which varies from one individual to the next. However, a new approach to PRS allows for the selection of variants, based on molecular pathways (Choi et al., 2023). This feature could open new opportunities to apply PRS across related manifestations to identify which pathways are important in the development of comorbidity.

A primary constraint of GWAS methodology lies in the assumption that genetic susceptibility primarily stems from common occurring variants (common disease common variant hypothesis). Although common variants clearly do contribute to susceptibility, it is acknowledged that a substantial portion, up to 70%, of the aggregate genetic predisposition to AxSpA remains elusive (Ellinghaus et al., 2016). Within this unexplained genetic component, it is thought that approximately half can be accounted for by SNP heritability (Ellinghaus et al., 2016), leaving the remainder comprising other types of variation, such as rare variants of notable effect size that evade detection via standard SNP array platforms. The issue of missing heritability attributable to common variants may be partially addressed using alternative chip designs. The Illumina ImmunoChip was favoured in previous AxSpA GWAS due to its targeted coverage of recognised immunologically associated regions of the genome. However, the ImmunoChip design was largely predicated upon prior GWAS findings, thereby restricting their coverage to a small proportion of the genome with established immunological relevance (Cortes and Brown, 2011).

It is conceivable that unexplored genomic territories may yet contain a portion of the latent genetic risk associated with AxSpA. Chips such as the Illumina Omnichip which provide low resolution genome-wide coverage may be able to identify currently elusive genetic signals. However, these platforms will require high-powered studies performed in collaboration with other AxSpA research cohorts with the further need for subsequent fine-mapping endeavours to pinpoint associated loci with greater precision. Ultimately, all chip designs are limited by their coverage. This dilemma highlights the imperative for a paradigm shift towards comprehensive genome sequencing methodologies, although cost and throughput limitations are current challenges to this approach.

The exploration of rare variant contributions to AxSpA remains an area of nascent inquiry, with sparse investigations conducted thus far. The conventional case-control study design, a cornerstone of GWAS, is inherently ill-suited toward such investigations owing to its constrained statistical power (McCarthy et al., 2008). In contrast, family-based methodologies offer a range of advantages, chief among them being causative variant co-segregation with the phenotypic trait of interest, thereby streamlining variant filtration processes (Costantino and Breban, 2023). Moreover, family-based studies are less susceptible to confounding effects stemming from population stratification.

Recent family-based investigations have typically employed co-segregation analysis and genomic sequencing to pinpoint rare variants of potential significance, although some studies have opted for SNP chip-based approaches. Notable discoveries using such approaches include the identification of non-synonymous mutations within immunologically relevant genes such as insulin receptor substrate 1 and MEFV Innate Immunity Regulator, Pyrin (Rong et al., 2015, Li et al., 2019).

Despite the advantages conferred by family-based designs in uncovering rare variants implicated in AxSpA, the validation of these findings through replication studies has posed considerable challenges (Costantino and Breban, 2023). This difficulty may be attributed, at least in part, to incomplete penetrance exhibited by these rare variants, compounded by the challenge of identifying additional familial cohorts carrying the same genetic aberrations. There is a clear need for more extensive family-based cohorts to effectively unravel further the contribution of rare variants in AxSpA and confirm the findings made by previous studies.

Within the ASRI cohort, several family groups were identified during the initial quality control screening of genetic information. However, these families were small (consisting of 3 or fewer members). Nonetheless, the ongoing recruitment efforts within the ASRI study should allow for increased familial representation, preferably including pedigrees with a pronounced history of AxSpA. Thus, the ongoing expansion of the ASRI cohort presents an opportunity to initiate

investigations into rare genetic variants associated with AxSpA disease using pedigree-based methods.

Another constraint within our investigative scope pertains to the unravelling of disease-associated haplotypes within the HLA gene region. Within our investigations, we found that specific SNPs within the HLA region may be more strongly linked with uveitis than with AxSpA. Similarly, our observations indicated a likelihood of multiple disease-associated variants within the *ERAP1* gene region. Due to the high linkage disequilibrium within these regions and the relatively homogenous Caucasian population employed in our study, we were unable to effectively delineate many of these signals. A conceivable approach to addressing this limitation involves the amalgamation of our genetic datasets with those derived from diverse population cohorts, particularly those of non-Caucasian ancestry. It is anticipated that such divergent population groups may harbour distinct common haplotypes absent in Caucasian cohorts, thereby facilitating the refined mapping of genetic signals. Although the prospect of inter-population variability raises the spectre of disparate causative variants, current evidence suggests a considerable degree of shared genetic architecture across populations (Amariuta et al., 2020), a phenomenon recently showcased in IBD (Gettler et al., 2022). Such a study could be supported by haplotype associated gene expression and chromatin accessibility analysis to identify which genes may be linked to specific variants.

Collaborative work may also benefit our findings in the clinical presentation of AxSpA. Concurrently, ongoing investigations across Europe, such as the RESPONSER and DESIR cohorts, are structurally similar to ASRI. While our investigations within ASRI have yielded numerous findings that align with those of our European counterparts, we have also unearthed discordant observations and delved into previously unexplored areas. Several of our initial findings, including strong HLA-B27 positivity in women with a history of uveitis should preferably be replicated in other cohorts to confirm these associations, while other observations, such as patient outcomes following cessation of smoking require increased sample sizes to resolve effectively. Likewise, anecdotal findings from these other studies may benefit from increased study power through collaboration with ASRI. Through collaborative frameworks, we anticipate not only the validation and consolidation of existing findings but also the generation of novel insights into the clinical trajectory and prognostic markers of AxSpA.

Our preliminary work into gene expression profiles of AxSpA patients using single cell RNA analysis has identified a wide range of differentially expressed genes spanning diverse cellular subsets. Building upon these foundational insights, our forthcoming investigations aim to extend this analytical framework to encompass T cells subjected to activating stimuli to identify potential divergent genetically determined responses between AxSpA patients and healthy controls. To

this end, we will perform a comparison of gene expression profiling within activated T cell subsets in AxSpA patients and controls using differential expression studies. Since certain disease associated variants may be associated with gene expression only during activated or resting states (Calderon et al., 2019), we will also focus on whether GWAS variants may be associated with altered expression of nearby genes uniquely in stimulated T cells or their subsets. Pathway enrichment analysis of the differentially expressed genes within stimulated T cells will be used to explore the perturbed biological pathways inherent to AxSpA pathogenesis. By comparing these findings with those derived from unstimulated cells, we aim to identify common differentially expressed pathways in AxSpA patients, providing mechanistic insights and possibly identifying additional genes which contribute to these pathways, likely to be active only in stimulated cells.

Furthermore, our investigative scope extends to exploring potential disparities in gene expression profiles between older biologic-naïve AxSpA patients presenting with more senescent disease and their younger counterparts recently diagnosed with active disease. We also aim to identify whether differential molecular pathways in early onset disease differ from those in late onset. In undertaking this comparative analysis, we aim to ascertain whether molecular pathways associated with AxSpA exhibit uniformity in terms of their presence and the magnitude of their association across these distinct demographic cohorts. This inquiry presents not only the possibility of delineating age-related variations in the transcriptional landscape characterising AxSpA pathogenesis but also offers a unique vantage point for discerning potential age-specific molecular signatures implicated in disease aetiology.

Finally, some recent studies have suggested that the genetic risk profile of autoimmune disease may not be static, but rather evolving over time. This is a finding which was first reported in type 1 diabetes and later replicated in coeliac disease (Vehik et al., 2013, Dominguez Castro et al., 2017, Dominguez Castro et al., 2020). Although the underlying cause for this phenomenon has not yet been resolved, it is suspected that changing environmental exposures over time lead to variation in the cast of susceptibility alleles that can contribute to disease aetiology. For example, shifts in gut microbiota relating to changing diets may be an underlying factor in some diseases, and not solely those involving predominantly the gastro-intestinal tract. As previously discussed, there is a rational link between AxSpA and the microbiome given the documented pervasiveness of gastro-intestinal inflammation in these patients. We are therefore interested to investigate whether the genetic profile of AxSpA patients may be changing with time. While we initially endeavoured to explore these possibilities as part of this work, we encountered challenges related to survivorship bias due to the cross-sectional study design of ASRI. However, we are actively planning follow-up studies with currently registered ASRI participants, in

conjunction with ongoing recruitment efforts, which may open avenues for conducting this type of research in the future.

Supplementary material

Chapter 1 supplementary

Table S1.1: Genetic concordance between associated diseases

Gene	AS	AAU	Psoriasis	UC	Crohn's	Mean Assoc. Score	Putative role
ERAP1	+	+	+	+	+	0.92	Ag processing/Pres
GPR65	+	-	+	+	+	0.95	Phagocytosis
TNFSF15	+	-	+	+	+	1.00	Apoptosis
B3GNT2	+	+	+	+	+	0.79	Protein metabolism
ANKRD55	+	-	+	+	+	1.00	Cytoskeletal
BACH2	+	-	+	+	+	0.96	Lymphocyte.
C1orf141	+	+	+	+	+	0.88	Unknown
SMAD3	+	-	+	+	+	1.00	TGF beta superfamily
FCGR2A	+	-	+	+	+	0.79	Phagocytosis
RUNX3	+	-	+	+	+	0.83	T cell
GPR35	+	-	+	+	+	0.92	Metabolism
IL2RA	+	-	+	+	+	0.96	T cell
NKX2-3	+	-	+	+	+	0.95	Cellular differentiation
HLA-B	+	+	+	+	+	0.82	Unknown
IL23R	+	+	+	+	+	0.89	Lymphocyte IL23/IL17
INAVA	+	-	+	+	+	0.91	Macrophage inflammatory
MUC19	+	-	+	+	+	0.96	Other
OSMR	+	-	+	+	+	0.94	Lymphocyte T cell IL23/IL17
IL6R	+	-	+	+	+	0.71	Lymphocyte IL23/IL17
CDKAL1	+	-	+	+	+	0.97	Other
TNFRSF1A	+	-	+	+	+	0.72	Inflammatory mediator
KSR1	+	-	+	+	+	0.91	Cellular proliferation
NRSF2	+	-	+	+	+	0.91	Transcriptional regulation
CD226	+	-	+	+	+	0.87	Ag processing/Pres
SPRED2	+	-	+	+	+	0.90	Other
IL12B	+	-	+	+	+	0.94	Lymphocyte T-cell IL23/IL17
ANKRD30A	+	-	+	+	+	0.86	Cytoskeletal
JAZF1	+	-	+	+	+	0.88	Transcriptional regulation
TLR4	+	-	+	+	+	0.83	Receptor inflammatory
PTPN2	+	-	+	+	+	0.96	Lymphocyte T cell IL23/IL17
CD6	+	-	+	+	+	0.90	T cell
ADCY7	+	-	+	+	+	0.89	Metabolism
IL18R1	+	-	+	+	+	0.87	TGF beta superfamily
ZNFR31	+	-	+	+	+	0.86	Transcriptional regulation
SLC9A8	+	-	+	+	+	0.82	Transporter
MAP3K8	+	-	+	+	+	0.85	Inflammatory mediator
FUT2	+	-	+	+	+	0.87	Other
ADGRL2	+	-	+	+	+	0.80	Cellular adhesion
IRF8	+	-	+	+	+	0.89	Transcriptional regulation
PUS10	+	-	+	+	+	0.92	Transcriptional regulation
ACD67752.1	+	-	+	+	+	0.89	Unknown
TAGAP	+	-	+	+	+	0.88	T cell
IKZF1	+	-	+	+	+	0.89	Transcriptional regulation
UBASH3A	+	-	+	+	+	0.77	Protein degradation
RIPK2	+	-	+	+	+	0.79	Protein metabolism
DLD	+	-	+	+	+	0.82	Protein metabolism
ASAP2	+	-	+	+	+	0.76	Other
HDAC7	+	-	+	+	+	0.76	Transcriptional regulation
BANK1	+	-	+	+	+	0.80	B cell cytoskeletal
C7orf33	+	-	+	+	+	0.80	Unknown
TNFAIP3	+	-	+	+	+	0.83	Protein degradation
FOSL2	+	-	+	+	+	0.81	Transcriptional regulation
ADCY3	+	-	+	+	+	0.82	Cell signalling
AC005697.1	+	-	-	-	-	0.74	Unknown
PRDM1	+	-	+	+	+	0.91	Apoptosis
SOX14	+	-	+	+	+	0.73	Transcriptional regulation
IFNLRI	+	-	+	+	+	0.80	T-cell JAK/STAT
NFKBIA	+	-	+	+	+	0.80	Inflammatory mediator
ZBTB40	+	-	+	+	+	0.79	Transcriptional regulation
PSMG1	+	-	+	+	+	0.69	Chaperone
LNPEP	+	-	-	-	+	0.86	Ag processing/Pres
IFNG	+	-	+	+	+	0.85	Inflammatory mediator
IL12RB2	+	-	+	+	+	0.75	T-cell IL23/IL17
UBE2E3	+	-	+	-	-	0.80	Protein degradation
IRF5	+	-	+	+	+	0.78	Apoptosis inflammatory mediator
ZC3H12C	+	-	+	+	+	0.73	Transcriptional regulation
SLC39A11	+	-	+	+	+	0.78	Other
PRKCB	+	-	+	+	+	0.78	Cell signalling
PTPN22	+	-	+	+	+	0.74	Ag processing/Pres
CD28	+	-	+	+	+	0.76	Lymphocyte
CCL2	+	-	+	+	+	0.78	Macrophage inflammatory
ANTXR2	+	-	+	+	+	0.70	Cellular adhesion
NXPE1	+	-	+	+	+	0.77	Unknown
PARK7	+	-	+	+	+	0.75	Other
NFKB1	+	-	+	+	+	0.76	Macrophage inflammatory
IFNGR2	+	-	+	+	+	0.78	Inflammatory mediator JAK/STAT
MICA	+	+	-	-	-	0.68	Ag processing/Pres
SP140	+	-	+	+	+	0.74	Transcriptional regulation
FOXO1	+	-	+	+	+	0.73	Transcriptional regulation
FAP	+	-	+	+	+	0.77	Tissue remodelling
TRIB1	+	-	+	+	+	0.74	Protein degradation
JAK2	+	-	+	+	+	0.85	JAK/STAT
TNFSF18	+	-	+	+	+	0.74	Inflammatory mediator
KIF21B	+	-	-	+	+	0.55	Other
ILZ7	+	-	+	+	+	0.65	Lymphocyte T cell
GNA12	+	-	+	+	+	0.73	Other
CPEB4	+	-	+	+	+	0.74	Protein degradation
RORC	+	-	+	+	+	0.70	Transcriptional regulation
SOX4	+	-	+	+	+	0.70	Transcriptional regulation
CREM	+	-	+	+	+	0.76	Transcriptional regulation
TMEM17	+	+	+	+	+	0.55	Other
PPP5C	+	-	+	+	+	0.67	Other
TNFRSF14	+	-	+	+	+	0.72	Ag processing/Pres
CARD9	+	-	+	+	+	0.75	Inflammatory mediator
STAT3	+	-	+	+	+	0.80	JAK/STAT
PDGFB	+	-	+	+	+	0.75	Lymphocyte B cell
GAL3ST2	+	-	+	+	+	0.66	Other
LINC02694	+	-	+	+	+	0.63	Other
PLCL1	+	-	+	+	+	0.78	Macrophage inflammatory
DAP	+	-	+	+	+	0.70	Apoptosis

Table S1: Top 100 non-MHC genes associated with AS and their commonality with AAU, psoriasis, UC and CD. Prepared from data obtained through Open Targets and Genecards (Ochoa et al., 2020, Stelzer et al., 2016).

Chapter 3 supplementary

Table S3.1: Haplotype associations within ERAP1 500Kb region, based on 1,000 permutations.

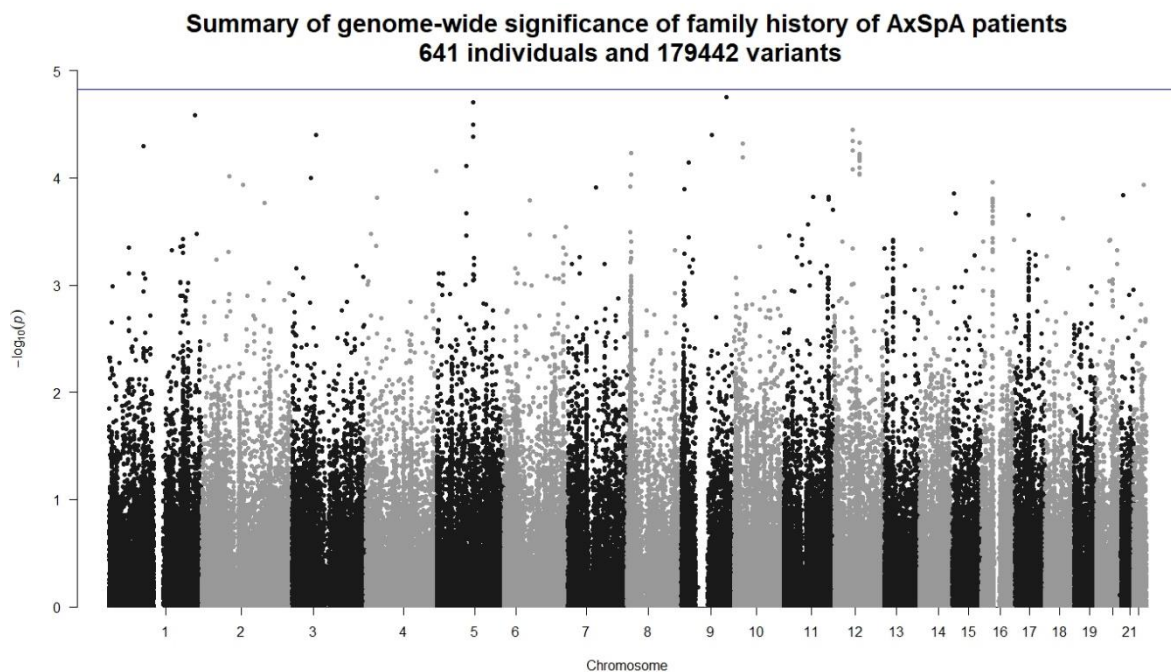
Name	Chi Square	Permutation p-value
Block 6: AAGCGGCAACAA	25.623	<<0.001
Block 16: ATCGGTC	19.415	<0.001
Block 7: TGGATATATCGAGCCCTTCATTGACCATGCCATGTGCAACGGTGTACCTTCGTAGCTCC	15.562	0.001
Block 8: TCCTAAGAGCCTGAAAACGGCAGTC	17.407	0.001
Block 5: TTTGAACAA	13.842	0.008
Block 4: TG	12.898	0.019
Block 11: TAATATACTTTGCGGACCTACCTTTTAAGAGAAGAATGCAGCGAGTGGTCTATTACAGTATGCCCTAAGGATTAC	12.976	0.019
Block 16: AGTGCTA	11.836	0.033
Block 6: AGGTGGTCGGAA	11.199	0.056
Block 8: CTCTGAGTCGCTGGAACCGCAATC	11.137	0.057
Block 4: CG	11.052	0.06
Block 7: AGGATATATCGAGCCCTCACCAGCTAGATCATGCCTAGTCGCGCTTGCCTCACTAC	8.894	0.177
Block 11: TGATATACTTTGCGGACCTACCTTTTAAGAGAAGAATGCAGCGAGTGGTCTATTACAGTATGCCCTAAGGATTAC	8.379	0.229
Block 3: GGGACCT	8.235	0.244
Block 9: TTACCGATGTC	8.231	0.244
Block 2: AGGAAAG	7.981	0.27
Block 9: GTACCGATGCT	7.309	0.352
Block 6: AAACAGCAACGA	6.679	0.443
Block 5: TCTAAACAG	6.571	0.459
Block 10: ATCTAT	5.516	0.629
Block 11: GGGGTGCTCTATGATTATGCCGCTGGACAGAATGGCATGTAAGATGTACTAATATCATGCGCCCGAGAGGCTGC	4.917	0.782
Block 7: AGGATATACCGGGCCCTCACCAGCTAGATCATGCCTAGTCGCGCTCACTTGCCTCACTAC	4.86	0.791
Block 5: TCTAAACAA	4.497	0.849
Block 10: GTCTAT	4.105	0.89
Block 11: TAATATACTTTGCGGACCTACCTTTTAAGAGAAGAATGCAGCGAGCGTCTATTACAGTATGCCCTAAGGATTAC	3.868	0.919
Block 16: AGTGGAA	3.266	0.977
Block 7: TGGATATATCGAGCCCTCACTGACCAGGTCACATCTAACGGCGTTACCTGCGTAGCCAT	3.205	0.981
Block 3: GAGACTT	2.94	0.99
Block 3: GAGACCT	2.655	0.995
Block 7: AGGATATATCGAGCCCTCGCTGACCAGGTCACGTGTGACGGCGTTACCATCATAGCTCC	2.66	0.995
Block 7: AAATCGGTTTAAAGTCTTACC GGCTGGGTGCGCTCTAACGACGCCACCTGCGTAGCTAC	2.546	0.998
Block 2: GGCAAAA	2.261	0.999
Block 16: GGTGTGA	2.062	0.999
Block 1: GAA	1.153	1
Block 1: AGC	1.706	1
Block 2: GGGGAAA	0.549	1
Block 2: GCCACGA	0.023	1
Block 3: CGTGTC	0.002	1
Block 4: CT	0.006	1
Block 5: TCTAAATA	0.446	1
Block 5: CCTAGTCAA	0.033	1
Block 5: TCAGAACAA	0.048	1
Block 5: TCTAAACTA	0.185	1
Block 6: GAGCAACAACGA	0.689	1
Block 6: AAGCAGCAACGG	0.483	1
Block 6: GAGCAACAACAA	0.079	1
Block 7: AGGATATACCGGGCCCTCACCAGCTAGATCATGCCTAGTCGCGTCACTTGCCTCAATAC	2.02	1
Block 7: AGGATATATCGAGCCCCCACTCACCAGGTCACGTCTAACCGCGTTACCTGTGCCACTAC	0.399	1
Block 7: AAATCGGATTAAAGTCTTACC GGCTGGGTGCGCTCTAACGACGCCACCTGCGTAGCTAC	0.003	1
Block 7: AGGATATATCGAGCCCTCACTGAACAGGTTACATCTAACGGCCTCATCTTCGTAGCTCC	0.196	1
Block 8: TCTTAATACGCTGAAAACCGCAGTC	1.094	1
Block 8: TTCTAGGAGCCTGAAAACGGCAGTC	0.933	1
Block 8: TTCTAAGACGCTGAAAACCAAGTC	0.26	1
Block 8: TCCCAAGAGCCTGAAAACGGCAGTC	1.716	1
Block 8: TCCTAATACGCTGAAAACCGCAGTC	0.005	1
Block 8: CTCTGAGTCGCTGGGCTCGGTGCA	0.18	1
Block 9: GTACGGACGCT	0.115	1
Block 9: GTTCCGATTCT	0.685	1
Block 9: GCAGCTGCGCT	0.399	1
Block 10: GATCGA	0.121	1
Block 11: TAATATACTTTGCGGACCTACTTTTAAGAGAAGAATGCAGCGAGTGGTCTATTACAGTATGCCCTAAGGATTAC	0.071	1
Block 11: TAATATACTTTGCTGACCTACCTTTTAAGAGAAGAATGCAGCGAGTGGTCTATTACAGTATGCCCTAAGGATTAC	0.001	1
Block 11: GGGGTGCTCTATGATTATGCCGCTGGACAGAATGGCATGTAAGATGTACTAATATCATGCGCCCGAGAGGCTGC	0.362	1
Block 12: GCA	0.465	1
Block 12: ATG	0.376	1
Block 13: AACATAACGTTTC	1.146	1
Block 13: AACATAACTTCGT	1.829	1
Block 13: AACATAACGCTTC	0.471	1
Block 13: AACATAACTTCTT	1.126	1
Block 13: GTTGCGCTGTTTC	1.389	1
Block 14: TCTTCT	0.931	1
Block 14: CAACTC	0.018	1
Block 14: TCTTTC	1.663	1
Block 15: AA	1.433	1
Block 15: AG	0.356	1
Block 15: GA	0.542	1
Block 16: AGTGGTA	0.108	1

Table S3.2: Univariate and multivariate analysis between rs30187 allelotypes in HLA-B27- patients.

Characteristic	Genotype CC		Genotype CT		Genotype TT		Cochran-Armitage or Ordinal Logistic Regression (P)	Multivariate Analysis (P)
	n	Median (IQR) or n (%)	n	Median (IQR) or n (%)	n	Median (IQR) or n (%)		
Male (%)	44	32 (72.7)	52	39 (75.0)	2	1 (50.0)	0.951	0.879
Family history, n (%)	43	0 (0.0%)	52	0 (0.0%)	2	0 (0.0%)	0.722	0.687
Age in years, median (IQR)	44	50 (37 - 57)	52	48 (40 - 58)	2	44 (42 - 45)	0.797	--
Age at first symptom onset, median (IQR)	44	33 (24 - 42)	52	31 (24 - 43)	2	41 (39 - 42)	0.795	--
Age at diagnosis, median (IQR)	44	42 (32 - 52)	52	40 (34 - 47)	2	44 (42 - 45)	0.585	0.427
mNY criteria* n (%)	44	33 (75.0%)	52	45 (86.5%)	2	2 (100.0%)	0.109	0.111
Uveitis**, n (%)	43	9 (20.9%)	52	7 (13.5%)	2	0 (0.0%)	0.252	0.419
Psoriasis, n (%)	43	12 (27.9%)	52	12 (23.1%)	2	0 (0.0%)	0.414	0.454
Inflammatory bowel disease, n (%)	43	8 (18.6%)	52	7 (13.5%)	2	0 (0.0%)	0.383	0.614
Depression, n (%)	44	3 (6.8%)	52	7 (13.5%)	2	0 (0.0%)	0.422	--
Clinical Instruments:								
BASDAI score, median (IQR)	44	4.1 (2.3 - 5.8)	50	4.1 (2.2 - 6.3)	2	6.5 (5.8 - 7.3)	0.681	--
BASMI score, median (IQR)	35	3.2 (2.4 - 5.6)	46	3.8 (2.6 - 5.4)	1	1.0 (1.0 - 1.0)	0.656	--
BASFI score, median (IQR)	42	4.3 (2.2 - 5.6)	48	3.8 (1.5 - 6.4)	2	4.9 (3.8 - 5.9)	0.951	--
ASQoL score, median (IQR)	44	7.0 (1.0 - 10.5)	52	5.0 (1.0 - 13.3)	2	8.5 (8.3 - 8.8)	1	--
HAQ-s score, median (IQR)	44	0.25 (0.00 - 0.63)	51	0.50 (0.06 - 0.75)	2	0.19 (0.09 - 0.28)	0.382	--
CRP, median (IQR)	38	2.80 (0.9 - 8.5)	47	1.30 (0.8 - 5.1)	2	3.90 (2.9 - 4.9)	0.114	--
ESR, median (IQR)	33	19.18 (6.0 - 20.0)	47	14.12 (4.0 - 22.0)	2	11.00 (11.0 - 11.0)	0.205	--

* Refers to modified New York criteria 1984 for AS.

** At least one reported episode of acute anterior uveitis.

**Fig. S3.1:** Genome-wide scan of variants associated with family history of AxSpA, based on 180 patients with and 461 patients without family history of AxSpA.

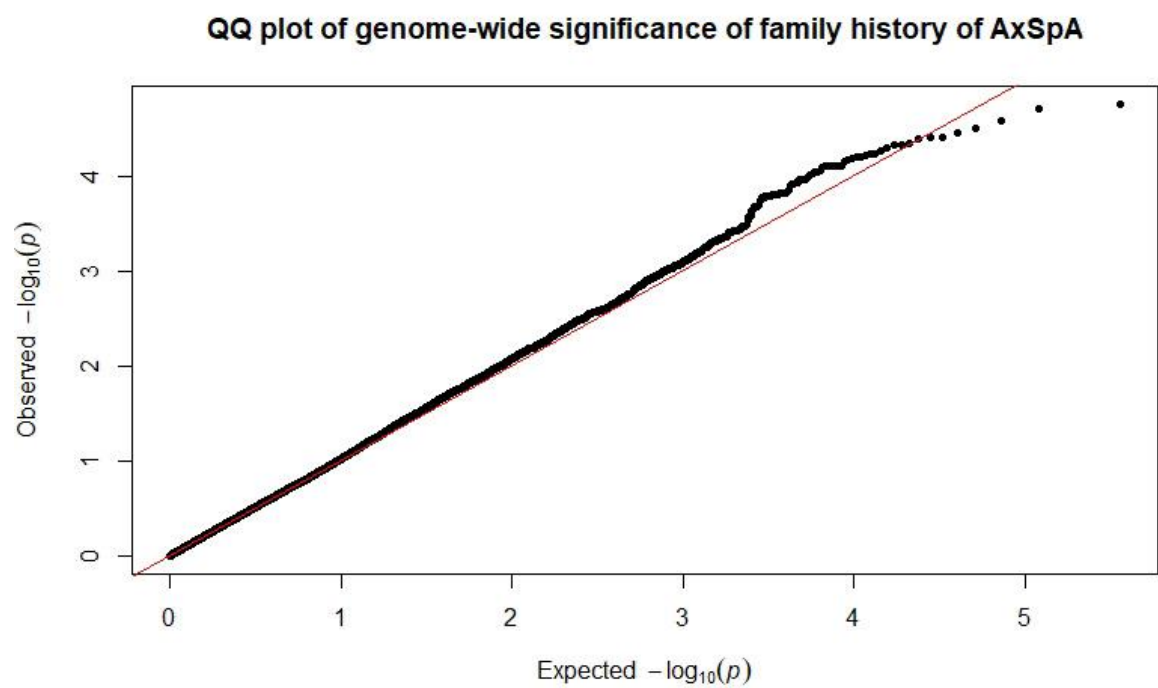


Fig. S3.2 QQ-plot of distribution of significance presented in Fig. S3.1.

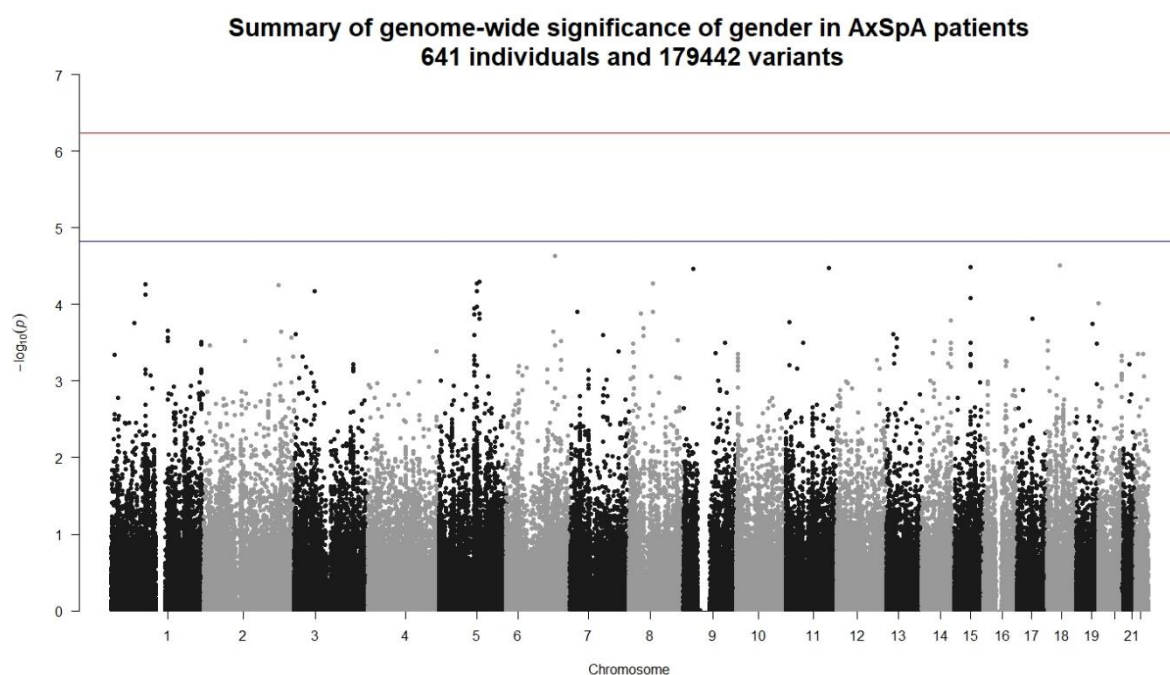


Fig. S3.3: Genome-wide scan of significance of variants associated with gender assigned at birth in AxSpA, based on 497 male and 138 female patients.

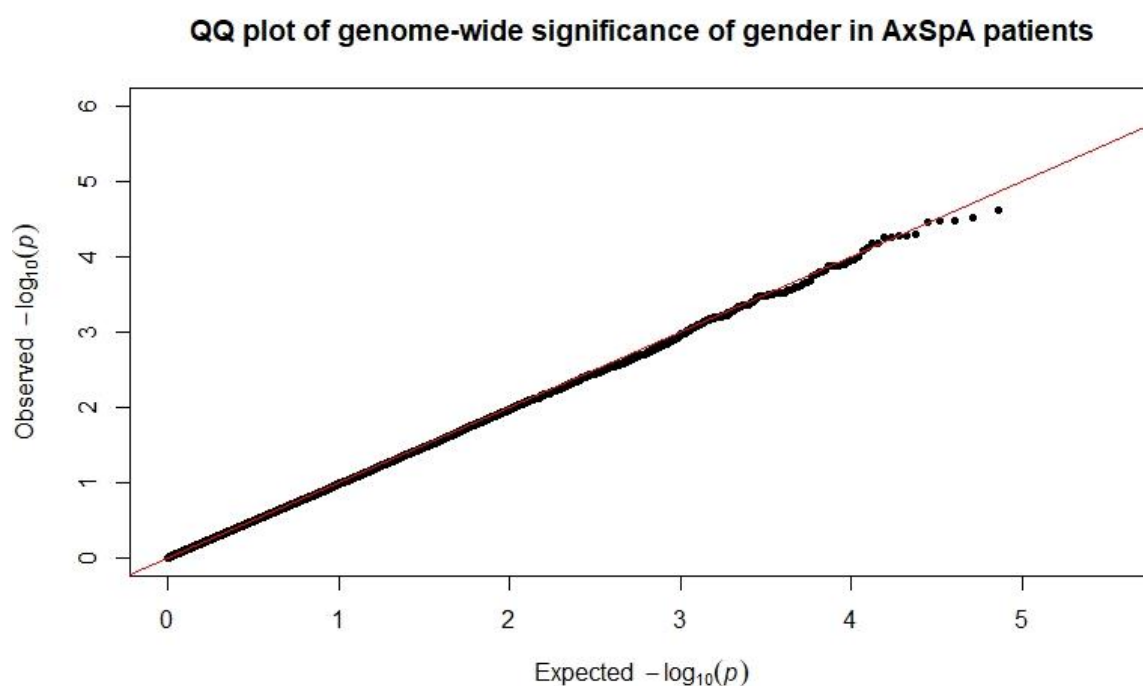


Fig. S3.4 QQ-plot of distribution of significance presented in Fig. S3.3.

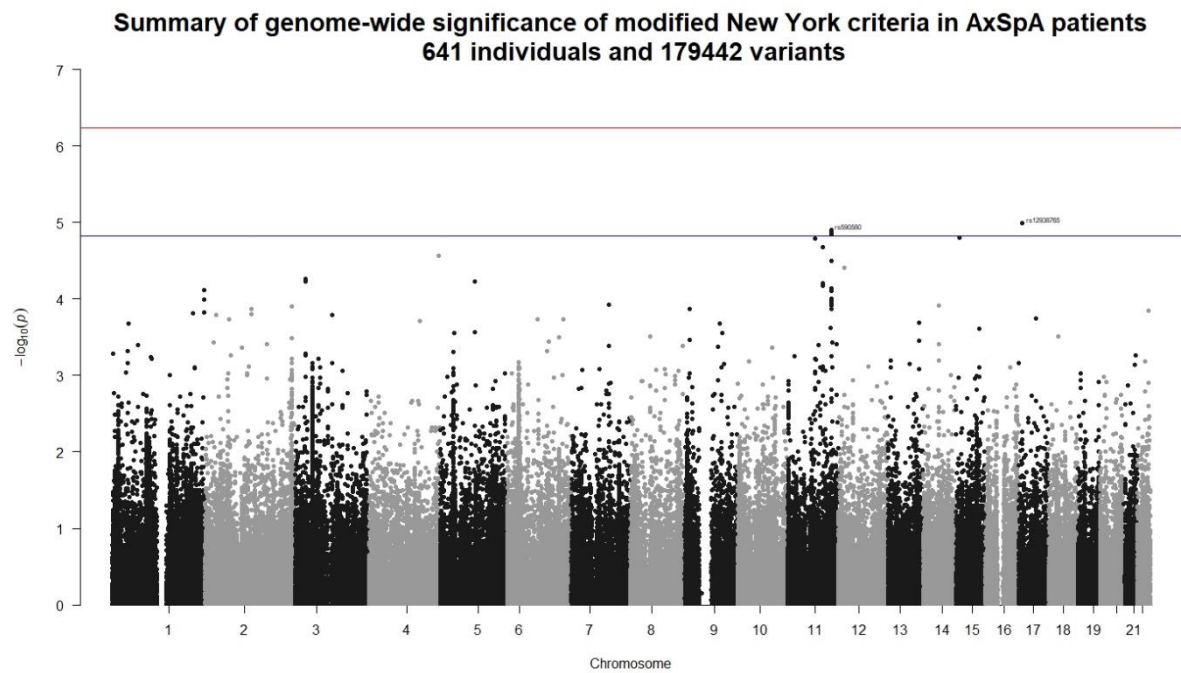


Fig. S3.5: Genome-wide scan of variants associated with modified New York criteria (diagnostic criteria for AS), based on 180 patients with and 461 patients without family history of AxSpA.

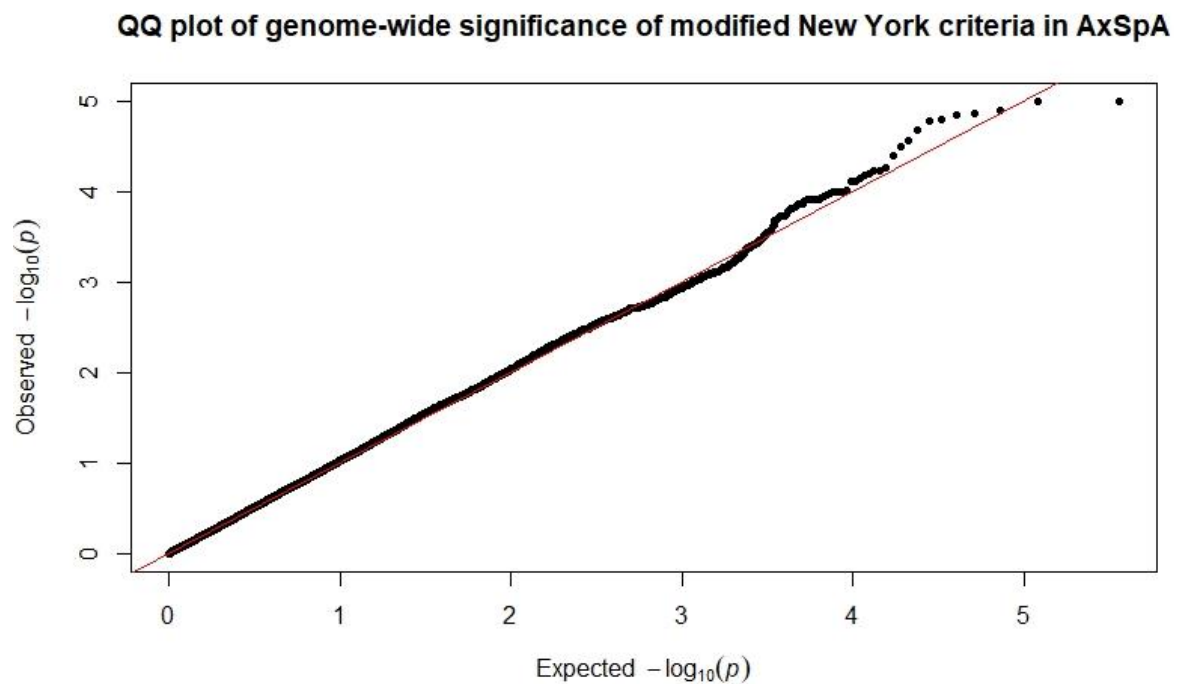


Fig. S3.6 QQ-plot of distribution of significance presented in Fig. S3.5. Fig. S3.6

Chapter 4 supplementary

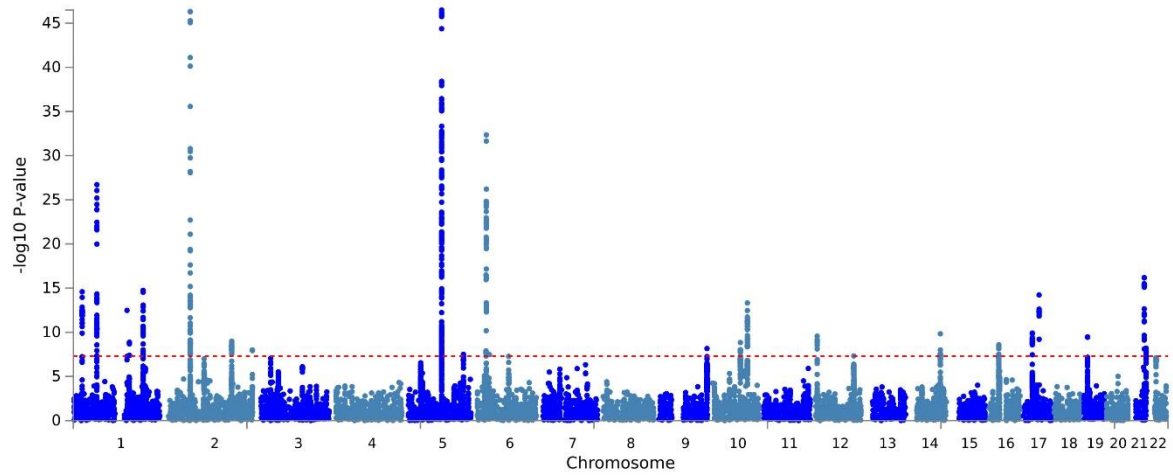


Fig. S4.1: Manhattan plot reconstructed from IGAS summary data (Cortes et al., 2013b, Watanabe et al., 2017).

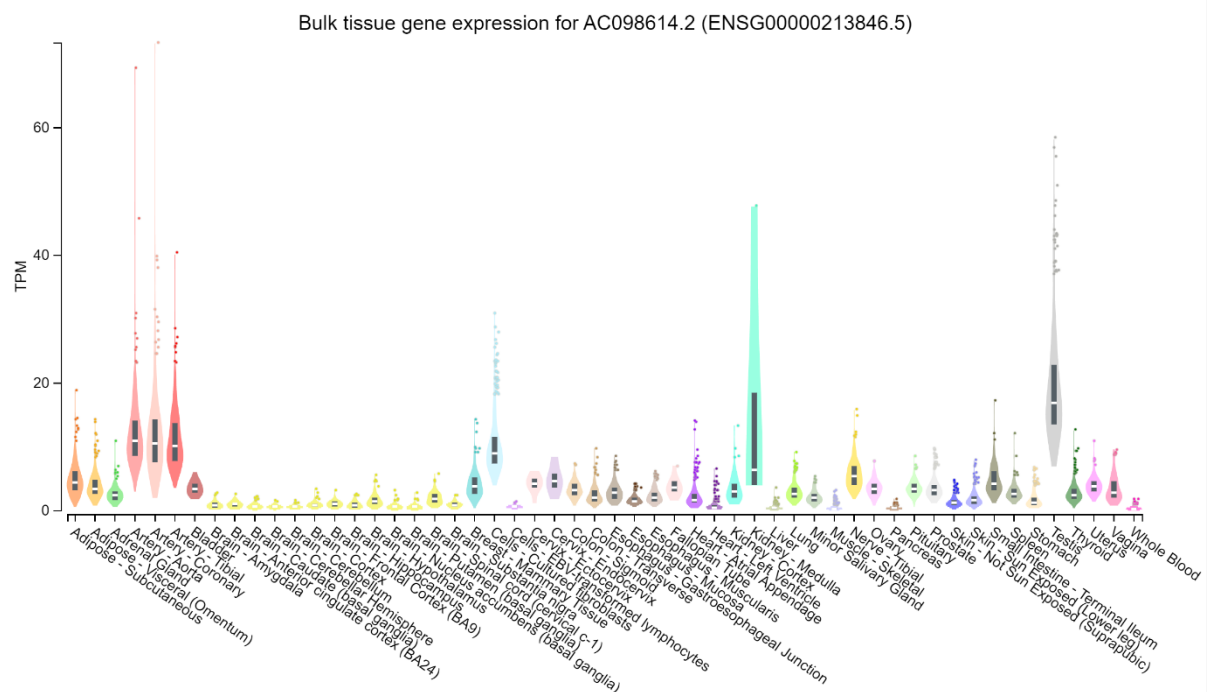


Fig. S4.2: Tissue expression of AC098614.2 (Carithers and Moore, 2015).

Table S4.1: LD between rs28653390 and both rs572715043 and rs6804441 (International Multiple Sclerosis Genetics et al., 2013) in CEU and GBR populations (Machiela and Chanock, 2015).

rs57271503

chr3:119244593

A

G

rs28653390

chr3:119259163

A

G

1

54

65

260

55 (0.145)

325 (0.855)

66 (0.174)

314 (0.826)

380

Haplotypes

Statistics

G_G: 260 (0.684)

G_A: 65 (0.171)

A_G: 54 (0.142)

A_A: 1 (0.003)

D': 0.8953

R²: 0.0285

Chi-sq: 10.835

p-value: 0.001

rs6804441

chr3:119260944

A

G

rs28653390

chr3:119259163

A

G

55

0

263

62

55 (0.145)

325 (0.855)

318 (0.837)

62 (0.163)

380

Haplotypes

Statistics

G_A: 263 (0.692)

G_G: 62 (0.163)

A_A: 55 (0.145)

A_G: 0 (0.0)

D': 1.0

R²: 0.033

Chi-sq: 12.538

p-value: 0.0004

rs28653390 and rs57271503 are in linkage equilibrium

rs28653390 and rs6804441 are in linkage equilibrium

Table S4.2: Investigation of further *IL1R2* region SNPs.

	Published				O/R (ASRI	P (Meta-
Disease	SNP	ASRI proxy SNP	r ²	D'	GWAS)	analysis)
Coeliac and Crohn's	rs917997	rs2110734	1	1	1.118	6.3x10 ⁻⁴
IBD	rs10185424	rs7588933	0.99	1	0.9664	1.4x10 ⁻⁵
IBD	rs13022757	rs11123902	0.98	1	0.9237	2.7*10 ⁻⁴

Chapter 5 supplementary

Table S5.1: List of antibodies sourced from BD Biosciences used in single cell identification.

Marker name	Cat #	clone	Cell targets
CCR2 (CD192)	940286	LS132.1D9	Monocytes
CCR7 (CD197)	940014	3D12	naïve/CM T cells
CD10	940045	HI10a	Neutrophil
CD102	940241	CBR-IC2/2	ICAM2
CD115	940225	9-4D2-1E4	CSF1R, monocytes & Macrophages
CD117	940250	104D2	cKIT, mast cells
CD11b	940266	ICRF44	Mac1 integrin, nrutrophils, eosinophils and monocytes/macrophages
CD11c	940024	B-ly6	Monocytes, neutrophils, B cells
CD123	940020	7G3	IL-3R
CD127	940012	HIL-7R-M21	Effector and memory T cells
CD14	940005	MΦP9	Monocytes
CD141	940079	1A4	Neutrophils, monocytes
CD142	940280	HTF-1	Tissue Factor
CD15	940274	W6D3	Neutrophils
CD16	940006	3G8	NK cells, monocytes
CD161	940283	HP-3G10	MAIT cells
CD163	940058	GHI/61	Monocytes
CD18	940086	6.7	Most immune cells
CD183 (CXCR3)	940030	1C6/CXCR3	Th1 cells
CD184 (CXCR4)	940056	12G5	Neutrophil maturity
CD19	940247	HIB19	B cells
CD193	940245	5E8-G9-B4	Eosinophils
CD194	940047	1G1	CCR4, expressed on many cells
CD196 (CCR6)	940033	11A9	many cells, promotes migration to sites of inflammation
CD1a	940063	HI149	Antigen-presenting cells
CD1d	940296	CD1d42	Antigen-presenting cells
CD206	940068	19.2	Mannose receptor
CD226	940075	DX11	Some moDC, CD8 T cells, NK cells
CD235a	940040	GA-R2 (HIR2)	Red blood cells
CD25	940009	2A3	Treg cells
CD26	940101	M-A261	cDC
CD27	940018	M-T271	IFN-g+ NK cells, B cell memory marker
CD273	940071	MIH18	PDL-2 inhibitory receptor
CD274	940035	MIH1	PD-L1
CD279 (PD-1)	940015	EH12.1	Inhibitory receptor on T and B cells
CD28	940017	CD28.2	naïve T cells
CD294 (CRTH2)	940098	BM16	ILC2
CD3	940000	SK7	T cells
CD31	940254	WM-59	Platelets
CD32	940069	FLI8.26	FcγRII
CD335	940064	9E2/NKp46	NKp46
CD38	940013	HIT2	T, B and NK cells
CD4	940001	SK3	CD4 T cells
CD44	940364	515	HA receptor, many cells,
CD45	940002	HI30	All leukocytes
CD45RA	940011	HI100	naïve/TD T cells
CD45RO	940022	UCHL1	Memory T cells
CD5	940318	L17F12	Certain lymphocytes
CD56	940007	NCAM16.2	NK cells
CD62L	940041	DREG-56	Memory T cell marker, Neutrophils
CD62P	940309	AC1.2	
CD64	940023	10.1	Macrophages
CD73	940294	AD2	
CD8	940003	RPA-T8	CD8 T cells
CD86	940025	2331 (FUN-1)	Antigen-presenting cells
CD94	940081	HP-3D9	NK cells, CD8 T cells
CXCR5 (CD185)	940042	RF8B2	Follicular helper T cells
E-Cadherin	940210	67A4	Epithelial
FcεR1a	940220	AER-37	Mast cells and basophils
HLA ABC	940062	G46-2.6	MHC I
HLA DR	940010	G46-6	Monocytes
IgD	940026	IA6-2	Unswitched B cells
IgM	940276	G20-127	B cells
Podoplanin	940270	LpMab-17	Epithelial, endothelial cells
TCR gamma delta	940057	B1	γδ T cells
TIM3 (CD366)	940066	7D3	Inhibitory receptor
CD117	940051	YB5.B8	Mast cells and basophils

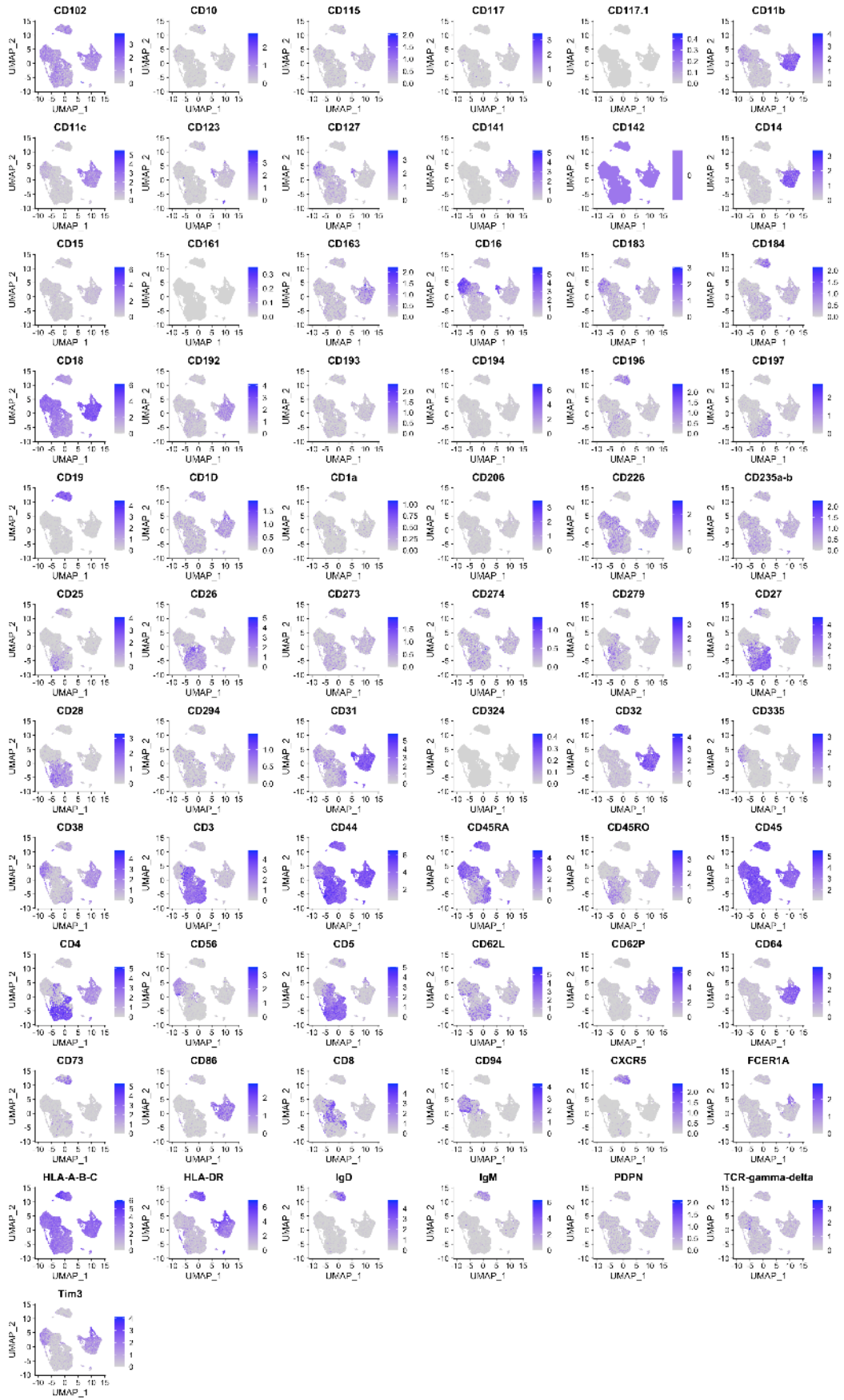


Fig. S5.1: Antibody binding in cell clusters.

Table S5.2: Differential genes in monocytes.

Monocytes up-regulated			Monocytes up-regulated			Monocytes up-regulated			Monocytes down-regulated		
Gene	Log2 fold change	P _{adj}	Gene	Log2 fold change	P _{adj}	Gene	Log2 fold change	P _{adj}	Gene	Log2 fold change	P _{adj}
IFI6	1.53	1.99E-25	GPCPD1	0.48	2.95E-10	MYO9B	0.30	2.08E-02	GPSM3	-0.25	1.74E-04
IFITM3	1.46	5.92E-29	CALR	0.48	1.97E-09	TMEM123	0.30	1.50E-02	VAMP2	-0.25	1.25E-02
IL1RN	1.17	4.32E-26	PLK3	0.47	1.53E-04	NAAA	0.29	2.58E-02	VSIR	-0.25	3.41E-04
RGS1	1.15	4.97E-13	KDM6B	0.47	1.39E-09	RILPL2	0.29	1.68E-04	CDC42EP3	-0.26	3.12E-03
GADD45B	1.15	1.59E-32	ARHGAP26	0.47	3.79E-06	CHD1	0.29	3.66E-02	AP1S2	-0.26	1.25E-06
CXCL2	1.14	1.46E-12	SLC25A37	0.45	6.74E-05	RPN2	0.29	1.79E-02	ANXA2	-0.27	8.49E-09
ZNF331	1.07	4.24E-25	FOXO3	0.45	2.64E-05	IRF2BP2	0.29	9.30E-07	TMEM170B	-0.27	2.64E-02
LMNA	1.00	2.51E-21	UBAP1	0.44	3.95E-05	LIMS1	0.29	1.68E-02	MYH9	-0.27	9.01E-04
IER3	0.98	1.15E-40	NEDD9	0.44	1.80E-06	ID2	0.29	4.00E-05	CALM2	-0.27	6.29E-05
MAP3K8	0.98	2.15E-43	AL390957.1	0.44	4.63E-03	LAPTM5	0.28	5.93E-11	SUMO2	-0.27	5.10E-04
OLR1	0.97	4.95E-21	RASGEF1B	0.44	2.10E-11	C16orf72	0.28	9.95E-03	GRK2	-0.28	5.91E-03
PNP	0.96	7.84E-30	PID1	0.43	3.15E-05	PPP1CB	0.28	3.07E-02	BRD2	-0.28	8.60E-03
RGCC	0.93	5.65E-45	GNA13	0.43	5.20E-05	FOSL2	0.28	6.19E-04	FGL2	-0.28	4.26E-04
CCL3	0.93	5.05E-13	NR3C1	0.43	1.96E-04	LDHA	0.28	7.67E-03	KIAA0930	-0.29	4.69E-02
PHACTR1	0.93	8.38E-30	TLE3	0.43	1.96E-04	MCL1	0.28	3.66E-09	HLA-A	-0.30	6.54E-26
HSPA5	0.92	5.36E-31	TSC22D2	0.43	9.90E-03	JUND	0.27	6.85E-15	BTG2	-0.30	8.22E-05
HES1	0.85	7.21E-05	KLF4	0.43	1.15E-18	CSNK1A1	0.26	4.35E-03	OGFRL1	-0.30	2.05E-03
BHLHE40	0.84	2.09E-23	ADAM17	0.42	1.25E-05	RNF19B	0.26	9.74E-03	SRSF7	-0.31	1.28E-02
SOC3	0.82	3.28E-23	MAPK6	0.42	2.82E-04	BACH1	0.26	3.37E-02	MEF2C	-0.31	4.08E-02
LONRF3	0.82	2.59E-12	HAVCR2	0.42	1.43E-02	ARF6	0.25	1.96E-04	YPEL3	-0.32	3.64E-03
LUCAT1	0.82	1.72E-11	ATP2B1-AS1	0.42	3.09E-11	CD63	0.25	1.63E-03	SNX10	-0.32	4.77E-03
GPR183	0.80	5.39E-20	NR4A3	0.42	3.51E-02	TPM4	0.25	1.78E-03	ZFP36L2	-0.32	1.82E-03
OSM	0.79	8.51E-13	MEF2D	0.42	3.33E-05	YBX3	0.25	9.74E-03	LASP1	-0.33	1.70E-03
LITAF	0.77	2.33E-20	CD58	0.42	5.20E-06				CD99	-0.33	1.74E-06
HBEGF	0.77	4.58E-24	B3GNT5	0.42	3.01E-03				HLA-C	-0.33	2.56E-25
GABARAPL1	0.77	2.13E-23	YPEL5	0.42	1.39E-09				CORO1A	-0.33	5.06E-05
CD69	0.76	3.55E-09	ZMIZ1	0.41	9.04E-05				ADD3	-0.34	2.93E-02
DSE	0.75	7.39E-15	ETV3	0.41	1.21E-02				NADK	-0.34	4.21E-02
THBD	0.75	3.40E-07	PLIN2	0.41	1.66E-02				GMFG	-0.34	3.47E-07
PNPLA8	0.74	6.84E-17	HIF1A	0.41	6.24E-07				LIMD2	-0.34	1.59E-05
NR4A1	0.73	5.77E-19	AL691403.1	0.40	9.65E-06				MAR11	-0.34	8.69E-03
FOSB	0.72	1.09E-59	HMOX1	0.40	4.32E-03				ANP32A	-0.35	2.37E-04
ATF3	0.72	2.69E-16	ZBTB43	0.40	2.48E-03				CAPG	-0.35	1.08E-04
AREG	0.72	6.60E-05	TUBA1B	0.40	2.08E-04				ARPC1B	-0.35	6.47E-05
CXCR4	0.71	2.07E-20	ADGRE2	0.40	3.01E-03				CCNY	-0.35	1.77E-05
PIF	0.71	1.09E-30	JUN	0.40	9.20E-05				MEGF9	-0.36	1.58E-03
NR4A2	0.68	7.43E-23	RFX2	0.39	1.14E-03				CNBP	-0.36	7.42E-06
INSIG1	0.68	9.47E-10	METRN1	0.38	2.89E-04				CLEC4E	-0.36	1.77E-03
EREG	0.68	4.51E-08	SKI	0.38	9.64E-03				WASF2	-0.36	1.68E-05
EMP1	0.67	2.15E-09	TREM1	0.38	8.43E-08				CPPED1	-0.37	2.92E-04
ITGA5	0.67	1.06E-11	DENND5A	0.38	2.08E-04				RAB27A	-0.37	1.58E-03
SKIL	0.66	7.47E-12	JMJD1C	0.38	6.53E-06				ARPC5	-0.37	3.50E-09
HMGA1	0.66	5.41E-12	CLEC7A	0.38	1.67E-13				PRKCB	-0.37	2.90E-06
PHLDA1	0.65	4.93E-08	CD83	0.38	1.95E-09				RASGRP4	-0.37	5.24E-03
RUNX3	0.64	1.73E-08	RYBP	0.38	5.69E-04				RCSO1	-0.38	1.79E-04
DHFR	0.64	1.02E-08	MAML3	0.38	3.88E-03				DUSP6	-0.38	1.40E-10
CSRN1	0.63	9.73E-17	EAF1	0.38	8.32E-05				GTF2I	-0.38	3.54E-04
LCP2	0.63	7.19E-14	ANXA1	0.38	3.82E-16				SELL	-0.38	1.19E-09
SLC2A3	0.61	3.28E-16	SH3BP5	0.38	1.58E-03				CD52	-0.39	8.56E-10
BTG1	0.60	1.59E-31	NLRP3	0.37	1.46E-04				AC020656.1	-0.39	1.03E-03
TUBB6	0.60	1.79E-06	MAPKAPK2	0.37	1.01E-03				LYZ	-0.40	1.12E-41
AC016831.6	0.59	2.95E-06	SYAP1	0.36	5.23E-03				ITGAL	-0.41	4.26E-04
HLA-DQA2	0.59	2.51E-02	TOP1	0.36	7.53E-04				PLEK	-0.41	4.51E-18
MX2	0.58	1.54E-06	CHSY1	0.36	3.27E-02				NFKBIZ	-0.41	9.15E-06
CTNBN1	0.58	6.08E-11	AC245128.3	0.36	4.87E-07				LYST	-0.42	3.48E-06
VEGFA	0.58	5.04E-07	MAP1LC3B	0.35	5.52E-05				CNN2	-0.43	1.59E-09
RUNX1	0.57	2.13E-09	NAPA	0.35	1.58E-04				MNDA	-0.43	1.17E-14
TOM1	0.57	2.98E-10	RHOB	0.35	1.89E-03				ARL6IP5	-0.43	2.92E-09
CHMP1B	0.56	6.97E-15	ETF1	0.35	2.00E-04				CYBB	-0.45	4.31E-17
ABCA1	0.56	2.29E-03	CXCL16	0.35	9.15E-06				RIPOR2	-0.45	1.42E-05
SLA	0.55	1.43E-06	PDIA3	0.35	2.00E-04				ZFP36L1	-0.47	3.41E-11
ARL4A	0.55	1.39E-09	RAB1A	0.34	8.97E-06				MAP3K1	-0.47	5.03E-05
WARS	0.54	1.07E-07	GAB2	0.34	2.29E-02				POU2F2	-0.48	5.49E-09
NFKB1	0.54	8.35E-06	RARA	0.34	1.01E-03				TMEM176B	-0.50	1.87E-09
BCL2A1	0.54	2.16E-10	CDKN1A	0.34	2.68E-07				IL17RA	-0.55	5.06E-11
NRIP1	0.53	3.00E-05	DUSP2	0.34	2.12E-03				RPS26	-0.57	2.28E-25
CXCL8	0.50	2.87E-06	C9orf72	0.32	4.54E-04				IRF1	-0.58	3.32E-11
ELL2	0.49	2.68E-04	AC020916.1	0.32	2.29E-02				TMEM176A	-0.63	6.98E-12
PLAUR	0.49	1.15E-19	MAP3K2	0.32	3.79E-03				SMAP2	-0.66	2.70E-23
HSP90B1	0.48	1.27E-09	TGFB1	0.32	9.86E-04				HLA-DQB1	-0.70	1.09E-27
ERCC1	0.48	3.50E-06	AHR	0.32	3.67E-04				NCF1	-0.80	1.21E-19
ANKRD28	0.48	1.89E-05	EZR	0.31	5.33E-03				THBS1	-0.84	3.92E-13
LDLR	0.48	6.53E-06	ATF4	0.31	1.72E-05				TXNIP	-1.15	1.03E-59
FAM102B	0.48	1.55E-03	ENSA	0.31	4.60E-04						
CCL3L1	0.48	3.20E-02	RAPGEF1	0.31	4.78E-02						
GPCPD1	0.48	2.95E-10	UBE2J1	0.30	3.30E-03						

Table S5.3: Differential genes in T cells

Naïve T up-regulated			Naïve T up-regulated			CD4+ Memory up-regulated			CD8+ Memory up-regulated		
Gene	Log2 fold change	P _{adj}	Gene	Log2 fold change	P _{adj}	Gene	Log2 fold change	P _{adj}	Gene	Log2 fold change	P _{adj}
KLF6	1.06	6.91E-46	AMBRA1	0.29	6.89E-05	MYADM	0.81	3.77E-19	MYADM	0.76	6.68E-24
RGCC	0.87	1.96E-29	EIF4G2	0.29	9.71E-05	KLF6	0.75	2.50E-36	RGCC	0.50	6.06E-11
TNFAIP3	0.84	7.94E-55	MYLIP	0.29	3.86E-03	RGCC	0.61	7.55E-19	NFKBIA	0.47	1.27E-13
NFKBIA	0.79	5.44E-26	PRDM2	0.29	1.82E-02	LMNA	0.53	7.03E-06	KLF6	0.44	1.46E-15
MYADM	0.79	2.14E-12	SRSF7	0.29	1.71E-04	IRF1	0.50	2.81E-08	TRBC1	0.41	1.11E-05
IRF1	0.77	3.07E-24	NKTR	0.29	9.60E-03	TNFAIP3	0.48	9.46E-29	MCL1	0.41	1.92E-07
YPEL5	0.73	1.52E-29	EZR	0.29	9.07E-03	YPEL5	0.47	2.19E-12	IRF1	0.41	7.71E-09
ANXA1	0.72	2.14E-12	MBP	0.29	8.61E-03	GATA3	0.44	2.30E-05	JUND	0.37	3.73E-19
HSPA5	0.71	1.24E-20	BTG1	0.29	1.33E-10	CCNL1	0.44	1.90E-07	PPP1R15A	0.37	2.11E-05
ZFP36	0.69	1.12E-20	SF1	0.28	7.11E-06	HSPA5	0.43	1.44E-09	HSPA5	0.36	6.15E-08
JUND	0.62	4.10E-51	PCGF5	0.28	1.21E-02	NFKBIA	0.43	7.11E-08	CCL4	0.35	3.01E-02
CD69	0.59	3.16E-16	GNAS	0.28	6.24E-08	BHLHE40	0.43	4.12E-04	YPEL5	0.34	4.52E-08
IRF2BP2	0.57	2.26E-20	RPL9	0.28	1.30E-27	JUND	0.40	2.91E-26	TSC22D3	0.33	2.82E-07
GPR183	0.57	1.92E-09	CDC42SE2	0.27	2.18E-04	TSC22D3	0.39	7.47E-09	CD69	0.30	1.36E-04
CCNL1	0.54	1.81E-14	CCR7	0.27	2.48E-05	GPR183	0.37	1.05E-04	AC016831.6	0.29	2.30E-02
BIRC3	0.53	2.24E-09	SIK3	0.27	3.72E-02	C16orf54	0.35	1.79E-02	BCLAF1	0.29	6.08E-05
ZNF331	0.53	1.72E-08	LRRFIP1	0.27	2.24E-03	PPP1R15A	0.35	4.80E-05	TNFAIP3	0.26	7.01E-07
ARL4C	0.53	1.72E-20	ZFP36L1	0.27	3.59E-03	IRF2BP2	0.32	2.97E-06	CALR	0.26	8.04E-04
PTGER4	0.53	1.62E-08	EEF2	0.27	6.73E-07	FOSB	0.31	2.73E-03	NR4A2	0.26	1.08E-03
IFI6	0.51	9.34E-05	EIF3A	0.27	3.00E-03	BCLAF1	0.28	3.72E-04	HSP90B1	0.25	9.27E-04
PPP1R15A	0.50	4.85E-08	DDX3X	0.27	6.50E-04	EIF5A	0.28	6.31E-03			
DDX21	0.49	1.21E-08	CFLAR	0.27	1.75E-02	HLA-B	0.27	9.79E-21	CD8+ Memory down-regulated		
SOC53	0.48	1.39E-06	KLF12	0.27	2.74E-02	AMBRA1	0.27	2.36E-03	FAM107B	-0.25	1.18E-03
AHNAK	0.48	2.43E-06	GSPT1	0.27	1.58E-02	HSP90B1	0.27	3.48E-03	PDE4D	-0.25	7.08E-03
STK17B	0.48	1.94E-17	RUNX3	0.27	3.29E-03	CALR	0.27	4.04E-03	GNLY	-0.26	3.16E-03
PIM2	0.47	1.71E-10	AC013394.1	0.26	2.36E-04	RUNX3	0.26	2.06E-03	KIF2A	-0.26	3.46E-02
SBD5	0.47	5.80E-11	PRRC2C	0.26	2.35E-04				CD3G	-0.26	1.32E-03
PDCD4	0.47	5.59E-09	KDM2A	0.26	3.71E-03	CD4+ Memory down-regulated			PRF1	-0.26	5.18E-03
RGS1	0.45	4.35E-03	STK17A	0.26	2.14E-02	FOS	0.26	2.36E-05	ITGB1	-0.26	4.99E-03
SLC2A3	0.43	5.62E-08	SRSF11	0.26	6.81E-04	RASAL2	0.25	1.17E-02	GMFG	-0.27	2.20E-02
TSC22D3	0.43	1.23E-12	VIM	0.26	6.73E-06	LCP1	-0.25	2.71E-03	PPIA	-0.27	3.19E-07
HSP90B1	0.43	1.05E-07	C12orf57	0.26	8.16E-03	YWHAB	-0.26	3.12E-04	YWHAB	-0.27	2.98E-05
RBM39	0.42	1.43E-12	TMEM123	0.26	8.51E-05	CNN2	-0.27	3.99E-03	DEK	-0.27	1.39E-02
HSP90AA1	0.42	4.21E-12	MALAT1	0.26	1.06E-12	CD74	-0.27	2.54E-03	S100A6	-0.27	3.40E-04
BCLAF1	0.41	2.82E-09	SLC38A1	0.25	1.38E-03	ADD3	-0.28	2.51E-03	TRBC2	-0.27	1.34E-05
AC016831.6	0.41	8.67E-04	CCSER2	0.25	4.35E-02	HP1BP3	-0.28	6.28E-03	LCP1	-0.28	1.24E-05
IL6ST	0.41	8.00E-07	SRSF5	0.25	9.05E-06	TCF7	-0.29	6.22E-05	CNN2	-0.29	1.13E-03
HLA-B	0.39	5.84E-43	HNRNPUL1	0.25	8.08E-04	CAST	-0.29	3.48E-03	IL32	-0.31	1.19E-11
SFPQ	0.38	2.43E-08				TXNIP	-0.29	1.93E-04	SSBP3	-0.32	1.45E-04
EIF1	0.38	6.32E-27	Naïve T down-regulated			AQP3	-0.29	2.84E-02	HP1BP3	-0.35	6.03E-05
RASAL2	0.38	1.72E-08	RPS26	-0.69	1.25E-44	MT-ND3	-0.29	7.17E-90	JUNB	-0.35	8.00E-09
S100A4	0.38	1.86E-05				HLA-C	-0.29	2.95E-16	CD2	-0.35	7.79E-05
TAGLN2	0.37	1.49E-06	Regulatory T up-regulated			PTPRC	-0.30	6.31E-07	PLEK	-0.38	2.37E-04
SRGN	0.37	6.73E-07	JUND	0.66	9.29E-04	ETS1	-0.30	1.06E-05	IL7R	-0.39	7.06E-06
DUSP2	0.36	3.71E-03	EIF1	0.58	9.29E-04	ACTB	-0.34	2.59E-14	NSA2	-0.39	8.28E-08
ARID4B	0.36	7.59E-06	HLA-B	0.53	8.38E-05	MAF	-0.35	2.59E-02	KLRB1	-0.44	3.12E-04
EIF5A	0.36	1.55E-04				SELL	-0.38	1.32E-04	CAST	-0.44	4.58E-08
WHAMM	0.35	4.05E-04	Regulatory T down-regulated			RPS26	-1.01	5.36E-62	HLA-C	-0.54	4.92E-76
SYNE2	0.35	4.22E-03	RPS26	-0.82	9.29E-04				RPS26	-1.19	5.24E-111
CXCR4	0.35	4.29E-14									
RNF125	0.34	1.32E-04	Aberrant T up-regulated								
RSRC2	0.34	6.28E-05	GNG2	1.09	4.98E-04						
RHOH	0.34	7.03E-05	PPP1R16B	0.93	1.53E-03						
CALR	0.34	4.90E-05	FAM129A	0.88	3.04E-02						
DDX24	0.34	4.92E-05	ATP2B1	0.87	5.21E-03						
BTG2	0.34	5.93E-06	SYTL3	0.79	1.48E-02						
FOS	0.34	1.13E-07	TGFB1	0.79	1.91E-02						
PIM1	0.33	4.05E-04	TNFAIP3	0.76	1.53E-03						
MORF4L1	0.33	3.30E-06	AC016831.6	0.75	1.08E-05						
BCL2	0.33	7.18E-04	NFKBIA	0.73	3.00E-02						
FTH1	0.32	1.43E-12	NFKB1	0.72	3.00E-02						
MTRNR2L8	0.32	4.27E-14	SIK3	0.69	6.41E-03						
N4BP2L2	0.32	1.08E-03	JUND	0.53	1.48E-02						
ARHGAP15	0.32	1.33E-03	HLA-B	0.44	1.49E-02						
IRS2	0.31	2.19E-04									
SH3BGR13	0.31	4.86E-07	Aberrant T down-regulated								
PDE4B	0.31	4.12E-03	MT-ND3	-0.26	1.05E-06						
HSPA8	0.31	5.89E-06	HLA-C	-0.50	2.61E-02						
YWHAZ	0.30	1.05E-07	RIPOR2	-0.54	4.21E-02						
ITPKB	0.30	4.90E-04									
NDFIP1	0.30	1.98E-04									
SP100	0.30	2.59E-03									
S1PR1	0.30	2.59E-04									
FNBP4	0.30	2.03E-03									
SMCHD1	0.30	8.44E-04									
SPOCK2	0.29	1.37E-06									

B cells up-regulated			B cells down-regulated			NK cells up-regulated			NK cells down-regulated		
Gene	Log2 fold change	P _{adj}	Gene	Log2 fold change	P _{adj}	Gene	Log2 fold change	P _{adj}	Gene	Log2 fold change	P _{adj}
FOSB	0.89	1.53E-16	ACTB	-0.25	2.05E-03	CD69	1.45	5.49E-58	IL2RG	-0.25	4.55E-04
NFKB1	0.80	2.80E-05	IGHM	-0.32	1.02E-03	NFKBIA	1.37	8.20E-67	SPN	-0.25	2.51E-02
MYADM	0.66	3.48E-04	SEPT6	-0.33	3.92E-02	FOSB	1.37	3.02E-39	CCL5	-0.26	5.21E-04
TNFAIP3	0.60	5.56E-04	ZFP36L1	-0.35	1.24E-02	PPP1R15A	1.32	3.73E-38	RAC2	-0.26	3.12E-05
CDKN1A	0.59	5.33E-04	FCMR	-0.38	8.06E-03	MYADM	1.18	3.18E-31	ARPC5	-0.26	4.54E-02
CD69	0.50	3.94E-06	BCL11A	-0.38	3.38E-02	JUN	1.16	3.60E-28	COX6C	-0.27	6.67E-03
SLC38A1	0.50	1.51E-03	RHOB	-0.40	3.93E-02	TSC22D3	0.99	8.40E-32	CAPZB	-0.27	6.67E-03
REL	0.50	3.68E-05	RCSO1	-0.47	7.33E-04	TNFAIP3	0.92	2.63E-43	SH3KBP1	-0.28	9.07E-03
YPEL5	0.49	2.00E-07	RPS26	-0.52	4.83E-09	NFKBIZ	0.90	4.95E-10	NSA2	-0.29	1.48E-02
NR4A2	0.49	5.33E-04	SELL	-0.58	1.64E-05	JUND	0.87	7.30E-60	PPP1R18	-0.29	1.80E-02
AC016831.6	0.48	1.62E-02				KLF6	0.84	4.82E-28	IKZF1	-0.29	4.83E-03
PPP1R15A	0.48	9.02E-05				IER5	0.84	5.50E-15	MT-ND3	-0.29	3.33E-51
RASGEF1B	0.48	2.10E-03				FOS	0.82	3.59E-26	ACTN4	-0.29	6.67E-03
CD83	0.47	1.01E-07				ZC3H12A	0.81	6.30E-12	TPM3	-0.29	1.29E-06
HLA-B	0.46	1.09E-21				CCL4	0.78	4.87E-08	MYH9	-0.29	5.10E-03
IER2	0.44	3.39E-03				YPEL5	0.78	2.18E-23	SEPTO9	-0.30	1.57E-02
MOB3A	0.44	2.65E-03				MCL1	0.78	1.85E-16	RGS2	-0.30	1.05E-02
CHMP1B	0.42	1.17E-02				RGCC	0.74	5.38E-09	S100A4	-0.30	1.50E-05
MCL1	0.42	1.27E-02				DUSP1	0.71	1.18E-14	KLRB1	-0.31	9.76E-04
KDM6B	0.41	2.32E-02				MAFF	0.70	5.33E-10	ARL6IP5	-0.32	3.32E-03
GPR183	0.41	2.00E-02				IER2	0.70	5.60E-15	ANP32B	-0.33	6.53E-04
HSP90B1	0.40	8.12E-03				BHLHE40	0.69	3.73E-16	CD3E	-0.34	2.92E-02
DDX3Y	0.40	1.71E-02				NFKB1	0.61	1.05E-04	GNLY	-0.34	9.17E-14
C16orf72	0.39	2.12E-02				RG51	0.60	1.62E-05	MYL12A	-0.34	2.10E-04
TRIR	0.38	1.09E-02				IRF1	0.59	1.18E-14	FLNA	-0.35	1.65E-03
NFKBID	0.38	3.36E-02				CCNL1	0.56	2.01E-08	SET	-0.35	1.17E-06
FOS	0.38	5.56E-04				IFITM3	0.51	5.03E-05	ADD3	-0.35	8.91E-04
JUND	0.37	4.51E-13				PMAI1	0.50	5.16E-04	ID2	-0.36	2.52E-05
CCNL1	0.37	7.00E-03				DDX3X	0.49	1.12E-06	ACTR2	-0.36	7.05E-06
DDX3X	0.36	1.92E-02				REL	0.46	5.00E-06	TLN1	-0.36	9.48E-05
HSP90AA1	0.36	7.54E-03				PTGER4	0.45	2.69E-04	GNAI2	-0.36	2.99E-06
RBM39	0.35	1.93E-04				TAGLN2	0.45	9.40E-08	TMBIM6	-0.36	1.88E-05
ARID4B	0.35	4.21E-02				EIF1	0.43	1.02E-20	LCP1	-0.38	2.98E-08
MAP3K8	0.34	4.87E-02				DNAJA1	0.42	8.01E-04	FAM107B	-0.39	3.88E-05
BCLAF1	0.33	4.68E-03				AC016831.6	0.42	8.17E-03	S100A6	-0.39	2.10E-04
UBC	0.33										

Bibliography

- An Efficacy and Safety Study of Ustekinumab in Participants With Active Nonradiographic Axial Spondyloarthritis* [Online]. Available: <https://clinicaltrials.gov/ct2/show/NCT02407223> [Accessed 22/08/2022].
- A Study to Evaluate the Efficacy and Safety of Ustekinumab in the Treatment of Anti-TNF(Alpha) Refractory Participants With Active Radiographic Axial Spondyloarthritis* [Online]. Available: <https://clinicaltrials.gov/ct2/show/NCT02438787> [Accessed 22/08/2022].
- Abdelrahman, M. H., Mahdy, S., Khanjar, I. A., Siam, A. M., Malallah, H. A., et al. 2012. Prevalence of HLA-B27 in Patients with Ankylosing Spondylitis in Qatar. *Int J Rheumatol*, 2012, 860213.
- Adams, H.-P. & Koziol, J. A. 1995. Prediction of binding to MHC class I molecules. *Journal of Immunological Methods*, 185, 181-190.
- Akar, S., Igci, Y. Z., Sari, I., Pala, E., Geyik, E., et al. 2017. Do major histocompatibility complex tag single nucleotide polymorphisms accurately identify HLA-B27 in the Turkish population? *International journal of rheumatic diseases*, 20, 2035-2039.
- Akassou, A., Yacoubi, H., Jamil, A., Dakka, N., Amzazi, S., et al. 2015. Prevalence of HLA-B27 in Moroccan healthy subjects and patients with ankylosing spondylitis and mapping construction of several factors influencing AS diagnosis by using multiple correspondence analysis. *Rheumatology International*, 35, 1889-1894.
- Akkoç, N., Yarkan, H., Kenar, G. & Khan, M. A. 2017. Ankylosing Spondylitis: HLA-B*27-Positive Versus HLA-B*27-Negative Disease. *Current Rheumatology Reports*, 19, 26.
- Aksenova, A. Y., Zhuk, A. S., Lada, A. G., Zotova, I. V., Stepchenkova, E. I., et al. 2021. Genome Instability in Multiple Myeloma: Facts and Factors. *Cancers* [Online], 13.
- Alvarez-Navarro, C. & Lopez De Castro, J. A. 2014. ERAP1 structure, function and pathogenetic role in ankylosing spondylitis and other MHC-associated diseases. *Mol Immunol*, 57, 12-21.
- Amariuta, T., Ishigaki, K., Sugishita, H., Ohta, T., Koido, M., et al. 2020. Improving the trans-ancestry portability of polygenic risk scores by prioritizing variants in predicted cell-type-specific regulatory elements. *Nature Genetics*, 52, 1346-1354.
- Anderson, C. A., Boucher, G., Lees, C. W., Franke, A., D'amato, M., et al. 2011. Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. *Nature genetics*, 43, 246-252.
- Andrés, A. M., Dennis, M. Y., Kretzschmar, W. W., Cannons, J. L., Lee-Lin, S.-Q., et al. 2010. Balancing selection maintains a form of ERAP2 that undergoes nonsense-mediated decay and affects antigen presentation. *PLoS genetics*, 6, e1001157.

- Anne Constanze, R., Anja, W., Fabian, P., Xenofon, B., Frank, B., et al. 2023. Comparison of patients with axial PsA and patients with axSpA and concomitant psoriasis: an analysis of the German register RABBIT-SpA. *RMD Open*, 9, e002837.
- Antoniou, A. N., Ford, S., Taurog, J. D., Butcher, G. W. & Powis, S. J. 2004. Formation of HLA-B27 homodimers and their relationship to assembly kinetics. *Journal of Biological Chemistry*, 279, 8895-8902.
- Appel, H., Kuhne, M., Spiekermann, S., Ebhardt, H., Grozdanovic, Z., et al. 2006a. Immunohistologic analysis of zygapophyseal joints in patients with ankylosing spondylitis. *Arthritis & Rheumatism*, 54, 2845-2851.
- Appel, H., Loddenkemper, C., Grozdanovic, Z., Ebhardt, H., Dreimann, M., et al. 2006b. Correlation of histopathological findings and magnetic resonance imaging in the spine of patients with ankylosing spondylitis. *Arthritis research & therapy*, 8, 1-7.
- Appel, H., Ruiz-Heiland, G., Listing, J., Zwerina, J., Herrmann, M., et al. 2009. Altered skeletal expression of sclerostin and its link to radiographic progression in ankylosing spondylitis. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 60, 3257-3262.
- Arévalo, M., Gratacós Masmitjà, J., Moreno, M., Calvet, J., Orellana, C., et al. 2018. Influence of HLA-B27 on the Ankylosing Spondylitis phenotype: results from the REGISPONSER database. *Arthritis research & therapy*, 20, 1-6.
- Aslam, M. L., Boison, S. A., Lillehammer, M., Norris, A. & Gjerde, B. 2020. Genome-wide association mapping and accuracy of predictions for amoebic gill disease in Atlantic salmon (*Salmo salar*). *Scientific Reports*, 10, 6435.
- Atagunduz, P., Aydin, S. Z., Bahadir, C., Erer, B. & Direskeneli, H. 2010. Determinants of Early Radiographic Progression in Ankylosing Spondylitis. *The Journal of Rheumatology*, 37, 2356.
- Atzeni, F., Beasley, M., Dean, L. E., Jones, G., Gibson, J., et al. 2013. Do patients with ankylosing spondylitis have an excess prevalence of chronic widespread pain?: Results from the Scotland and Ireland Registry for Ankylosing Spondylitis (SIRAS) and the MUSICIAN study. *Arthritis & Rheumatism*, 65, S652-S652.
- Aydin, S. Z., Maksymowych, W. P., Bennett, A. N., McGonagle, D., Emery, P. & Marzo-Ortega, H. 2012. Validation of the ASAS criteria and definition of a positive MRI of the sacroiliac joint in an inception cohort of axial spondyloarthritis followed up for 8 years. *Ann Rheum Dis*, 71, 56-60.
- Babaie, F., Mohammadi, H., Hemmatzadeh, M., Ebrazeh, M., Torkamandi, S., et al. 2020. Evaluation of ERAP1 gene single nucleotide polymorphisms in immunomodulation of pro-inflammatory and anti-inflammatory cytokines profile in ankylosing spondylitis. *Immunology Letters*, 217, 31-38.

- Bacchiega, A. B. S., Balbi, G. G. M., Ochtrop, M. L. G., De Andrade, F. A., Levy, R. A. & Baraliakos, X. 2017. Ocular involvement in patients with spondyloarthritis. *Rheumatology*, 56, 2060-2067.
- Baeten, D., Kruithof, E., Van Den Bosch, F., Demetter, P., Van Damme, N., et al. 2001. Immunomodulatory effects of anti-tumor necrosis factor alpha therapy on synovium in spondylarthropathy: histologic findings in eight patients from an open-label pilot study. *Arthritis Rheum*, 44, 186-95.
- Baeten, D., Ostergaard, M., Wei, J. C., Sieper, J., Jarvinen, P., et al. 2018. Risankizumab, an IL-23 inhibitor, for ankylosing spondylitis: results of a randomised, double-blind, placebo-controlled, proof-of-concept, dose-finding phase 2 study. *Ann Rheum Dis*, 77, 1295-1302.
- Bakirci, S., Dabague, J., Eder, L., Mcgonagle, D. & Aydin, S. Z. 2020. The role of obesity on inflammation and damage in spondyloarthritis: a systematic literature review on body mass index and imaging. *Clin Exp Rheumatol*, 38, 144-8.
- Bakland, G., Nossent, H. C. & Gran, J. T. 2005. Incidence and prevalence of ankylosing spondylitis in Northern Norway. *Arthritis Rheum*, 53, 850-5.
- Ball, J. 1971. Enthesopathy of rheumatoid and ankylosing spondylitis. *Annals of the rheumatic diseases*, 30, 213.
- Bang, S., Donnelly, C. R., Luo, X., Toro-Moreno, M., Tao, X., et al. 2021. Activation of GPR37 in macrophages confers protection against infection-induced sepsis and pain-like behaviour in mice. *Nature Communications*, 12, 1704.
- Bang, S., Xie, Y.-K., Zhang, Z.-J., Wang, Z., Xu, Z.-Z. & Ji, R.-R. 2018. GPR37 regulates macrophage phagocytosis and resolution of inflammatory pain. *The Journal of clinical investigation*, 128, 3568-3582.
- Baraliakos, X. & Braun, J. 2015. Non-radiographic axial spondyloarthritis and ankylosing spondylitis: what are the similarities and differences? *RMD Open*, 1, e000053.
- Baraliakos, X., Coates, L. C. & Braun, J. 2015. The involvement of the spine in psoriatic arthritis. *Clin Exp Rheumatol*, 33, S31-5.
- Barrett, J. C., Fry, B., Maller, J. & Daly, M. J. 2004. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*, 21, 263-265.
- Bauer, O., Sharir, A., Kimura, A., Hantisteanu, S., Takeda, S. & Groner, Y. 2015. Loss of Osteoblast Runx3 Produces Severe Congenital Osteopenia. *Molecular and Cellular Biology*, 35, 1097-1109.
- Benavent, D., Plasencia-Rodríguez, C., Franco-Gómez, K., Nieto, R., Monjo-Henry, I., et al. 2020. Axial spondyloarthritis and axial psoriatic arthritis: similar or different disease spectrum? *Therapeutic Advances in Musculoskeletal Disease*, 12, 1759720X20971889.

- Benjamini, Y. & Hochberg, Y. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57, 289-300.
- Beurel, E., Grieco, S. F. & Jope, R. S. 2015. Glycogen synthase kinase-3 (GSK3): Regulation, actions, and diseases. *Pharmacology & Therapeutics*, 148, 114-131.
- Bjornevik, K., Cortese, M., Healy, B. C., Kuhle, J., Mina, M. J., et al. 2022. Longitudinal analysis reveals high prevalence of Epstein-Barr virus associated with multiple sclerosis. *Science*, 375, 296-301.
- Blair, H. A. & Deeks, E. D. 2017. Abatacept: A Review in Rheumatoid Arthritis. *Drugs*, 77, 1221-1233.
- Blumberg, B. 1958. Bernard connor's description of the pathology of ankyloeing spondylitis. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 1, 553-563.
- Bogers, L., Kuiper, K. L., Smolders, J., Rip, J. & Van Luijn, M. M. 2023. Epstein–Barr virus and genetic risk variants as determinants of T-bet+ B cell-driven autoimmune diseases. *Immunology Letters*, 261, 66-74.
- Bollow, M., Fischer, T., Reißhauer, H., Backhaus, M., Sieper, J., et al. 2000. Quantitative analyses of sacroiliac biopsies in spondyloarthropathies: T cells and macrophages predominate in early and active sacroiliitis—cellularity correlates with the degree of enhancement detected by magnetic resonance imaging. *Annals of the rheumatic diseases*, 59, 135-140.
- Boonen, A., Van Der Heijde, D., Landewé, R., Guillemin, F., Spoorenberg, A., et al. 2003. Costs of ankylosing spondylitis in three European countries: the patient's perspective. *Annals of the Rheumatic Diseases*, 62, 741.
- Bouziat, R., Hinterleitner, R., Brown, J. J., Stencel-Baerenwald, J. E., Ikizler, M., et al. 2017. Reovirus infection triggers inflammatory responses to dietary antigens and development of celiac disease. *Science*, 356, 44-50.
- Bowness, P. 2015. HLA-B27. *Annual review of immunology*, 33.
- Boyce, B. F., Yamashita, T., Yao, Z., Zhang, Q., Li, F. & Xing, L. 2005. Roles for NF- κ B and c-Fos in osteoclasts. *Journal of bone and mineral metabolism*, 23, 11-15.
- Boyle, L. H. & Hill Gaston, J. S. 2003. Breaking the rules: the unconventional recognition of HLA-B27 by CD4+ T lymphocytes as an insight into the pathogenesis of the spondyloarthropathies. *Rheumatology*, 42, 404-412.
- Braakenburg, A. M., De Valk, H. W., De Boer, J. & Rothova, A. 2008. Human leukocyte antigen-B27-associated uveitis: long-term follow-up and gender differences. *Am J Ophthalmol*, 145, 472-9.

- Braun, J., Baraliakos, X., Brandt, J., Listing, J., Zink, A., et al. 2005. Persistent clinical response to the anti-TNF- α antibody infliximab in patients with ankylosing spondylitis over 3 years. *Rheumatology*, 44, 670-676.
- Braun, J., Baraliakos, X., Deodhar, A., Baeten, D., Sieper, J., et al. 2017. Effect of secukinumab on clinical and radiographic outcomes in ankylosing spondylitis: 2-year results from the randomised phase III MEASURE 1 study. *Ann Rheum Dis*, 76, 1070-1077.
- Braun, J., Baraliakos, X., Golder, W., Brandt, J., Rudwaleit, M., et al. 2003. Magnetic resonance imaging examinations of the spine in patients with ankylosing spondylitis, before and after successful therapy with infliximab: evaluation of a new scoring system. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 48, 1126-1136.
- Braun, J., Bollow, M., Neure, L., Seipelt, E., Seyrekbasan, F., et al. 1995. Use of immunohistologic and in situ hybridization techniques in the examination of sacroiliac joint biopsy specimens from patients with ankylosing spondylitis. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 38, 499-505.
- Braun, J. & Sieper, J. 2007. Ankylosing spondylitis. *The Lancet*, 369, 1379-1390.
- Brenner, L. N., Mercader, J. M., Robertson, C. C., Cole, J., Chen, L., et al. 2020. Analysis of Glucocorticoid-Related Genes Reveal CCHCR1 as a New Candidate Gene for Type 2 Diabetes. *Journal of the Endocrine Society*, 4, bvaa121.
- Brophy, S., Hickey, S., Menon, A., Taylor, G., Bradbury, L., et al. 2004. Concordance of disease severity among family members with ankylosing spondylitis? *The Journal of Rheumatology*, 31, 1775.
- Brown, A. C., Cohen, C. J., Mielczarek, O., Migliorini, G., Costantino, F., et al. 2023. Comprehensive epigenomic profiling reveals the extent of disease-specific chromatin states and informs target discovery in ankylosing spondylitis. *Cell Genomics*, 3.
- Brown, M. A., Laval, S. H., Brophy, S. & Calin, A. 2000. Recurrence risk modelling of the genetic susceptibility to ankylosing spondylitis. *Ann Rheum Dis*, 59, 883-6.
- Brown, M. A. & Wordsworth, B. P. 2017. Genetics in ankylosing spondylitis - Current state of the art and translation into clinical outcomes. *Best Practice & Research in Clinical Rheumatology*, 31, 763-776.
- Calderon, D., Nguyen, M. L. T., Mezger, A., Kathiria, A., Müller, F., et al. 2019. Landscape of stimulation-responsive chromatin across diverse human immune cells. *Nature Genetics*, 51, 1494-1505.
- Calin, A., Garrett, S., Whitelock, H., Kennedy, L. G., O'hea, J., et al. 1994. A new approach to defining functional ability in ankylosing spondylitis: the development of the Bath Ankylosing Spondylitis Functional Index. *J rheumatol*, 21, 2281-2285.

- Cantini, F., Nannini, C., Cassara, E., Kaloudi, O. & Niccoli, L. 2015. Uveitis in Spondyloarthritis: An Overview. *J Rheumatol Suppl*, 93, 27-9.
- Capelusnik, D., Ramiro, S., Schneeberger, E. E. & Citera, G. 2021. Peripheral arthritis and higher disease activity lead to more functional impairment in axial spondyloarthritis: Longitudinal analysis from ESPAXIA. *Seminars in Arthritis and Rheumatism*, 51, 553-558.
- Carithers, L. J. & Moore, H. M. 2015. The Genotype-Tissue Expression (GTEx) Project. *Biopreserv Biobank*, 13, 307-8.
- Castellanos-Rubio, A., Fernandez-Jimenez, N., Kratchmarov, R., Luo, X., Bhagat, G., et al. 2016. A long noncoding RNA associated with susceptibility to celiac disease. *Science*, 352, 91-95.
- Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M. & Lee, J. J. 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience*, 4.
- Chava, K. R., Karpurapu, M., Wang, D., Bhanoori, M., Kundumani-Sridharan, V., et al. 2009. CREB-Mediated IL-6 Expression Is Required for 15(S)-Hydroxyeicosatetraenoic Acid-Induced Vascular Smooth Muscle Cell Migration. *Arteriosclerosis, Thrombosis, and Vascular Biology*, 29, 809-815.
- Chen, C., Peng, J., Ma, S., Ding, Y., Huang, T., et al. 2021a. Ribosomal protein S26 serves as a checkpoint of T-cell survival and homeostasis in a p53-dependent manner. *Cellular & Molecular Immunology*, 18, 1844-1846.
- Chen, C. & Zhang, X. 2015. ERAP1 variants are associated with ankylosing spondylitis in East Asian population: a new Chinese case-control study and meta-analysis of published series. *International Journal of Immunogenetics*, 42, 168-173.
- Chen, J., Wright, K., Davis, J. M., Jeraldo, P., Marietta, E. V., et al. 2016. An expansion of rare lineage intestinal microbes characterizes rheumatoid arthritis. *Genome Medicine*, 8, 43.
- Chen, L., Wang, Y.-F., Liu, L., Bielowka, A., Ahmed, R., et al. 2020. Genome-wide assessment of genetic risk for systemic lupus erythematosus and disease severity. *Human Molecular Genetics*, 29, 1745-1756.
- Chen, S., Zhang, J., Yu, W. B., Zhuang, J. C., Xiao, W., et al. 2021b. Eomesodermin in CD4(+)T cells is essential for Ginkgolide K ameliorating disease progression in experimental autoimmune encephalomyelitis. *Int J Biol Sci*, 17, 50-61.
- Chen, W.-S., Chang, Y.-S., Lin, K.-C., Lai, C.-C., Wang, S.-H., et al. 2012. Association of serum interleukin-17 and interleukin-23 levels with disease activity in Chinese patients with ankylosing spondylitis. *Journal of the Chinese Medical Association*, 75, 303-308.
- Chen, W. M., Chen, Y. M., Jiang, S. Y., Tao, Y. P., Hong, Y. G., et al. 2021c. LncRNA POT1-AS1 accelerates the progression of gastric cancer by serving as a competing

- endogenous RNA of microRNA-497-5p to increase PDK3 expression. *J Gastrointest Oncol*, 12, 2728-2742.
- Chimenti, M. S., Conigliaro, P., Navarini, L., Martina, F. M., Peluso, G., et al. 2020. Demographic and clinical differences between ankylosing spondylitis and non-radiographic axial spondyloarthritis: results from a multicentre retrospective study in the Lazio region of Italy. *Clinical and experimental rheumatology*, 38, 88-93.
- Choi, E. Y., Lee, M. & Lee, C. S. 2020a. Uveitis occurrence in patients with ankylosing spondylitis according to the type of tumour necrosis factor inhibitor: A cohort study of 175 patients. *Clin Exp Rheumatol*, 38, 1132-7.
- Choi, S. W., García-González, J., Ruan, Y., Wu, H. M., Porras, C., et al. 2023. PRSet: Pathway-based polygenic risk score analyses and software. *PLoS genetics*, 19, e1010624.
- Choi, S. W., Mak, T. S.-H. & O'reilly, P. F. 2020b. Tutorial: a guide to performing polygenic risk score analyses. *Nature protocols*, 15, 2759-2772.
- Choi, S. W. & O'reilly, P. F. 2019. PRSice-2: Polygenic Risk Score software for biobank-scale data. *GigaScience*, 8.
- Chung, H. Y., Machado, P., Van Der Heijde, D., D'agostino, M. A. & Dougados, M. 2011. HLA-B27 positive patients differ from HLA-B27 negative patients in clinical presentation and imaging: results from the DESIR cohort of patients with recent onset axial spondyloarthritis. *Ann Rheum Dis*, 70, 1930-6.
- Clarke, L., Zheng-Bradley, X., Smith, R., Kulesha, E., Xiao, C., et al. 2012. The 1000 Genomes Project: data management and community access. *Nature Methods*, 9, 459-462.
- Cleynen, I., Boucher, G., Jostins, L., Schumm, L. P., Zeissig, S., et al. 2016. Inherited determinants of Crohn's disease and ulcerative colitis phenotypes: a genetic association study. *The Lancet*, 387, 156-167.
- Colbert, R. A., Rowland-Jones, S. L., Mcmichael, A. J. & Frelinger, J. A. 1994. Differences in peptide presentation between B27 subtypes: The importance of the P1 side chain in maintaining high affinity peptide binding to B★2703. *Immunity*, 1, 121-130.
- Colonna, M. 2007. All roads lead to CARD9. *Nature Immunology*, 8, 554-555.
- Conceição, I., Luscher-Dias, T., Queiroz, L. R., De Melo, A. G. B., Machado, C. R., et al. 2022. Metformin Treatment Modulates Long Non-Coding RNA Isoforms Expression in Human Cells. *Noncoding RNA*, 8.
- Cortes, A. & Brown, M. A. 2011. Promise and pitfalls of the Immunochip. *Arthritis Research & Therapy*, 13, 101.

- Cortes, A., Hadler, J., Pointon, J. P., Robinson, P. C., Karaderi, T., et al. 2013a. Identification of multiple risk variants for ankylosing spondylitis through high-density genotyping of immune-related loci.
- Cortes, A., Hadler, J., Pointon, J. P., Robinson, P. C., Karaderi, T., et al. 2013b. Identification of multiple risk variants for ankylosing spondylitis through high-density genotyping of immune-related loci. *Nature Genetics*, 45, 730-+.
- Costantino, F. & Breban, M. 2023. Family studies: A useful tool to better understand spondyloarthritis. *Joint Bone Spine*, 90, 105588.
- Costantino, F., Talpin, A., Evnouchidou, I., Kadi, A., Leboime, A., et al. 2015. ERAP1 gene expression is influenced by nonsynonymous polymorphisms associated with predisposition to spondyloarthritis. *Arthritis & rheumatology*, 67, 1525-1534.
- Cresswell, P., Bangia, N., Dick, T. & Diedrich, G. 1999. The nature of the MHC class I peptide loading complex. *Immunological Reviews*, 172, 21-28.
- Csukasi, F., Duran, I., Barad, M., Barta, T., Gudernova, I., et al. 2018. The PTH/PTHrP-SIK3 pathway affects skeletogenesis through altered mTOR signaling. *Science Translational Medicine*, 10, eaat9356.
- Cunha, A. R., Aguiar, R., Santos, M. J., Afreixo, V. & Barcelos, A. 2021. POS0950 ARE THERE GENDER DIFFERENCES IN AXIAL SPONDYLOARTHRITIS: DATA FROM A PORTUGUESE SPONDYLOARTHRITIS COHORT. *Annals of the Rheumatic Diseases*, 80, 738-738.
- D'amato, M., Fiorillo, M. T., Carcassi, C., Mathieu, A., Zuccarelli, A., et al. 1995. Relevance of residue 116 of HLA-B27 in determining susceptibility to ankylosing spondylitis. *European Journal of Immunology*, 25, 3199-3201.
- Daltroy, L., Larson, M., Roberts, N. & Liang, M. 1990. A modification of the Health Assessment Questionnaire for the spondyloarthropathies. *The Journal of Rheumatology*, 17, 946-950.
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., et al. 2011. The variant call format and VCFtools. *Bioinformatics*, 27, 2156-2158.
- Danoy, P., Pryce, K., Hadler, J., Bradbury, L. A., Farrar, C., et al. 2010. Association of variants at 1q32 and STAT3 with ankylosing spondylitis suggests genetic overlap with Crohn's disease. *PLoS genetics*, 6, e1001195.
- Daoussis, D., Liossis, S. N. C., Solomou, E. E., Tsanakti, A., Bounia, K., et al. 2010. Evidence that Dkk-1 is dysfunctional in ankylosing spondylitis. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 62, 150-158.
- Dare, S., Mackay, D. F. & Pell, J. P. 2015. Relationship between smoking and obesity: a cross-sectional study of 499,504 middle-aged adults in the UK general population. *PloS one*, 10, e0123579.

- De Bosscher, K., Vanden Berghe, W. & Haegeman, G. 2003. The Interplay between the Glucocorticoid Receptor and Nuclear Factor- κ B or Activator Protein-1: Molecular Mechanisms for Gene Repression. *Endocrine Reviews*, 24, 488-522.
- De Lange, K. M., Moutsianas, L., Lee, J. C., Lamb, C. A., Luo, Y., et al. 2017. Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease. *Nature genetics*, 49, 256-261.
- De Vos, M., Cuvelier, C., Mielants, H., Veys, E., Barbier, F. & Elewaut, A. 1989. Ileocolonoscopy in seronegative spondylarthropathy. *Gastroenterology*, 96, 339-44.
- Dean, L. E., Jones, G. T., Macdonald, A. G., Downham, C., Sturrock, R. D. & Macfarlane, G. J. 2014. Global prevalence of ankylosing spondylitis. *Rheumatology*, 53, 650-657.
- Degerman, E., Ahmad, F., Chung, Y. W., Guirguis, E., Omar, B., et al. 2011. From PDE3B to the regulation of energy homeostasis. *Current Opinion in Pharmacology*, 11, 676-682.
- Delaneau, O., Zagury, J.-F., Robinson, M. R., Marchini, J. L. & Dermitzakis, E. T. 2019. Accurate, scalable and integrative haplotype estimation. *Nature Communications*, 10, 5436.
- Dennehy, K. M., Willment, J. A., Williams, D. L. & Brown, G. D. 2009. Reciprocal regulation of IL-23 and IL-12 following co-activation of Dectin-1 and TLR signaling pathways. *European Journal of Immunology*, 39, 1379-1386.
- Deveci, H., Turk, A. C., Ozmen, Z. C., Demir, A. K. & Say Coskun, S. U. 2019. Biological and genetic evaluation of IL-23/IL-17 pathway in ankylosing spondylitis patients. *Cent Eur J Immunol*, 44, 433-439.
- Di, G., Kong, L., Zhao, Q. & Ding, T. 2018. MicroRNA-146a knockdown suppresses the progression of ankylosing spondylitis by targeting dickkopf 1. *Biomedicine & Pharmacotherapy*, 97, 1243-1249.
- Ding, Y., Yang, Y. & Xue, L. 2023. Immune cells and their related genes provide a new perspective on the common pathogenesis of ankylosing spondylitis and inflammatory bowel diseases. *Frontiers in Immunology*, 14.
- Dobрева, Z. G., Stanilova, S. A. & Miteva, L. D. 2009. The influence of JNK and P38 MAPK inhibition on IL-12P40 and IL-23 production depending on IL12B promoter polymorphism. *Cellular & Molecular Biology Letters*, 14, 609-621.
- Dominguez Castro, P., Harkin, G., Hussey, M., Christopher, B., Kiat, C., et al. 2020. Prevalence of coexisting autoimmune thyroidal diseases in coeliac disease is decreasing. *United European Gastroenterology Journal*, 8, 148-156.
- Dominguez Castro, P., Harkin, G., Hussey, M., Christopher, B., Kiat, C., et al. 2017. Changes in Presentation of Celiac Disease in Ireland From the 1960s to 2015. *Clinical Gastroenterology and Hepatology*, 15, 864-871.e3.

- Doran, M. F., Brophy, S., Mackay, K., Taylor, G. & Calin, A. 2003. Predictors of longterm outcome in ankylosing spondylitis. *The Journal of Rheumatology*, 30, 316.
- Dougados, M. & Baeten, D. 2011. Spondyloarthritis. *The Lancet*, 377, 2127-2137.
- Doward, L. C., Spoorenberg, A., Cook, S. A., Whalley, D., Helliwell, P. S., et al. 2003. Development of the ASQoL: a quality of life instrument specific to ankylosing spondylitis. *Annals of the Rheumatic Diseases*, 62, 20.
- Dubois, P. C. A., Trynka, G., Franke, L., Hunt, K. A., Romanos, J., et al. 2010. Multiple common variants for celiac disease influencing immune gene expression. *Nature Genetics*, 42, 295-302.
- Duncan, L., Shen, H., Gelaye, B., Meijsen, J., Ressler, K., et al. 2019. Analysis of polygenic risk score usage and performance in diverse human populations. *Nat Commun*, 10, 3328.
- Duvallet, E., Semerano, L., Assier, E., Falgarone, G. & Boissier, M. C. 2011. Interleukin-23: a key cytokine in inflammatory diseases. *Ann Med*, 43, 503-11.
- Duverle, D. A., Yotsukura, S., Nomura, S., Aburatani, H. & Tsuda, K. 2016. CellTree: an R/bioconductor package to infer the hierarchical structure of cell populations from single-cell RNA-seq data. *BMC Bioinformatics*, 17, 363.
- Ebringer, A., Rashid, T., Wilson, C., Ptaszynska, T. & Fielder, M. 2006. Ankylosing spondylitis as an auto-immune disease linked to intestinal Klebsiella infection: prospects for a new therapeutic approach. *Current Rheumatology Reviews*, 2, 55-68.
- El Maghraoui, A. 2004. Osteoporosis and ankylosing spondylitis. *Joint Bone Spine*, 71, 291-295.
- El Maghraoui, A. 2011. Extra-articular manifestations of ankylosing spondylitis: prevalence, characteristics and therapeutic implications. *Eur J Intern Med*, 22, 554-60.
- El Maghraoui, A., Borderie, D., Cherruau, B., Edouard, R., Dougados, M. & Roux, C. 1999. Osteoporosis, body composition, and bone turnover in ankylosing spondylitis. *The Journal of rheumatology*, 26, 2205-2209.
- Elkayam, O., Ophir, J., Yaron, M. & Caspi, D. 2000. Psoriatic Arthritis: Interrelationships between Skin and Joint Manifestations Related to Onset, Course and Distribution. *Clinical Rheumatology*, 19, 301-305.
- Ellinghaus, D., Jostins, L., Spain, S. L., Cortes, A., Bethune, J., et al. 2016. Analysis of five chronic inflammatory diseases identifies 27 new associations and highlights disease-specific patterns at shared loci. *Nature Genetics*, 48, 510-+.
- Erber, R., Eichelsbacher, U., Powajbo, V., Korn, T., Djonov, V., et al. 2006. EphB4 controls blood vascular morphogenesis during postnatal angiogenesis. *The EMBO journal*, 25, 628-641.

- Evans, D. M., Spencer, C. C., Pointon, J. J., Su, Z., Harvey, D., et al. 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, 761-767.
- Evnouchidou, I., Weimershaus, M., Saveanu, L. & Van Endert, P. 2014. ERAP1–ERAP2 Dimerization Increases Peptide-Trimming Efficiency. *The Journal of Immunology*, 193, 901.
- Fadista, J., Manning, A. K., Florez, J. C. & Groop, L. 2016. The (in)famous GWAS P-value threshold revisited and updated for low-frequency variants. *European Journal of Human Genetics*, 24, 1202-1205.
- Fairfax, B. P., Humburg, P., Makino, S., Naranbhai, V., Wong, D., et al. 2014. Innate Immune Activity Conditions the Effect of Regulatory Variants upon Monocyte Gene Expression. *Science*, 343, 1246949.
- Fam, A. G., Rubenstein, J. D., Chin-Sang, H. & Leung, F. Y. K. 1985. Computed tomography in the diagnosis of early ankylosing spondylitis. *Arthritis & Rheumatism*, 28, 930-937.
- Fan, J.-B., Miyauchi, S., Xu, H.-Z., Liu, D., Kim, L. J. Y., et al. 2020. Type I Interferon Regulates a Coordinated Gene Network to Enhance Cytotoxic T Cell–Mediated Tumor Killing. *Cancer Discovery*, 10, 382-393.
- Fang, H.-S., Wang, C.-C., Chao, C.-Y., Fan, W.-L., Su, S.-C. & Wu, Y.-R. 2022a. Association of ITPKB, IL1R2 and COQ7 with Parkinson's disease in Taiwan. *Journal of the Formosan Medical Association*, 121, 679-686.
- Fang, Y., Liu, J., Long, Y., Wen, J., Huang, D. & Xin, L. 2022b. Knockdown of circular RNA hsa_circ_0003307 inhibits synovial inflammation in ankylosing spondylitis by regulating the PI3K/AKT pathway. *Advances in Clinical and Experimental Medicine: Official Organ Wroclaw Medical University*.
- Faour, W. H., Gomi, K. & Kennedy, C. R. J. 2008. PGE2 induces COX-2 expression in podocytes via the EP4 receptor through a PKA-independent mechanism. *Cellular Signalling*, 20, 2156-2164.
- Fawzy, R. M., Ganab, S. S., Said, E. A. & Fouad, N. A. 2016. Serum Level of Interleukin-37 and Expression of Its mRNA in Ankylosing Spondylitis Patients: Possible Role in Osteoporosis. *The Egyptian journal of immunology*, 23, 19-29.
- Feld, J., Chandran, V., Haroon, N., Inman, R. & Gladman, D. 2018. Axial disease in psoriatic arthritis and ankylosing spondylitis: a critical comparison. *Nature Reviews Rheumatology*, 14, 363-371.
- Feldtkeller, E., Khan, M. A., Van Der Heijde, D., Van Der Linden, S. & Braun, J. 2003. Age at disease onset and diagnosis delay in HLA-B27 negative vs. positive patients with ankylosing spondylitis. *Rheumatology International*, 23, 61-66.

- Ferrand, A., Al Nabhani, Z., Tapias, N. S., Mas, E., Hugot, J.-P. & Barreau, F. 2019. NOD2 Expression in Intestinal Epithelial Cells Protects Toward the Development of Inflammation and Associated Carcinogenesis. *Cellular and Molecular Gastroenterology and Hepatology*, 7, 357-369.
- Floss, D. M., Schröder, J., Franke, M. & Scheller, J. 2015. Insights into IL-23 biology: From structure to function. *Cytokine & Growth Factor Reviews*, 26, 569-578.
- Franck, H. & Keck, E. 1993. Serum osteocalcin and vitamin D metabolites in patients with ankylosing spondylitis. *Annals of the Rheumatic Diseases*, 52, 343-346.
- Frantz, C., Portier, A., Etcheto, A., Monnet, D., Brezin, A., et al. 2019. Acute anterior uveitis in spondyloarthritis: a monocentric study of 301 patients. *Clin Exp Rheumatol*, 37, 26-31.
- Freedman, M. L., Monteiro, A. N. A., Gayther, S. A., Coetzee, G. A., Risch, A., et al. 2011. Principles for the post-GWAS functional characterization of cancer risk loci. *Nature Genetics*, 43, 513-518.
- Freeston, J., Barkham, N., Hensor, E., Emery, P. & Fraser, A. 2007. Ankylosing spondylitis, HLA-B27 positivity and the need for biologic therapies. *Joint Bone Spine*, 74, 140-143.
- Fu, D., Song, X., Hu, H., Sun, M., Li, Z. & Tian, Z. 2016. Downregulation of RUNX3 moderates the frequency of Th17 and Th22 cells in patients with psoriasis. *Molecular medicine reports*, 13, 4606-4612.
- Gabriel, S. B., Schaffner, S. F., Nguyen, H., Moore, J. M., Roy, J., et al. 2002. The Structure of Haplotype Blocks in the Human Genome. *Science*, 296, 2225-2229.
- Galloway, M. & Machado, P. M. 2022. Axial spondyloarthritis. *Medicine*, 50, 159-166.
- Gao, S., Xu, T., Liang, W., Xun, C., Deng, Q., et al. 2020. Association of rs27044 and rs30187 polymorphisms in endoplasmic reticulum aminopeptidase 1 gene and ankylosing spondylitis susceptibility: A meta-analysis. *Int J Rheum Dis*, 23, 499-510.
- Gao, S., Xu, T., Mao, C., Cheng, J., Xun, C., et al. 2022. Lack of Associations between Endoplasmic Reticulum Aminopeptidase 2 Gene Polymorphisms and Ankylosing Spondylitis: A Meta-analysis with Trial Sequential Analysis. *Immunological Investigations*, 51, 715-726.
- García, R., Rognan, D., Lamas, J. R., Marina, A. & De Castro, J. a. L. 1998. An HLA-B27 polymorphism (B*2710) that is critical for T-cell recognition has limited effects on peptide specificity. *Tissue Antigens*, 51, 1-9.
- Garrett, S., Jenkinson, T., Kennedy, L. G., Whitelock, H., Gaisford, P. & Calin, A. 1994. A new approach to defining disease status in ankylosing spondylitis: the Bath Ankylosing Spondylitis Disease Activity Index. *The Journal of rheumatology*, 21, 2286-2291.
- Gaur, T., Lengner, C. J., Hovhannisyanyan, H., Bhat, R. A., Bodine, P. V., et al. 2005. Canonical WNT signaling promotes osteogenesis by directly stimulating Runx2 gene expression. *Journal of Biological Chemistry*, 280, 33132-33140.

- Gauzzi, M. C., Velazquez, L., Mckendry, R., Mogensen, K. E., Fellous, M. & Pellegrini, S. 1996. Interferon- α -dependent activation of Tyk2 requires phosphorylation of positive regulatory tyrosines by another kinase. *Journal of Biological Chemistry*, 271, 20494-20500.
- Gettler, K., Levantovsky, R., Moscati, A., Giri, M., Wu, Y., Hsu, N.Y., Chuang, L.S., Sazonovs, A., Venkateswaran, S., Korie, U. and Chasteau, C., 2021. Common and rare variant prediction and penetrance of IBD in a large, multi-ethnic, health system-based biobank cohort. *Gastroenterology*, 160(5), pp.1546-1557.
- Gilgil, E., Kaçar, C., Tuncer, T. & Bütün, B. 2005. The association of syndesmophytes with vertebral bone mineral density in patients with ankylosing spondylitis. *The Journal of Rheumatology*, 32, 292-294.
- Ginhoux, F., Yalin, A., Dutertre, C. A. & Amit, I. 2022. Single-cell immunology: Past, present, and future. *Immunity*, 55, 393-404.
- Ginsburg, W. W. & Cohen, M. D. 1983. Peripheral arthritis in ankylosing spondylitis. A review of 209 patients followed up for more than 20 years. *Mayo Clinic proceedings*, 58, 593-596.
- Glintborg, B., Sørensen, I. J., Østergaard, M., Dreyer, L., Mohamoud, A. A., et al. 2017. Ankylosing Spondylitis versus Nonradiographic Axial Spondyloarthritis: Comparison of Tumor Necrosis Factor Inhibitor Effectiveness and Effect of HLA-B27 Status. An Observational Cohort Study from the Nationwide DANBIO Registry. *The Journal of Rheumatology*, 44, 59.
- Gómez-García, I., Ladehesa-Pineda, M. L., Puche-Larrubia, M. Á., Ortega-Castro, R., Font-Ugalde, P., et al. 2021. Uveitis as the first symptom in spondyloarthritis and its association with the evolution of the disease. Results from the REGISPONSER registry. *Joint Bone Spine*, 88, 105136.
- Goto, Y., Hattori, A., Ishii, Y. & Tsujimoto, M. 2006. Reduced activity of the hypertension-associated Lys528Arg mutant of human adipocyte-derived leucine aminopeptidase (A-LAP)/ER-aminopeptidase-1. *FEBS Letters*, 580, 1833-1838.
- Gracey, E., Hromadova, D., Lim, M., Qiayum, Z., Zeng, M., et al. 2020. TYK2 inhibition reduces type 3 immunity and modifies disease progression in murine spondyloarthritis. *J Clin Invest*, 130, 1863-1878.
- Gran, J. T. & Husby, G. The epidemiology of ankylosing spondylitis. *Seminars in arthritis and rheumatism*, 1993. Elsevier, 319-334.
- Gratacos, J., Collado, A., Filella, X., Sanmarti, R., Canete, J., et al. 1994. Serum cytokines (IL-6, TNF- α , IL-1 β and IFN- γ) in ankylosing spondylitis: a close correlation between serum IL-6 and disease activity and severity. *Rheumatology*, 33, 927-931.

- Grisar, J., Bernecker, P. M., Aringer, M., Redlich, K., Sedlak, M., et al. 2002. Ankylosing spondylitis, psoriatic arthritis, and reactive arthritis show increased bone resorption, but differ with regard to bone formation. *The Journal of Rheumatology*, 29, 1430-1436.
- Grivennikov, S. I. & Karin, M. 2010. Dangerous liaisons: STAT3 and NF-kappaB collaboration and crosstalk in cancer. *Cytokine Growth Factor Rev*, 21, 11-9.
- Gross, O., Gewies, A., Finger, K., Schäfer, M., Sparwasser, T., et al. 2006. Card9 controls a non-TLR signalling pathway for innate anti-fungal immunity. *Nature*, 442, 651-656.
- Han, G.-W., Zeng, L.-W., Liang, C.-X., Cheng, B.-L., Yu, B.-S., et al. 2011. Serum levels of IL-33 is increased in patients with ankylosing spondylitis. *Clinical Rheumatology*, 30, 1583-1588.
- Hanson, A. L., Cuddihy, T., Haynes, K., Loo, D., Morton, C. J., et al. 2018. Genetic Variants in ERAP1 and ERAP2 Associated With Immune-Mediated Diseases Influence Protein Expression and the Isoform Profile. *Arthritis & Rheumatology*, 70, 255-265.
- Hara, H., Ishihara, C., Takeuchi, A., Imanishi, T., Xue, L., et al. 2007. The adaptor protein CARD9 is essential for the activation of myeloid cells through ITAM-associated and Toll-like receptors. *Nature Immunology*, 8, 619-629.
- Haroon, N., Tsui, F. W., Uchanska-Ziegler, B., Ziegler, A. & Inman, R. D. 2012. Endoplasmic reticulum aminopeptidase 1 (ERAP1) exhibits functionally significant interaction with HLA-B27 and relates to subtype specificity in ankylosing spondylitis. *Ann Rheum Dis*, 71, 589-95.
- 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, 409-416.
- Harvey, D., Pointon, J. J., Evans, D. M., Karaderi, T., Farrar, C., et al. 2009. Investigating the genetic association between ERAP1 and ankylosing spondylitis. *Human Molecular Genetics*, 18, 4204-4212.
- Haywood, K., Garratt, A. & Dawes, P. 2005. Patient-assessed health in ankylosing spondylitis: a structured review. *Rheumatology*, 44, 577-586.
- Hazlett, J., Stamp, L. K., Merriman, T., Highton, J. & Hessian, P. A. 2012. IL-23R rs11209026 polymorphism modulates IL-17A expression in patients with rheumatoid arthritis. *Genes Immun*, 13, 282-7.
- Hébert, H. L., Bowes, J., Smith, R. L., Flynn, E., Parslew, R., et al. 2015. Identification of loci associated with late-onset psoriasis using dense genotyping of immune-related regions. *British Journal of Dermatology*, 172, 933-939.
- Hernandez, J. M., Floyd, D. H., Weillbaecher, K. N., Green, P. L. & Boris-Lawrie, K. 2008. Multiple facets of junD gene expression are atypical among AP-1 family members. *Oncogene*, 27, 4757-4767.

- Heuft-Dorenbosch, L., Van Tubergen, A., Spoorenberg, A., Landewé, R., Dougados, M., et al. 2004. The influence of peripheral arthritis on disease activity in ankylosing spondylitis patients as measured with the Bath Ankylosing Spondylitis Disease Activity Index. *Arthritis Care & Research*, 51, 154-159.
- Hitomi, Y., Nakatani, K., Kojima, K., Nishida, N., Kawai, Y., et al. 2019a. NFKB1 and MANBA Confer Disease Susceptibility to Primary Biliary Cholangitis via Independent Putative Primary Functional Variants. *Cellular and Molecular Gastroenterology and Hepatology*, 7, 515-532.
- Hitomi, Y., Ueno, K., Kawai, Y., Nishida, N., Kojima, K., et al. 2019b. POGLUT1, the putative effector gene driven by rs2293370 in primary biliary cholangitis susceptibility locus chromosome 3q13.33. *Scientific Reports*, 9, 102.
- Hohler, T., Hug, R., Schneider, P. M., Krummenauer, F., Gripenberg-Lerche, C., et al. 1999. Ankylosing spondylitis in monozygotic twins: studies on immunological parameters. *Ann Rheum Dis*, 58, 435-40.
- Hong, C., Kwan, Y. H., Leung, Y.-Y., Lui, N. L. & Fong, W. 2019. Comparison of ankylosing spondylitis and non-radiographic axial spondyloarthritis in a multi-ethnic Asian population of Singapore. *International Journal of Rheumatic Diseases*, 22, 1506-1511.
- Hong, J., Yan, J., Chen, J., Li, S., Huang, Y., et al. 2021. Identification of key potential targets for TNF- α /TNFR1-related intervertebral disc degeneration by bioinformatics analysis. *Connective Tissue Research*, 62, 531-541.
- Hromadova, D., Elewaut, D., Inman, R. D., Strobl, B. & Gracey, E. 2021. From Science to Success? Targeting Tyrosine Kinase 2 in Spondyloarthritis and Related Chronic Inflammatory Diseases. *Front Genet*, 12, 685280.
- Hsu, Y.-M. S., Zhang, Y., You, Y., Wang, D., Li, H., et al. 2007. The adaptor protein CARD9 is required for innate immune responses to intracellular pathogens. *Nature Immunology*, 8, 198-205.
- Huang, J., Song, G., Yin, Z., Fu, Z. & Zhang, L. 2019. Altered expression of microRNAs targeting Dkk-1 in peripheral blood mononuclear cells of patients with ankylosing spondylitis. *Central-european Journal of Immunology*, 44, 59.
- Huang, T., Yin, L., Wu, J., Gu, J. J., Ding, K., et al. 2017. TNFAIP3 inhibits migration and invasion in nasopharyngeal carcinoma by suppressing epithelial mesenchymal transition. *Neoplasma*, 64, 389-394.
- Huang, X. F., Li, Z., De Guzman, E., Robinson, P., Gensler, L., et al. 2020. Genomewide Association Study of Acute Anterior Uveitis Identifies New Susceptibility Loci. *Invest Ophthalmol Vis Sci*, 61, 3.

- Hulur, I., Gamazon, E. R., Skol, A. D., Xicola, R. M., Llor, X., et al. 2015. Enrichment of inflammatory bowel disease and colorectal cancer risk variants in colon expression quantitative trait loci. *BMC Genomics*, 16, 138.
- Hutchinson, A., Asimit, J. & Wallace, C. 2020. Fine-mapping genetic associations. *Human Molecular Genetics*, 29, R81-R88.
- Ichiyama, K., Sekiya, T., Inoue, N., Tamiya, T., Kashiwagi, I., et al. 2011. Transcription Factor Smad-Independent T Helper 17 Cell Induction by Transforming-Growth Factor- β Is Mediated by Suppression of Eomesodermin. *Immunity*, 34, 741-754.
- International Multiple Sclerosis Genetics, C., Beecham, A. H., Patsopoulos, N. A., Xifara, D. K., Davis, M. F., et al. 2013. Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. *Nature genetics*, 45, 1353-1360.
- International Multiple Sclerosis Genetics, C., Wellcome Trust Case Control, C., Sawcer, S., Hellenthal, G., Pirinen, M., et al. 2011. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature* [Online], 476. [Accessed 2011/08/].
- Ipsos, M. 2021. Healthy Ireland survey 2021: summary report.
- Jaakkola, E., Herzberg, I., Laiho, K., Barnardo, M. C. N. M., Pointon, J. J., et al. 2006. Finnish HLA studies confirm the increased risk conferred by HLA-B27 homozygosity in ankylosing spondylitis. *Annals of the Rheumatic Diseases*, 65, 775.
- Jadav, R. S., Kumar, D., Buwa, N., Ganguli, S., Thampatty, S. R., et al. 2016. Deletion of inositol hexakisphosphate kinase 1 (IP6K1) reduces cell migration and invasion, conferring protection from aerodigestive tract carcinoma in mice. *Cellular Signalling*, 28, 1124-1136.
- Jansen, R., Hottenga, J.-J., Nivard, M. G., Abdellaoui, A., Laport, B., et al. 2017. Conditional eQTL analysis reveals allelic heterogeneity of gene expression. *Human Molecular Genetics*, 26, 1444-1451.
- Janssen, F., Bardoutsos, A. & Vidra, N. 2020. Obesity Prevalence in the Long-Term Future in 18 European Countries and in the USA. *Obesity Facts*, 13, 514-527.
- Jardetzky, T. S., Lane, W. S., Robinson, R. A., Madden, D. R. & Wiley, D. C. 1991. Identification of self peptides bound to purified HLA-B27. *Nature*, 353, 326-9.
- Järvinen, P. 1995. Occurrence of ankylosing spondylitis in a nationwide series of twins. *Arthritis & Rheumatism*, 38, 381-383.
- Jenkinson, T. R., Mallorie, P. A., Whitelock, H. C., Kennedy, L. G., Garrett, S. L. & Calin, A. 1994. Defining spinal mobility in ankylosing spondylitis (AS). The Bath AS Metrology Index. *The Journal of rheumatology*, 21, 1694-1698.

- Jo, S., Won, E. J., Kim, M.-J., Lee, Y. J., Jin, S.-H., et al. 2021. STAT3 phosphorylation inhibition for treating inflammation and new bone formation in ankylosing spondylitis. *Rheumatology*, 60, 3923-3935.
- Jones, G. T., Ratz, T., Dean, L. E., Macfarlane, G. J. & Atzeni, F. 2017. Disease Severity in Never Smokers, Ex-Smokers, and Current Smokers With Axial Spondyloarthritis: Results From the Scotland Registry for Ankylosing Spondylitis. *Arthritis Care & Research*, 69, 1407-1413.
- Jostins, L., Ripke, S., Weersma, R. K., Duerr, R. H., McGovern, D. P., et al. 2012. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature*, 491, 119-24.
- Kachuri, L., Graff, R. E., Smith-Byrne, K., Meyers, T. J., Rashkin, S. R., et al. 2020. Pan-cancer analysis demonstrates that integrating polygenic risk scores with modifiable risk factors improves risk prediction. *Nature Communications*, 11, 6084.
- Kaeley, G. S., Eder, L., Aydin, S. Z., Gutierrez, M. & Bakewell, C. 2018. Dactylitis: A hallmark of psoriatic arthritis. *Seminars in Arthritis and Rheumatism*, 48, 263-273.
- Kampstra, P. 2008. Beanplot: A Boxplot Alternative for Visual Comparison of Distributions. *Journal of Statistical Software, Code Snippets*, 28, 1 - 9.
- Karberg, K., Zochling, J., Sieper, J., Felsenberg, D. & Braun, J. 2005. Bone loss is detected more frequently in patients with ankylosing spondylitis with syndesmophytes. *The Journal of rheumatology*, 32, 1290-1298.
- Kasela, S., Kisand, K., Tserel, L., Kaleviste, E., Remm, A., et al. 2017. Pathogenic implications for autoimmune mechanisms derived by comparative eQTL analysis of CD4+ versus CD8+ T cells. *PLoS genetics*, 13, e1006643.
- Kavanaugh, A., Baraliakos, X., Gao, S., Chen, W., Sweet, K., et al. 2023. Genetic and Molecular Distinctions Between Axial Psoriatic Arthritis and Radiographic Axial Spondyloarthritis: Post Hoc Analyses from Four Phase 3 Clinical Trials. *Advances in Therapy*, 40, 2439-2456.
- Kennedy, A., Waters, E., Rowshanravan, B., Hinze, C., Williams, C., et al. 2022. Differences in CD80 and CD86 transendocytosis reveal CD86 as a key target for CTLA-4 immune regulation. *Nature Immunology*, 23, 1365-1378.
- Kent, W. J., Sugnet, C. W., Furey, T. S., Roskin, K. M., Pringle, T. H., et al. 2002. The human genome browser at UCSC. *Genome research*, 12, 996-1006.
- Kenyon, M., Maguire, S., Rueda Pujol, A., O'shea, F. & Mcmanus, R. 2022. The genetic backbone of ankylosing spondylitis: how knowledge of genetic susceptibility informs our understanding and management of disease. *Rheumatology International*.
- Khan, M. A. 1995. HLA-B27 and its subtypes in world populations. *Current opinion in rheumatology*, 7, 263-269.

- Khera, A. V., Chaffin, M., Zekavat, S. M., Collins, R. L., Roselli, C., et al. 2019. Whole-Genome Sequencing to Characterize Monogenic and Polygenic Contributions in Patients Hospitalized With Early-Onset Myocardial Infarction. *Circulation*, 139, 1593-1602.
- Kim, J.-H., Cheong, H. S., Sul, J. H., Seo, J.-M., Kim, D.-Y., et al. 2014. A genome-wide association study identifies potential susceptibility loci for Hirschsprung disease. *PloS one*, 9, e110292.
- Kimura, A. & Kishimoto, T. 2010. IL-6: Regulator of Treg/Th17 balance. *European Journal of Immunology*, 40, 1830-1835.
- Klasen, C., Meyer, A., Wittekind, P. S., Waqué, I., Nabhani, S. & Kofler, D. M. 2019. Prostaglandin receptor EP4 expression by Th17 cells is associated with high disease activity in ankylosing spondylitis. *Arthritis Research & Therapy*, 21, 159.
- Kleijmeer, M. J., Kelly, A., Geuze, H. J., Slot, J. W., Townsend, A. & Trowsdale, J. 1992. Location of MHC-encoded transporters in the endoplasmic reticulum and cis-Golgi. *Nature*, 357, 342-344.
- Klein, W., Tromm, A., Folwaczny, C., Hagedorn, M., Duerig, N., et al. 2004. A polymorphism of the NFKBIA gene is associated with Crohn's disease patients lacking a predisposing allele of the CARD15 gene. *International Journal of Colorectal Disease*, 19, 153-156.
- Koga, T., Ichinose, K., Kawakami, A. & Tsokos, G. C. 2021. Current insights and future prospects for targeting IL-17 to treat patients with systemic lupus erythematosus. *Frontiers in Immunology*, 11, 624971.
- Koo, A. S., Chiu, R., Soong, J., Dekernion, J. B. & Belldgrun, A. 1992. The expression of C-jun and junB mRNA in renal cell cancer and in vitro regulation by transforming growth factor beta 1 and tumor necrosis factor alpha 1. *The Journal of urology*, 148, 1314-1318.
- Korinfskaya, S., Parameswaran, S., Weirauch, M. T. & Barski, A. 2021. Runx Transcription Factors in T Cells—What Is Beyond Thymic Development? *Frontiers in Immunology*, 12.
- La Borda, J. P. Y., Szczypiorska, M., Bartolomé, N., Campos, J., Flores-Robles, B. J., et al. 2019. Clinical and genetic characteristics of ankylosing spondylitis patients with peripheral arthritis at disease onset. *Clin. Exp. Rheumatol*, 37, 215-221.
- Lambert, S. A., Abraham, G. & Inouye, M. 2019. Towards clinical utility of polygenic risk scores. *Human Molecular Genetics*, 28, R133-R142.
- Landi, M., Maldonado-Ficco, H., Perez-Alamino, R., Maldonado-Cocco, J. A., Citera, G., et al. 2016. Gender differences among patients with primary ankylosing spondylitis and spondylitis associated with psoriasis and inflammatory bowel disease in an iberoamerican spondyloarthritis cohort. *Medicine*, 95, e5652.

- Lange, U., Boss, B., Teichmann, J., Klör, H. & Neeck, G. 2000. Serum amyloid A—an indicator of inflammation in ankylosing spondylitis. *Rheumatology international*, 19, 119-122.
- Lange, U., Stapfer, G., Ditting, T., Geiger, H., Teichmann, J., et al. 2007. Pathologic alterations of the heart and the kidney in patients with ankylosing spondylitis. *European journal of medical research*, 12, 573.
- Laukens, D., Peeters, H., Marichal, D., Vander Cruyssen, B., Mielants, H., et al. 2005. CARD15 gene polymorphisms in patients with spondyloarthropathies identify a specific phenotype previously related to Crohn's disease. *Ann Rheum Dis*, 64, 930-5.
- Lee, E. G., Boone, D. L., Chai, S., Libby, S. L., Chien, M., et al. 2000. Failure to Regulate TNF-Induced NF- κ B and Cell Death Responses in A20-Deficient Mice. *Science*, 289, 2350-2354.
- Lee, K. M., Kang, B. S., Lee, H. L., Son, S. J., Hwang, S. H., et al. 2004. Spinal NF- κ B activation induces COX-2 upregulation and contributes to inflammatory pain hypersensitivity. *Eur J Neurosci*, 19, 3375-81.
- Lekpa, F. K., Farrenq, V., Canouï-Poitaine, F., Paul, M., Chevalier, X., et al. 2012. Lack of efficacy of abatacept in axial spondylarthropathies refractory to tumor-necrosis-factor inhibition. *Joint Bone Spine*, 79, 47-50.
- Lewis, A. C. F., Green, R. C. & Vassy, J. L. 2021. Polygenic risk scores in the clinic: Translating risk into action. *Human Genetics and Genomics Advances*, 2, 100047.
- Li, H. 2011. Tabix: fast retrieval of sequence features from generic TAB-delimited files. *Bioinformatics*, 27, 718-719.
- Li, H., Bradbury, J. A., Dackor, R. T., Edin, M. L., Graves, J. P., et al. 2011a. Cyclooxygenase-2 Regulates Th17 Cell Differentiation during Allergic Lung Inflammation. *American Journal of Respiratory and Critical Care Medicine*, 184, 37-49.
- Li, M.-X., Yeung, J. M. Y., Cherny, S. S. & Sham, P. C. 2012. Evaluating the effective numbers of independent tests and significant p-value thresholds in commercial genotyping arrays and public imputation reference datasets. *Human Genetics*, 131, 747-756.
- Li, Q., Makri, A., Lu, Y., Marchand, L., Grabs, R., et al. 2013. Genome-wide search for exonic variants affecting translational efficiency. *Nature Communications*, 4, 2260.
- Li, S., Schmid, K. T., De Vries, D. H., Korshevniuk, M., Losert, C., et al. 2023. Identification of genetic variants that impact gene co-expression relationships using large-scale single-cell data. *Genome Biology*, 24, 80.
- Li, X., Grisanti, M., Fan, W., Asuncion, F. J., Tan, H. L., et al. 2011b. Dickkopf-1 regulates bone formation in young growing rodents and upon traumatic injury. *Journal of Bone and Mineral Research*, 26, 2610-2621.

- Li, Z., Akar, S., Yarkan, H., Lee, S. K., Çetin, P., et al. 2019. Genome-wide association study in Turkish and Iranian populations identify rare familial Mediterranean fever gene (MEFV) polymorphisms associated with ankylosing spondylitis. *PLoS genetics*, 15, e1008038.
- Li, Z., Ju, X., Silveira, P. A., Abadir, E., Hsu, W.-H., et al. 2019. CD83: activation marker for antigen presenting cells and its therapeutic potential. *Frontiers in immunology*, 10, 1312.
- Li, Z., Van Der Linden, S. M., Khan, M. A., Baumberger, H., Zandwijk, H. V., et al. 2022. Heterogeneity of axial spondyloarthritis: genetics, sex and structural damage matter. *RMD Open*, 8, e002302.
- Li, Z., Wu, X., Leo, P. J., De Guzman, E., Akkoc, N., et al. 2021. Polygenic Risk Scores have high diagnostic capacity in ankylosing spondylitis. *Annals of the rheumatic diseases*.
- Libermann, T. A. & Baltimore, D. 1990. Activation of Interleukin-6 Gene-Expression through the Nf-Kappa-B Transcription Factor. *Molecular and Cellular Biology*, 10, 2327-2334.
- Lin, S. Y., Hsieh, S. C., Lin, Y. C., Lee, C. N., Tsai, M. H., et al. 2012. A whole genome methylation analysis of systemic lupus erythematosus: hypomethylation of the IL10 and IL1R2 promoters is associated with disease activity. *Genes & Immunity*, 13, 214-220.
- Lin, T.-T., Lu, J., Qi, C.-Y., Yuan, L., Li, X.-L., et al. 2015. Elevated serum level of IL-27 and VEGF in patients with ankylosing spondylitis and associate with disease activity. *Clinical and Experimental Medicine*, 15, 227-231.
- Linssen, A. & Meenken, C. 1995. Outcomes of HLA-B27-Positive and HLA-B27-Negative Acute Anterior Uveitis. *American Journal of Ophthalmology*, 120, 351-361.
- Liu, J. Z., Almarri, M. A., Gaffney, D. J., Mells, G. F., Jostins, L., et al. 2012. Dense fine-mapping study identifies new susceptibility loci for primary biliary cirrhosis. *Nature genetics*, 44, 1137-1141.
- Liu, J. Z., Van Sommeren, S., Huang, H., Ng, S. C., Alberts, R., et al. 2015. Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nature genetics*, 47, 979-986.
- Liu, X., Peng, L., Li, D., He, C., Xing, S., et al. 2021. The Impacts of IL1R1 and IL1R2 Genetic Variants on Rheumatoid Arthritis Risk in the Chinese Han Population: A Case–Control Study. *International Journal of General Medicine*, 14, 2147-2159.
- Lo, H. W., Cao, X. Y., Zhu, H. & Ali-Osman, F. 2010. Cyclooxygenase-2 Is a Novel Transcriptional Target of the Nuclear EGFR-STAT3 and EGFRvIII-STAT3 Signaling Axes. *Molecular Cancer Research*, 8, 232-245.

- Londono, J., Santos, A. M., Pena, P., Calvo, E., Espinosa, L. R., et al. 2015. Analysis of HLA-B15 and HLA-B27 in spondyloarthritis with peripheral and axial clinical patterns. *BMJ Open*, 5, e009092.
- Long, F., Wang, T., Li, Q., Xiong, Y. & Zeng, Y. 2022. Association between *Klebsiella pneumoniae* and ankylosing spondylitis: A systematic review and meta-analysis. *International Journal of Rheumatic Diseases*, 25, 422-432.
- Lubberts, E. 2015. The IL-23–IL-17 axis in inflammatory arthritis. *Nature Reviews Rheumatology*, 11, 415-429.
- Lucasson, F., Richette, P., Aouad, K., Ryussen-Witrand, A., Wendling, D., et al. 2022. Prevalence and consequences of psoriasis in recent axial spondyloarthritis: an analysis of the DESIR cohort over 6 years. *RMD Open*, 8, e001986.
- Luxán, G., Stewen, J., Díaz, N., Kato, K., Maney, S. K., et al. 2019. Endothelial EphB4 maintains vascular integrity and transport function in adult heart. *eLife*, 8, e45863.
- Maas, F., Arends, S., Van Der Veer, E., Wink, F., Efde, M., et al. 2016. Obesity Is Common in Axial Spondyloarthritis and Is Associated with Poor Clinical Outcome. *The Journal of Rheumatology*, 43, 383.
- Machiela, M. J. & Chanock, S. J. 2015. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics*, 31, 3555-7.
- Madden, D. R., Gorga, J. C., Strominger, J. L. & Wiley, D. C. 1992. The three-dimensional structure of HLA-B27 at 2.1 Å resolution suggests a general mechanism for tight peptide binding to MHC. *Cell*, 70, 1035-1048.
- Madhi, R., Rahman, M., Taha, D., Linders, J., Merza, M., et al. 2019. Platelet IP6K1 regulates neutrophil extracellular trap-microparticle complex formation in acute pancreatitis. *JCI insight*.
- Magrey, M. & Khan, M. A. 2010. Osteoporosis in Ankylosing Spondylitis. *Current Rheumatology Reports*, 12, 332-336.
- Maguire, S., Fitzgerald, G., Gallagher, P. & O'shea, F. 2022a. Advances in treatment of axial spondyloarthritis are associated with improved patient outcomes: data from the Ankylosing Spondylitis Registry of Ireland (ASRI). *Rheumatology International*, 42, 831-838.
- Maguire, S., Wilson, F., Gallagher, P. & O'shea, F. 2022b. Central Obesity in Axial Spondyloarthritis: The Missing Link to Understanding Worse Outcomes in Women? *The Journal of Rheumatology*, 49, 577.
- Maguire, S., Wilson, F., Gallagher, P. & O'shea, F. 2022c. The toll of unemployment in axial spondyloarthropathy: high prevalence and negative impact on outcomes captured in a national registry. *Scandinavian Journal of Rheumatology*, 51, 300-303.

- Maguire, S., Wilson, F., Gallagher, P. & O'shea, F. D. 2022d. Worse scores but similar patterns of disease activity: interpreting outcomes in women with axial spondyloarthritis. *Scandinavian Journal of Rheumatology*, 1-8.
- Maksymowych, W. P., Chiowchanwisawakit, P., Clare, T., Pedersen, S. J., Østergaard, M. & Lambert, R. G. 2009a. Inflammatory lesions of the spine on magnetic resonance imaging predict the development of new syndesmophytes in ankylosing spondylitis: evidence of a relationship between inflammation and new bone formation. *Arthritis & Rheumatism*, 60, 93-102.
- Maksymowych, W. P., Inman, R. D., Gladman, D. D., Reeve, J. P., Pope, A. & Rahman, P. 2009b. Association of a specific ERAP1/ARTS1 haplotype with disease susceptibility in ankylosing spondylitis. *Arthritis & Rheumatism*, 60, 1317-1323.
- Malynn, B. A. & Ma, A. 2019. A20: A multifunctional tool for regulating immunity and preventing disease. *Cellular Immunology*, 340, 103914.
- Mandola, A. B., Sharfe, N., Nagdi, Z., Dadi, H., Vong, L., et al. 2021. Combined immunodeficiency caused by a novel homozygous NFKB1 mutation. *Journal of Allergy and Clinical Immunology*, 147, 727-733.e2.
- Manry, J., Nédélec, Y., Fava, V. M., Cobat, A., Orlova, M., et al. 2017. Deciphering the genetic control of gene expression following Mycobacterium leprae antigen stimulation. *PLoS genetics*, 13, e1006952.
- Marden, J. R., Mayeda, E. R., Walter, S., Vivot, A., Tchetgen, E. J. T., et al. 2016. Using an Alzheimer's Disease polygenic risk score to predict memory decline in black and white Americans over 14 years of follow-up Running head: AD polygenic risk score predicting memory decline. *Alzheimer disease and associated disorders*, 30, 195.
- Márquez-Luna, C., Loh, P.-R., South Asian Type 2 Diabetes, C., The, S. T. D. C. & Price, A. L. 2017. Multiethnic polygenic risk scores improve risk prediction in diverse populations. *Genetic Epidemiology*, 41, 811-823.
- Martens, A. & Van Loo, G. 2020. A20 at the crossroads of cell death, inflammation, and autoimmunity. *Cold Spring Harbor Perspectives in Biology*, 12, a036418.
- Martín-Esteban, A., Sanz-Bravo, A., Guasp, P., Barnea, E., Admon, A. & López De Castro, J. A. 2017. Separate effects of the ankylosing spondylitis associated ERAP1 and ERAP2 aminopeptidases determine the influence of their combined phenotype on the HLA-B*27 peptidome. *Journal of Autoimmunity*, 79, 28-38.
- Martínez-Nava, G. A., Zamudio-Cuevas, Y., Terrazas-Ontiveros, N. A., Martínez-Flores, K., Espinosa-Morales, R., et al. 2021. A proposed HLA-B*27 screening method for ankylosing spondylitis detection based on tag-single nucleotide polymorphisms: a preliminary study. *Molecular Biology Reports*, 48, 7819-7829.

- Marusina, A. I., Ji-Xu, A., Le, S. T., Toussi, A., Tsoi, L. C., et al. 2023. Cell-Specific and Variant-Linked Alterations in Expression of ERAP1, ERAP2, and LNPEP Aminopeptidases in Psoriasis. *Journal of Investigative Dermatology*, 143, 1157-1167.e10.
- Mau, W., Zeidler, H., Mau, R., Majewski, A., Freyschmidt, J., et al. 1988. Clinical features and prognosis of patients with possible ankylosing spondylitis. Results of a 10-year followup. *The Journal of rheumatology*, 15, 1109-1114.
- Mccarthy, M. I., Abecasis, G. R., Cardon, L. R., Goldstein, D. B., Little, J., et al. 2008. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nature Reviews Genetics*, 9, 356-369.
- Mcgonagle, D., Watad, A., Sharif, K. & Bridgwood, C. 2021. Why Inhibition of IL-23 Lacked Efficacy in Ankylosing Spondylitis. *Front Immunol*, 12, 614255.
- Mease, P. & Van Den Bosch, F. 2021. IL-23 and axial disease: do they come together? *Rheumatology*, 60, iv28-iv33.
- Medeiros, M., Candido, M. F., Valera, E. T. & Brassesco, M. S. 2021. The multifaceted NF- κ B: are there still prospects of its inhibition for clinical intervention in pediatric central nervous system tumors? *Cellular and Molecular Life Sciences*, 78, 6161-6200.
- Mei, Y., Pan, F., Gao, J., Ge, R., Duan, Z., et al. 2011. Increased serum IL-17 and IL-23 in the patient with ankylosing spondylitis. *Clinical rheumatology*, 30, 269-273.
- Melroy-Greif, W. E., Gizer, I. R., Wilhelmsen, K. C. & Ehlers, C. L. 2017. Genetic Influences on Evening Preference Overlap with Those for Bipolar Disorder in a Sample of Mexican Americans and American Indians. *Twin Research and Human Genetics*, 20, 499-510.
- Mermerci Başkan, B., Pekin Doğan, Y., Sivas, F., Bodur, H. & Özoran, K. 2010. The relation between osteoporosis and vitamin D levels and disease activity in ankylosing spondylitis. *Rheumatology international*, 30, 375-381.
- Mevel, R., Draper, J. E., Lie-a-Ling, M., Kouskoff, V. & Lacaud, G. 2019. RUNX transcription factors: orchestrators of development. *Development*, 146, dev148296.
- Mielants, H., Veys, E., Cuvelier, C. & De Vos, M. 1988. Ileocolonoscopy findings in seronegative spondylarthropathies. *Rheumatology*, 27, 95-105.
- Mills, K. H. G. 2023. IL-17 and IL-17-producing cells in protection versus pathology. *Nature Reviews Immunology*, 23, 38-54.
- Min, H. K., Lee, J., Ju, J. H., Park, S.-H. & Kwok, S.-K. 2019. Predictors of Assessment of Spondyloarthritis International Society (ASAS) Health Index in Axial Spondyloarthritis and Comparison of ASAS Health Index between Ankylosing Spondylitis and Nonradiographic Axial Spondyloarthritis: Data from the Catholic Axial Spondyloarthritis Cohort (CASCO). *Journal of Clinical Medicine* [Online], 8.

- Mohanakrishnan, R., Beier, S. & Deodhar, A. 2022. Tofacitinib for the treatment of active ankylosing spondylitis in adults. *Expert Review of Clinical Immunology*, 18, 273-280.
- Montoya, J., Matta, N., Suchon, P., Guzman, M., Lambert, N., et al. 2016. Patients with ankylosing spondylitis have been breast fed less often than healthy controls: a case–control retrospective study. *Annals of the rheumatic diseases*, 75, 879-882.
- Morimoto, A., Ohno, Y., Tatsumi, Y., Nishigaki, Y., Maejima, F., et al. 2012. Impact of smoking cessation on incidence of diabetes mellitus among overweight or normal-weight Japanese men. *Diabetes Research and Clinical Practice*, 96, 407-413.
- Muller, M. & Demeret, C. 2014. CCHCR1 interacts specifically with the E2 protein of human papillomavirus type 16 on a surface overlapping BRD4 binding. *Plos one*, 9, e92581.
- Napolitani, G., Acosta-Rodriguez, E. V., Lanzavecchia, A. & Sallusto, F. 2009. Prostaglandin E2 enhances Th17 responses via modulation of IL-17 and IFN- γ production by memory CD4⁺ T cells. *European Journal of Immunology*, 39, 1301-1312.
- Narumiya, S., Sugimoto, Y. & Ushikubi, F. 1999. Prostanoid Receptors: Structures, Properties, and Functions. *Physiological Reviews*, 79, 1193-1226.
- Ni, J., Wang, P., Yin, K. J., Yang, X. K., Cen, H., et al. 2022. Novel insight into the aetiology of rheumatoid arthritis gained by a cross-tissue transcriptome-wide association study. *RMD Open*, 8.
- Nikiphorou, E. & Ramiro, S. 2020. Work Disability in Axial Spondyloarthritis. *Current Rheumatology Reports*, 22, 55.
- Nogueira, M., Puig, L. & Torres, T. 2020. JAK Inhibitors for Treatment of Psoriasis: Focus on Selective TYK2 Inhibitors. *Drugs*, 80, 341-352.
- Noisette, A. & Hochberg, M. C. 2018. Abatacept for the treatment of adults with psoriatic arthritis: patient selection and perspectives. *Psoriasis: Targets and Therapy*, 8, 31-39.
- Noren, N. K. & Pasquale, E. B. 2007. Paradoxes of the EphB4 Receptor in Cancer. *Cancer Research*, 67, 3994-3997.
- Nossent, J. C., Johnsen, S. & Bakland, G. 2016. The influence of ERAP1 gene variants on clinical phenotype in ankylosing spondylitis. *Scandinavian Journal of Rheumatology*, 45, 474-479.
- Ochoa, D., Hercules, A., Carmona, M., Suveges, D., Gonzalez-Uriarte, A., et al. 2020. Open Targets Platform: supporting systematic drug–target identification and prioritisation. *Nucleic Acids Research*, 49, D1302-D1310.
- Oh, M., Kim, J. G., Baek, I.-P., Ju, J. H., Sup & Sup 2022. Development of Spondyloarthritis After COVID-19 in HLA-B27-Positive Monozygotic Twins: Case Reports With Single Cell Transcriptome Profiling. *jrd*, 30, 58-64.

- Paladini, F., Fiorillo, M. T., Vitulano, C., Tedeschi, V., Piga, M., et al. 2018. An allelic variant in the intergenic region between ERAP1 and ERAP2 correlates with an inverse expression of the two genes. *Scientific Reports*, 8, 10398.
- Paramarta, J. E., De Rycke, L., Ambarus, C. A., Tak, P. P. & Baeten, D. 2013. Undifferentiated spondyloarthritis vs ankylosing spondylitis and psoriatic arthritis: a real-life prospective cohort study of clinical presentation and response to treatment. *Rheumatology*, 52, 1873-1878.
- Peh, C. A., Burrows, S. R., Barnden, M., Khanna, R., Cresswell, P., et al. 1998. HLA-B27–Restricted Antigen Presentation in the Absence of Tapasin Reveals Polymorphism in Mechanisms of HLA Class I Peptide Loading. *Immunity*, 8, 531-542.
- Pérez Alamino, R., Maldonado Cocco, J. A., Citera, G., Arturi, P., Vazquez-Mellado, J., et al. 2011. Differential Features Between Primary Ankylosing Spondylitis and Spondylitis Associated with Psoriasis and Inflammatory Bowel Disease. *The Journal of Rheumatology*, 38, 1656-1660.
- Poddubnyy, D., Haibel, H., Listing, J., Märker-Hermann, E., Zeidler, H., et al. 2012. Baseline radiographic damage, elevated acute-phase reactant levels, and cigarette smoking status predict spinal radiographic progression in early axial spondylarthritis. *Arthritis & Rheumatism*, 64, 1388-1398.
- Pouget, J. G., Schizophrenia Working Group of the Psychiatric Genomics, C., Han, B., Wu, Y., Mignot, E., et al. 2019. Cross-disorder analysis of schizophrenia and 19 immune-mediated diseases identifies shared genetic risk. *Human Molecular Genetics*, 28, 3498-3513.
- Pruim, R. J., Welch, R. P., Sanna, S., Teslovich, T. M., Chines, P. S., et al. 2010. LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics*, 26, 2336-2337.
- Przanowski, P., Dabrowski, M., Ellert-Miklaszewska, A., Kloss, M., Mieczkowski, J., et al. 2014. The signal transducers Stat1 and Stat3 and their novel target Jmjd3 drive the expression of inflammatory genes in microglia. *Journal of Molecular Medicine*, 92, 239-254.
- Przepiera-Będzak, H., Fischer, K. & Brzosko, M. 2015. Serum IL-6 and IL-23 levels and their correlation with angiogenic cytokines and disease activity in ankylosing spondylitis, psoriatic arthritis, and SAPHO syndrome. *Mediators of inflammation*, 2015.
- Puel, A., Picard, C., Cypowyj, S., Lilic, D., Abel, L. & Casanova, J.-L. 2010. Inborn errors of mucocutaneous immunity to *Candida albicans* in humans: a role for IL-17 cytokines? *Current Opinion in Immunology*, 22, 467-474.

- Purcell, A. W., Kelly, A. J., Peh, C. A., Dudek, N. L. & Mccluskey, J. 2000. Endogenous and exogenous factors contributing to the surface expression of HLA B27 on mutant APC. *Human Immunology*, 61, 120-130.
- Queiro, R., Morante, I., Cabezas, I. & Acasuso, B. 2015. HLA-B27 and psoriatic disease: a modern view of an old relationship. *Rheumatology*, 55, 221-229.
- Quinn, E. M., Coleman, C., Molloy, B., Dominguez Castro, P., Cormican, P., et al. 2015. Transcriptome analysis of CD4⁺ T cells in coeliac disease reveals imprint of BACH2 and IFN γ regulation. *PloS one*, 10, e0140049.
- Rahmati, S., Li, Q., Rahman, P. & Chandran, V. 2021. Insights into the pathogenesis of psoriatic arthritis from genetic studies. *Seminars in Immunopathology*, 43, 221-234.
- Rebeaud, F., Hailfinger, S., Posevitz-Fejfar, A., Tapernoux, M., Moser, R., et al. 2008. The proteolytic activity of the paracaspase MALT1 is key in T cell activation. *Nature Immunology*, 9, 272-281.
- Reich, K., Gooderham, M., Thaci, D., Crowley, J. J., Ryan, C., et al. 2019. Risankizumab compared with adalimumab in patients with moderate-to-severe plaque psoriasis (IMMvent): a randomised, double-blind, active-comparator-controlled phase 3 trial. *Lancet*, 394, 576-586.
- Ren, C., Li, M., Zheng, Y., Cai, B., Du, W., et al. 2022. Single-cell RNA-seq reveals altered NK cell subsets and reduced levels of cytotoxic molecules in patients with ankylosing spondylitis. *Journal of Cellular and Molecular Medicine*, 26, 1071-1082.
- Reveille, J. D., Sims, A.-M., Danoy, P., Evans, D. M., Leo, P., et al. 2010. Genome-wide association study of ankylosing spondylitis identifies non-MHC susceptibility loci. *Nature genetics*, 42, 123.
- Rich-Garg, N., Danve, A., Choi, D., Vakil-Gilani, K., Akkoc, N., et al. 2021. Assessing rheumatologists' attitudes and utilization of classification criteria for ankylosing spondylitis and axial spondyloarthritis: a global effort. *Clinical Rheumatology*, 40, 949-954.
- Riley, M. J., Ansell, B. M. & Bywaters, E. G. 1971. Radiological manifestations of ankylosing spondylitis according to age at onset. *Ann Rheum Dis*, 30, 138-48.
- Rizzo, A., Ferrante, A., Guggino, G. & Ciccia, F. 2017. Gut inflammation in spondyloarthritis. *Best Practice & Research Clinical Rheumatology*, 31, 863-876.
- Roberts, R. L., Wallace, M. C., Jones, G. T., Van Rij, A. M., Merriman, T. R., et al. 2013. Prevalence of HLA-B27 in the New Zealand population: effect of age and ethnicity. *Arthritis Res Ther*, 15, R158.
- Robinson, M. D., McCarthy, D. J. & Smyth, G. K. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, 26, 139-40.

- Robinson, P. C. & Brown, M. A. 2014. Genetics of ankylosing spondylitis. *Molecular Immunology*, 57, 2-11.
- Robinson, P. C., Claushuis, T. A., Cortes, A., Martin, T. M., Evans, D. M., et al. 2015a. Genetic dissection of acute anterior uveitis reveals similarities and differences in associations observed with ankylosing spondylitis. *Arthritis Rheumatol*, 67, 140-51.
- Robinson, P. C., Costello, M.-E., Leo, P., Bradbury, L. A., Hollis, K., et al. 2015b. ERAP2 is associated with ankylosing spondylitis in HLA-B27-positive and HLA-B27-negative patients. *Annals of the rheumatic diseases*, 74, 1627-1629.
- Robinson, P. C., Van Der Linden, S., Khan, M. A. & Taylor, W. J. 2021. Axial spondyloarthritis: concept, construct, classification and implications for therapy. *Nature Reviews Rheumatology*, 17, 109-118.
- Robinson, P. C., Wordsworth, B. P., Reveille, J. D. & Brown, M. A. 2013. Axial spondyloarthritis: a new disease entity, not necessarily early ankylosing spondylitis. *Annals of the rheumatic diseases*, 72, 162-164.
- Robinson, W. H. & Steinman, L. 2022. Epstein-Barr virus and multiple sclerosis. *Science*, 375, 264-265.
- Romanus, R. & Ydén, S. 1952. Destructive and ossifying spondylitic changes in rheumatoid ankylosing spondylitis (pelvo-spondylitis ossificans). *Acta orthopaedica Scandinavica*, 22, 88-99.
- Rong, J., Li, Q., Zhang, P., Wu, X., Huang, J., et al. 2015. A rare co-segregation-mutation in the insulin receptor substrate 1 gene in one Chinese family with ankylosing spondylitis. *PLoS One*, 10, e0126348.
- Rosas, J., Llinares-Tello, F., Senabre-Gallego, J. M., Barber-Vallés, X., Santos-Soler, G., et al. 2017. Obesity decreases clinical efficacy and levels of adalimumab in patients with ankylosing spondylitis. *Clin Exp Rheumatol*, 35, 145-148.
- Rostami, S., Hoff, M., Brown, M. A., Hveem, K., Holmen, O. L., et al. 2020. Prediction of Ankylosing Spondylitis in the HUNT Study by a Genetic Risk Score Combining 110 Single-nucleotide Polymorphisms of Genome-wide Significance. *The Journal of Rheumatology*, 47, 204.
- Rougerie, P., Largeteau, Q., Megreli, L., Carrette, F., Lejeune, T., et al. 2013. Fam65b Is a New Transcriptional Target of FOXO1 That Regulates RhoA Signaling for T Lymphocyte Migration. *The Journal of Immunology*, 190, 748-755.
- Rubinacci, S., Delaneau, O. & Marchini, J. 2020. Genotype imputation using the positional burrows wheeler transform. *PLoS genetics*, 16, e1009049.
- Rudwaleit, M. & Baeten, D. 2006. Ankylosing spondylitis and bowel disease. *Best Practice & Research Clinical Rheumatology*, 20, 451-471.

- Rudwaleit, M., Van Der Heijde, D., Landewe, R., Listing, J., Akkoc, N., et al. 2009. The development of Assessment of SpondyloArthritis international Society classification criteria for axial spondyloarthritis (part II): validation and final selection. *Ann Rheum Dis*, 68, 777-83.
- Rusman, T., Van Vollenhoven, R. F. & Van Der Horst-Bruinsma, I. E. 2018. Gender Differences in Axial Spondyloarthritis: Women Are Not So Lucky. *Current Rheumatology Reports*, 20, 35.
- Saadi, A., Dang, J., Shan, S., Ladjouze-Rezig, A., Lefkir-Tafiani, S., et al. 2019. Ankylosing spondylitis: analysis of gene-gene interactions between IL-12 β , JAK2, and STAT3 in Han Chinese and Algerian cohorts. *Cent Eur J Immunol*, 44, 65-74.
- Saevarsdottir, S., Olafsdottir, T. A., Ivarsdottir, E. V., Halldorsson, G. H., Gunnarsdottir, K., et al. 2020. FLT3 stop mutation increases FLT3 ligand level and risk of autoimmune thyroid disease. *Nature*, 584, 619-623.
- Safran, M., Rosen, N., Twik, M., Barshir, R., Stein, T. I., et al. 2021. The GeneCards Suite. In: ABUGESSAISA, I. & KASUKAWA, T. (eds.) *Practical Guide to Life Science Databases*. Singapore: Springer Nature Singapore.
- Sanz-Bravo, A., Campos, J., Mazariegos, M. S. & López De Castro, J. A. 2015. Dominant role of the ERAP1 polymorphism R528K in shaping the HLA-B27 peptidome through differential processing determined by multiple peptide residues. *Arthritis & rheumatology*, 67, 692-701.
- Saric, T., Chang, S.-C., Hattori, A., York, I. A., Markant, S., et al. 2002. An IFN- γ -induced aminopeptidase in the ER, ERAP1, trims precursors to MHC class I-presented peptides. *Nature Immunology*, 3, 1169-1176.
- Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. 2015. Spatial reconstruction of single-cell gene expression data. *Nature Biotechnology*, 33, 495-502.
- Saveanu, L., Carroll, O., Lindo, V., Del Val, M., Lopez, D., et al. 2005. Concerted peptide trimming by human ERAP1 and ERAP2 aminopeptidase complexes in the endoplasmic reticulum. *Nature immunology*, 6, 689-697.
- Schmitt, Anthony d., Hu, M., Jung, I., Xu, Z., Qiu, Y., et al. 2016. A Compendium of Chromatin Contact Maps Reveals Spatially Active Regions in the Human Genome. *Cell Reports*, 17, 2042-2059.
- Shaath, H. & Alajez, N. M. 2021. Identification of PBMC-based molecular signature associational with COVID-19 disease severity. *Heliyon*, 7, e06866.
- Sheibanie, A. F., Yen, J.-H., Khayrullina, T., Emig, F., Zhang, M., et al. 2007. The Proinflammatory Effect of Prostaglandin E₂ in Experimental Inflammatory Bowel Disease Is Mediated through the IL-23 $\rightarrow$ IL-17 Axis. *The Journal of Immunology*, 178, 8138-8147.

- Sherry, S. T., Ward, M. H., Kholodov, M., Baker, J., Phan, L., et al. 2001. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Research*, 29, 308-311.
- Singh, G., Lawrence, A., Agarwal, V., Misra, R. & Aggarwal, A. 2008. Higher Prevalence of Extra-Articular Manifestations in Ankylosing Spondylitis With Peripheral Arthritis. *JCR: Journal of Clinical Rheumatology*, 14.
- Singh, R. S. P., Pradhan, V., Roberts, E. S., Scaramozza, M., Kieras, E., et al. 2021. Safety and Pharmacokinetics of the Oral TYK2 Inhibitor PF-06826647: A Phase I, Randomized, Double-Blind, Placebo-Controlled, Dose-Escalation Study. *Clin Transl Sci*, 14, 671-682.
- Siva, N. 2008. 1000 genomes project. *Nature Biotechnology*, 26, 256.
- Skol, A. D., Scott, L. J., Abecasis, G. R. & Boehnke, M. 2006. Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. *Nature Genetics*, 38, 209-213.
- Smeets, T. J., Kraan, M. C., Van Loon, M. E. & Tak, P. P. 2003. Tumor necrosis factor α blockade reduces the synovial cell infiltrate early after initiation of treatment, but apparently not by induction of apoptosis in synovial tissue. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 48, 2155-2162.
- Sokhi, U. K., Liber, M. P., Frye, L., Park, S., Kang, K., et al. 2018. Dissection and function of autoimmunity-associated TNFAIP3 (A20) gene enhancers in humanized mouse models. *Nature Communications*, 9, 658.
- Sokol, H., Conway, K. L., Zhang, M., Choi, M., Morin, B., et al. 2013. Card9 Mediates Intestinal Epithelial Cell Restitution, T-Helper 17 Responses, and Control of Bacterial Infection in Mice. *Gastroenterology*, 145, 591-601.e3.
- Sonel, B., Tutkak, H. & Düzgün, N. 2002. Serum levels of IL-1beta, TNF-alpha, IL-8, and acute phase proteins in seronegative spondyloarthropathies. *Joint Bone Spine*, 69, 463-467.
- Song, I.-H., Sieper, J. & Rudwaleit, M. 2007. Diagnosing early ankylosing spondylitis. *Current rheumatology reports*, 9, 367-374.
- Stelzer, G., Rosen, N., Plaschkes, I., Zimmerman, S., Twik, M., et al. 2016. The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. *Current Protocols in Bioinformatics*, 54, 1.30.1-1.30.33.
- Sternes, P. R., Brett, L., Phipps, J., Ciccio, F., Kenna, T., et al. 2022. Distinctive gut microbiomes of ankylosing spondylitis and inflammatory bowel disease patients suggest differing roles in pathogenesis and correlate with disease activity. *Arthritis Research & Therapy*, 24, 163.

- Stolwijk, C., Van Tubergen, A., Castillo-Ortiz, J. D. & Boonen, A. 2015. Prevalence of extra-articular manifestations in patients with ankylosing spondylitis: a systematic review and meta-analysis. *Annals of the Rheumatic Diseases*, 74, 65.
- Strange, A., Capon, F., Spencer, C. C. A., Knight, J., Weale, M. E., et al. 2010. A genome-wide association study identifies new psoriasis susceptibility loci and an interaction between HLA-C and ERAP1. *Nature Genetics*, 42, 985-990.
- Su, W., Du, L., Liu, S., Deng, J., Cao, Q., et al. 2018. ERAP1/ERAP2 and RUNX3 polymorphisms are not associated with ankylosing spondylitis susceptibility in Chinese Han. *Clinical & Experimental Immunology*, 193, 95-102.
- Sudhakar, P., Andrighetti, T., Verstockt, S., Caenepeel, C., Ferrante, M., et al. 2022. Integrated analysis of microbe-host interactions in Crohn's disease reveals potential mechanisms of microbial proteins on host gene expression. *IScience*, 25.
- Sveinbjornsson, G., Albrechtsen, A., Zink, F., Gudjonsson, S. A., Oddson, A., et al. 2016. Weighting sequence variants based on their annotation increases power of whole-genome association studies. *Nature Genetics*, 48, 314-317.
- Sykes, M. P., Hamilton, L., Jones, C. & Gaffney, K. 2018. Prevalence of axial spondyloarthritis in patients with acute anterior uveitis: a cross-sectional study utilising MRI. *RMD Open*, 4, e000553.
- Team, R. C. 2013. R: A language and environment for statistical computing.
- Tervaniemi, M. H., Siitonen, H. A., Söderhäll, C., Minhas, G., Vuola, J., et al. 2012. Centrosomal localization of the psoriasis candidate gene product, CCHCR1, supports a role in cytoskeletal organization. *PloS one*, 7, e49920.
- Tian, F., Chen, Z. & Xu, T. 2019. Efficacy and safety of tofacitinib for the treatment of chronic plaque psoriasis: a systematic review and meta-analysis. *J Int Med Res*, 47, 2342-2350.
- Torgutalp, M., Rios Rodriguez, V., Proft, F., Protopopov, M., Verba, M., et al. 2022. Treatment With Tumor Necrosis Factor Inhibitors Is Associated With a Time-Shifted Retardation of Radiographic Sacroiliitis Progression in Patients With Axial Spondyloarthritis: 10-Year Results From the German Spondyloarthritis Inception Cohort. *Arthritis & Rheumatology*, 74, 1515-1523.
- Toussiot, E., Ricard-Blum, S., Dumoulin, G., Cedoz, J. P. & Wendling, D. 1999. Relationship between urinary pyridinium cross-links, disease activity and disease subsets of ankylosing spondylitis. *Rheumatology*, 38, 21-27.
- Tsoi, L. C., Spain, S. L., Knight, J., Ellinghaus, E., Stuart, P. E., et al. 2012. Identification of 15 new psoriasis susceptibility loci highlights the role of innate immunity. *Nature Genetics*, 44, 1341-1348.

- Tsui, F. W. L., Haroon, N., Reveille, J. D., Rahman, P., Chiu, B., et al. 2010. Association of an ERAP1 ERAP2; haplotype with familial ankylosing spondylitis. *Annals of the Rheumatic Diseases*, 69, 733.
- Turner, S. D. 2014. qqman: an R package for visualizing GWAS results using QQ and manhattan plots. *Biorxiv*, 005165.
- Uderhardt, S., Diarra, D., Katzenbeisser, J., David, J.-P., Zwerina, J., et al. 2010. Blockade of Dickkopf (DKK)-1 induces fusion of sacroiliac joints. *Annals of the rheumatic diseases*, 69, 592-597.
- Ueno, A., Jeffery, L., Kobayashi, T., Hibi, T., Ghosh, S. & Jijon, H. 2018. Th17 plasticity and its relevance to inflammatory bowel disease. *Journal of Autoimmunity*, 87, 38-49.
- Van Bentum, R. E., Heslinga, S. C., Nurmohamed, M. T., Gerards, A. H., Griep, E. N., et al. 2019. Reduced Occurrence Rate of Acute Anterior Uveitis in Ankylosing Spondylitis Treated with Golimumab — The GO-EASY Study. *The Journal of Rheumatology*, 46, 153-159.
- Van Berkum, N. L., Lieberman-Aiden, E., Williams, L., Imakaev, M., Gnirke, A., et al. 2010. Hi-C: a method to study the three-dimensional architecture of genomes. *JoVE (Journal of Visualized Experiments)*, e1869.
- Van Der Linden, S., Valkenburg, H. A. & Cats, A. 1984. Evaluation of diagnostic criteria for ankylosing spondylitis. A proposal for modification of the New York criteria. *Arthritis Rheum*, 27, 361-8.
- Vargas-Alarcón, G., Londoño, J. D., Hernández-Pacheco, G., Pacheco-Tena, C., Castillo, E., et al. 2002. Effect of HLA-B and HLA-DR genes on susceptibility to and severity of spondyloarthropathies in Mexican patients. *Annals of the Rheumatic Diseases*, 61, 714.
- Vecellio, M., Chen, L., Cohen, C. J., Cortes, A., Li, Y., et al. 2021. Functional Genomic Analysis of a RUNX3 Polymorphism Associated With Ankylosing Spondylitis. *Arthritis & Rheumatology*, 73, 980-990.
- Vehik, K., Ajami, N. J., Hadley, D., Petrosino, J. F. & Burkhardt, B. R. 2013. The Changing Landscape of Type 1 Diabetes: Recent Developments and Future Frontiers. *Current Diabetes Reports*, 13, 642-650.
- Verma, S. S., De Andrade, M., Tromp, G., Kuivaniemi, H., Pugh, E., et al. 2014. Imputation and quality control steps for combining multiple genome-wide datasets. *Frontiers in Genetics*, 5.
- Villaverde-García, V., Cobo-Ibáñez, T., Candelas-Rodríguez, G., Seoane-Mato, D., Campo-Fontecha, P. D. D., et al. 2017. The effect of smoking on clinical and structural damage in patients with axial spondyloarthritis: A systematic literature review. *Seminars in Arthritis and Rheumatism*, 46, 569-583.

- Wakefield, D., Chang, J. H., Amjadi, S., Maconochie, Z., El-Asrar, A. A. & McCluskey, P. 2011. What is new HLA-B27 acute anterior uveitis? *Ocular immunology and inflammation*, 19, 139-144.
- Wang, L., Wang, Y., Zhang, P., Song, C., Pan, F., et al. 2022a. Gut microbiota changes in patients with spondyloarthritis: A systematic review. *Seminars in Arthritis and Rheumatism*, 52, 151925.
- Wang, Q., Wang, F., Zhong, W., Ling, H., Wang, J., et al. 2019. RNA-binding protein RBM6 as a tumor suppressor gene represses the growth and progression in laryngocarcinoma. *Gene*, 697, 26-34.
- Wang, R., Gabriel, S. E. & Ward, M. M. 2016. Progression of Nonradiographic Axial Spondyloarthritis to Ankylosing Spondylitis: A Population-Based Cohort Study. *Arthritis Rheumatol*, 68, 1415-21.
- Wang, W.-L., Ouyang, C., Graham, N. M., Zhang, Y., Cassady, K., et al. 2022b. microRNA-142 guards against autoimmunity by controlling Treg cell homeostasis and function. *PLoS biology*, 20, e3001552.
- Wang, X., Ma, J., Ma, J., Wen, Y., Meng, L., et al. 2017. Bioinformatics analysis of genetic variants of endoplasmic reticulum aminopeptidase 1 in ankylosing spondylitis. *Mol Med Rep*, 16, 6532-6543.
- Ward, L. D. & Kellis, M. 2012. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Research*, 40, D930-D934.
- Watanabe, K., Taskesen, E., Van Bochoven, A. & Posthuma, D. 2017. Functional mapping and annotation of genetic associations with FUMA. *Nature Communications*, 8, 1826.
- Weber, A., Wasiliew, P. & Kracht, M. 2010. Interleukin-1 β (IL-1 β) Processing Pathway. *Science Signaling*, 3, cm2-cm2.
- Wendling, D. & Claudepierre, P. 2013. New bone formation in axial spondyloarthritis. *Joint Bone Spine*, 80, 454-458.
- Went, M., Sud, A., Försti, A., Halvarsson, B.-M., Weinhold, N., et al. 2018. Identification of multiple risk loci and regulatory mechanisms influencing susceptibility to multiple myeloma. *Nature Communications*, 9, 3707.
- Westra, H. J., Peters, M. J., Esko, T., Yaghootkar, H., Schurmann, C., et al. 2013. Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat Genet*, 45, 1238-1243.
- Willer, C. J., Li, Y. & Abecasis, G. R. 2010. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*, 26, 2190-1.

- Williams, D. B., Barber, B. H., Flavell, R. A. & Allen, H. 1989. Role of beta 2-microglobulin in the intracellular transport and surface expression of murine class I histocompatibility molecules. *The Journal of Immunology*, 142, 2796-2806.
- Wishart, D. S., Knox, C., Guo, A. C., Shrivastava, S., Hassanali, M., et al. 2006. DrugBank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic Acids Res*, 34, D668-72.
- Wordsworth, B. P., Cohen, C. J., Davidson, C. & Vecellio, M. 2021. Perspectives on the Genetic Associations of Ankylosing Spondylitis. *Frontiers in Immunology*, 12.
- Wray, N. R., Lee, S. H., Mehta, D., Vinkhuyzen, A. a. E., Dudbridge, F. & Middeldorp, C. M. 2014. Research Review: Polygenic methods and their application to psychiatric traits. *Journal of Child Psychology and Psychiatry*, 55, 1068-1087.
- Xu, H., Yu, H., Liu, L., Wu, H., Zhang, C., et al. 2021. Integrative single-cell RNA-seq and ATAC-seq analysis of peripheral mononuclear cells in patients with ankylosing spondylitis. *Frontiers in Immunology*, 12, 760381.
- Yang, M., Xu, M., Pan, X., Hu, Z., Li, Q., et al. 2013a. Epidemiological comparison of clinical manifestations according to HLA-B*27 carrier status of Chinese Ankylosing Spondylitis patients. *Tissue Antigens*, 82, 338-343.
- Yang, W., Tang, H., Zhang, Y., Tang, X., Zhang, J., et al. 2013b. Meta-analysis followed by replication identifies loci in or near CDKN1B, TET3, CD80, DRAM1, and ARID5B as associated with systemic lupus erythematosus in Asians. *Am J Hum Genet*, 92, 41-51.
- Yao, C., Hirata, T., Soontrapa, K., Ma, X., Takemori, H. & Narumiya, S. 2013. Prostaglandin E2 promotes Th1 differentiation via synergistic amplification of IL-12 signalling by cAMP and PI3-kinase. *Nature Communications*, 4, 1685.
- Yasuda, K., Takeuchi, Y. & Hirota, K. The pathogenicity of Th17 cells in autoimmune diseases. *Seminars in immunopathology*, 2019. Springer, 283-297.
- Yenmis, G., Oner, T., Cam, C., Koc, A., Kucuk, O. S., et al. 2015. Association of NFKB1 and NFKBIA Polymorphisms in Relation to Susceptibility of Behçet's Disease. *Scandinavian Journal of Immunology*, 81, 81-86.
- Yi, J. S., Rosa-Bray, M., Staats, J., Zakrofsky, P., Chan, C., et al. 2019. Establishment of normative ranges of the healthy human immune system with comprehensive polychromatic flow cytometry profiling. *PLoS One*, 14, e0225512.
- Yildirim, K., Erdal, A., Karatay, S., Melikoglu, M. A., Ugur, M. & Senel, K. 2004. Relationship between some acute phase reactants and the Bath Ankylosing Spondylitis Disease Activity Index in patients with Ankylosing Spondylitis. *Southern Medical Journal*.
- Yokoyama, U., Iwatsubo, K., Umemura, M., Fujita, T. & Ishikawa, Y. 2013. The Prostanoid EP4 Receptor and Its Signaling Pathway. *Pharmacological Reviews*, 65, 1010-1052.

- Yu, R. Y., Brazaitis, J. & Gallagher, G. 2015. The Human IL-23 Receptor rs11209026 A Allele Promotes the Expression of a Soluble IL-23R–Encoding mRNA Species. *The Journal of Immunology*, 194, 1062-1068.
- Zabarovsky, E. R., Lerman, M. I. & Minna, J. D. 2002. Tumor suppressor genes on chromosome 3p involved in the pathogenesis of lung and other cancers. *Oncogene*, 21, 6915-6935.
- Zeboulon, N., Dougados, M. & Gossec, L. 2008. Prevalence and characteristics of uveitis in the spondyloarthropathies: a systematic literature review. *Annals of the Rheumatic Diseases*, 67, 955-959.
- Zhang, H., Jiang, H.-L. & Dai, S.-M. 2022a. No Significant Effects of IL-23 on Initiating and Perpetuating the Axial Spondyloarthritis: The Reasons for the Failure of IL-23 Inhibitors. *Frontiers in Immunology*, 13, 818413.
- Zhang, H., Liu, M., Du, G., Yu, B., Ma, X., et al. 2022b. Immune checkpoints related-LncRNAs can identify different subtypes of lung cancer and predict immunotherapy and prognosis. *Journal of Cancer Research and Clinical Oncology*, 148, 1597-1612.
- Zhang, W., Cai, Z., Liang, D., Han, J., Wu, P., et al. 2023. Immune Cell-Related Genes in Juvenile Idiopathic Arthritis Identified Using Transcriptomic and Single-Cell Sequencing Data. *International Journal of Molecular Sciences* [Online], 24.
- Zhao, S., Jones, G. T., Macfarlane, G. J., Hughes, D. M., Dean, L. E., et al. 2019. Associations between smoking and extra-axial manifestations and disease severity in axial spondyloarthritis: results from the BSR Biologics Register for Ankylosing Spondylitis (BSRBR-AS). *Rheumatology*, 58, 811-819.
- Zheng, X., Shen, J., Cox, C., Wakefield, J. C., Ehm, M. G., et al. 2014. HIBAG—HLA genotype imputation with attribute bagging. *The Pharmacogenomics Journal*, 14, 192-200.
- Zhu, W., He, X., Cheng, K., Zhang, L., Chen, D., et al. 2019. Ankylosing spondylitis: etiology, pathogenesis, and treatments. *Bone Res*, 7, 22.
- Zhu, Y., Jiang, H. G., Chen, Z. H., Lu, B. H., Li, J. & Shen, X. N. 2020. Genetic association between IL23R rs11209026 and rs10889677 polymorphisms and risk of Crohn's disease and ulcerative colitis: evidence from 41 studies. *Inflammation Research*, 69, 87-103.
- Zhu, Z., Zhang, F., Hu, H., Bakshi, A., Robinson, M. R., et al. 2016. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nature Genetics*, 48, 481-487.
- Zochling, J. 2011. Measures of symptoms and disease status in ankylosing spondylitis: Ankylosing Spondylitis Disease Activity Score (ASDAS), Ankylosing Spondylitis Quality of Life Scale (ASQoL), Bath Ankylosing Spondylitis Disease Activity Index

(BASDAI), Bath Ankylosing Spondylitis Functional Index (BASFI), Bath Ankylosing Spondylitis Global Score (BAS-G), Bath Ankylosing Spondylitis Metrology Index (BASMI), Dougados Functional Index (DFI), and Health Assessment Questionnaire for the Spondylarthropathies (HAQ-S). *Arthritis Care & Research*, 63, S47-S58.

Zochling, J. & Braun, J. 2005. Assessment of ankylosing spondylitis. *Clinical and experimental rheumatology*, 23, S133.